



Алтынбек Шәріпбай

НЕЙРОНДЫҚ ЖЕЛІЛЕР

Қазақстан Республикасы Білім және ғылым Министрлігі
Л.Н.Гумилев атындағы Еуразия ұлттық университеті

Алтынбек ШӘРІПБАЙ

НЕЙРОНДЫҚ ЖЕЛІЛЕР
(Оқулық)

Нұр-Сұлтан, 2021

УДК 004.4(075.8)
ББК 32.973.26-018.1я73
Ш-25

**Қазақстан Республикасы Білім және ғылым Министрлігі,
Л.Н.Гумилев атындағы Еуразия ұлттық университетінің Ғылыми кеңесі
ҰСЫНҒАН**

Пікір жазғандар:

Адамов А. А., Л.Н.Гумилев атындағы ЕҰУ Математикалық және компьютерлік моделдеу кафедрасының меңгерушісі, т.ғ.д., профессор.

Казиев Ғ.З., Алматы энергетика және байланыс университетінің Ақпараттық жүйелер және киберқауіпсіздік кафедрасының професоры, т.ғ.д., профессор.

Өскенбаева Р.К., Халықаралық ақпараттық технологиялар университетінің ректоры, т.ғ.д., профессор.

Шәріпбай А.Ә. Нейрондық желілер – Оқулық, 2-ші басылым, Нұр-Сұлтан, 2021, –391 б.

Мемлекеттік сыйлығының лауреаты, Халықаралық ақпараттандыру академиясының және Қазақстан педагогикалық ғылымдар академиясының академигі А.Ә. Шәріпбайдың бұл оқулығында жасанды нейрондық желілердің анықтамалары, типтері, олардың құрылымдары, жұмыс істеу принциптері және қасиеттері қарастырылады, көптеген мысалдар көрсетіледі, тапсырмалар беріледі және сұрақтар қойылады.

Оқулық ақпараттық және коммуникативтік технологиялар бағытындағы “Жасанды интеллект технологиялары” атты үш деңгейлі оқыту бағдарамасын қамтамасыз етеді. Оны көрсетілген бағыттағы “Информатика”, “Есептеу техникасы және программалық қамтым”, “Ақпараттық жүйелер”, “Автоматтандыру және басқару”, “Ақпараттық қауіпсіздік жүйесі”, “Компьютерлік лингвистика” және басқа оқыту бағдарламаларында оқып білім алушыларға арналған. Сонымен қатар, ол оқытушыларға, ғалымдарға және нейрондық желілерге қызыққан барлық азаматтарға да пайдалы.

1404000000
Ш-34
ISBN 978-301-326-222-2

ӘОЖ 004.4(075.8)
ББК 32.973.26–018.1я73
© А.Ә. Шәріпбай, 2021

БІРІНШІ БАСЫЛЫМҒА АЛҒЫ СӨЗ

Қазіргі таңда нейрондық желілер теориясы ғылым ретінде қалыптасты және ол өзіндік аппаратқа және зерттеу пәніне ие. Нейрондық желілер информатика, математика, физика, схемотехника тәріздес ғылымның әр түрлі салалары бойынша сұрақтардың кең спектрін қамтиды және әртүрлі есептердің үлкен тобын шешуге мүмкіндік береді.

Нейрондық желілерге деген айтулы қызығушылық 1982 жылы Хопфилд (Hopfield J.J.) жұмыстарында нейрондардағы есептер физикада әзірленген реттелмеген жүйелердің бірқатар моделдерін жалпылауға әкелетіндігін көрсеткеннен кейін басталды. Хопфилд желісінің жұмысы Хеббтің оқыту ережесімен анықталатын стационарлық күйлердің біріне екілік кодтар матрицасының бастапқы «спиндік портретінің» релаксациясынан тұрады.

1986 жылы Румельхарттың (Rumelhart D.), Хинтонның (Hinton G.) және Вильямстың (Williams R) жұмысы пайда болды, ол нейроинформатиканың дамуына ұзақ уақыт кедергі болған және 1940-1950 жылдарда есептердің кең тобы үшін әмбебаптылығы классик ғалымдармен орнатылған—иерархиялық қабатталған нейрондық желілер қалай оқытылатындығы туралы сұрақтың жауабын қамтыды. Келесі жылдары Хинтон ұсынған қателерді кері тарату алгоритмі көптеген түрлендіру мен жетілдірулерге ұшырады.

Жасанды нейрондық желі көпөлшемді функцияларды аппроксимациялау, ассоциативті жадты құру, үдерістерді болжамдау мен диагностикалау, деректер жиымында ассоциация бойынша іздеу және заңнамаларды іздеу, бейімді басқару және статистикалық талдау, үлгілерді айқындау мен тану және т.б үшін модел, жабдық және құрал ретінде қолданылады.

Нейрондық желілер әртүрлі интеллектуалды жүйелер құру кезінде маңызды рөл атқарады және бейнелерді өңдеу, роботтар мен үздіксіз өндірістерді басқару, сөйлеуді тану мен тудыру,

адамдардың ауруларын және техникалық құралдардың ақауларын диагностикалау, валюта курстары мен ұтыстарды болжау үшін, сонымен қатар көптеген басқа да динамикалық есептерді шешу үшін қолданылады.

Ұсынылған оқулық автордың Л.Н.Гумилев атындағы ақпараттық технологиялар факультетінің Информатика, Ақпараттық жүйелер, Есептеу техникасы және программалық қамтым мамандықтары бойынша студенттер, магистранттар, докторанттар үшін әр жылдары оқылған дәрістер негізінде даярланды. Оған осы сала бойынша баспадан шыққан оқулықтың жоқтығы себеп болды. Сондықтан да, оқыту үрдісін толық және сапалы қамту үшін авторға мағынасы бір ұғым үшін әртүрлі термин, анықтама мен таңбалауы бар әртүрлі деректер көздері мен интернет ресурстарын пайдалануға тура келді. Осы аталған жағдайлар әрбір ұғым үшін бір терминді, анықтаманы және таңбалауды қолдану арқылы осы оқулықты жазуға итермеледі. Автор ұсынып отырған оқулық қазіргі кезде «Нейрондық желілер» пәні бойынша оқу материалдары мен оқу материалдарының жеткіліксіздігін түзетеді және бірыңғай терминдерді, анықтамаларды және таңбалауларды қалыптастыруға себебін тигізеді деп сенеді.

Оқу материалының мазмұндық тәсілі оқырманға нейрондық желілер теориясындағы іргелі фактілер туралы деректерді беру, зерттелініп отырған саладағы тұжырымдардың дәлелдеу әдістерін беру, сонымен қатар, оқытылған материалдарды бекіту үшін оған мысалдар мен жаттығулар қорын беру болады.

Оқу материалы кредиттік технологияға сәйкес ұйымдастырылған, ол үш бөлімнен тұрады: *бөлім (модул), тарау (блок), параграф (дәріс)*. Бұл төрт деңгейде білімді беру және бақылау жүргізуге мүмкіндік береді: дәріс деңгейінде – *ағымдық бақылау*, блок деңгейінде – *аралық бақылау*, модул деңгейінде – *белестік бақылау*, оқулық деңгейінде – *қорытынды бақылау*. Білімді бағалау жоғарыда көрсетілген деңгейлерге сәйкес құрылған тесттердің көмегімен жүзеге асырылады.

Оқулық, жоғарыда айтылғандай, үш модульден тұрады. Бірінші модуль нейрон жүйелерінің математикалық негіздерін талқылауға арналған үш блоктан тұрады, оларда жиындар теориясының элементтері, функциялар теориясының элементтері, векторлық есептеудің элементтері беріледі. Екінші модуль нейрондық желілер теориясының сұрақтарын қамтитын үш блоктан тұрады, оларда нейрондық желілердің негізгі түсініктері мен түрлері, белсендіру функциялары және нейрондардың моделдері қарастырылады. Үшінші модуль нейрон жүйелерін оқытуға арналған бір блоктан тұрады, онда әртүрлі парадигмалар мен оқу тәсілдері суреттеледі.

Қолданылған деректер көздерінде және интернет ресурстарында оқулықтағы оқу материалдарын даярлау кезінде қолданылған негізгі көздер көрсетіледі. Олардың арасында кеңінен таралған монографиялар, оқулықтар, оқу құралдары және т.б. бар. Оларға осыған байланысты оқулық мәтінінде әр тақырып басында оқу материалдарына қатысты деректер көздері мен интернет ресурстарына сілтемелер көрсетіліп тұрады.

Оқулықты Информатика, Ақпараттық қауіпсіздік жүйелері, Есептеу техникасы және программалық қамтым, Ақпараттық жүйелер, Автоматтандыру және басқару, Математикалық және компьютерлік моделдеу мамандықтары бойынша студенттер, магистранттар, докторанттар, оқытушылар мен ғалымдар, сонымен қатар жасанды нейрондық желілермен қызығатын және осы сала бойынша білімдерін кеңейткісі келетіндер қолдана алады.

Автор:

Алтынбек Амирович ШАРИПБАЙ,
техника ғылымдарының докторы, профессор,
Қазақстан Республикасының Мемлекеттік сыйлығының
лауреаты

ЕКІШІ БАСЫЛЫМҒА АЛҒЫ СӨЗ

Осы оқулықты қайта басуға оның көкейкестігі негіз болды, себебі оған сұраныс басылмақ түгіл, өсуде. Оқырмандардың бұл оқулықты кітапханаларда таба алатынына қарамастан, олар оны дүкендерден сатып ала алмайды, себебі оның 1-ші басылымындағы тиражы аз болды. Бұл жағдай авторды осы оқулықты қайта басуға ой салды. Сонымен қатар, заманауи қоғамда *жасанды интеллект*тің пайланылмайтын саласы жоқтығын және оның «ақылы – миы» нейрондық желілерден құралып отырғандығын ескерсек, оқулықтың екінші басылымын шығару өте қажет екендігі күмән тудырмайды.

Осы басылымда жоғарыда көрсетілген 1-ші басылымға жазылған алғы сөзді өзгертпейтін бодым. Бірақ соған қарамастан, мен оқулыққа қолдану аялары күннен күнге өсіп келе жатқан интеллектуалды программалық жүйелер мен аппараттық жабдықтарды даярлауға қажет заманауи нейрондық желілердің теориялық және практикалық проблемаларын талдауды қостым.

Әрбір параграф алдында оқу материалын құрастыруда пайдаланылған түпнұсқалар мен интернет ресурстарына сілтеме берілген. Сондықтан олар мазмұнның мәтінінде кездеспейді.

Оқулық Ақпараттық және коммуникациялық технологиялары (АКТ) бағытындағы "Жасанды интеллект технологиялары" атты үш деңгейлі оқыту бағдарламасына арналған. Сонымен қатар, бұл оқулықты АКТ-бағытының "Информатика", "Ақпараттық жүйелер", "Есептеу техникасы мен программалық қамтым" және басқа оқыту бағдарламаларында да пайдалануға болады.

Автор оқулықтың осы 2-ші басылымының алғашқы үлгісін елеулі жақсартуға бағалы ескертпелер мен ұсыныстар берген пікір берушілер Ә.Ә. Адамовқа, Ғ.З.Қазиевке, Р.К. Өскенбаеваға өз алғысын білдіреді.

Оқулықтың цифрленген үлгісі <http://www.e-zerde.kz/zerde> сайтында орналастырылған

Автор:

ШӘРІПБАЙ Алтынбек Әмірұлы

І. НЕЙРОНДЫҚ ЖЕЛІЛЕРДІҢ МАТЕМАТИКАЛЫҚ НЕГІЗДЕРІ

І. 1. Жиындар теориясының элементтері

І. 1.1. Жиындар және оларға амалдар

Бұл тақырыпта жиындар ұғымы және оларға қолданылатын амалдар мен олардың қасиеттері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалын даярлау кезінде [1] дерек көзі пайдаланылды.

І.1.1.1-анықтама. *Жиындар* деп кейбір типі бойынша ажыратылатын элементтердің жиынтығын айтады.

Жиындар үлкен (бас) латын әріптермен, ал оның элементтері – кіші (шағын) латын әріптерімен немесе араб цифрларымен таңбаланады. Мысалы, натурал сандардың жиындары N әріпімен, оның элементтері $-1, 2, 3, \dots$ цифрларымен таңбаланады.

Жиындар екі әдіспен беріледі: *барлық элементтерді көрсету* немесе *элементтердің қасиеттерін суреттеу* арқылы. Бұл ретте, оның элементтері $\{ \text{және} \}$ фигуралы жақшада көрсетіледі.

І.1.1.2-анықтама. A жиынының қуаты оның элементтер санына тең және $|A|$ арқылы таңбаланады.

І.1.1.1-мысалдар.

1. Барлық теріс емес бүтін сандар натурал сандар жиыны болады.

2. $D = \{0, 1\}$, D жиынының элементтері 0, 1 шамалары ғана.

3. $X = \{x: x > 0\}$, X жиынының элементтері x -тің оң мәндері ғана.

4. Егер $A = \{a, b, c, d, e\}$, онда $|A| = 5$.

$a \in A$ ($a \notin A$) – a элементінің A жиынына тиісті (тиісті емес).

Барлық жиындар ішінен екі жиын ерекшеленеді:

1. $\emptyset$ – *бос жиын*, бірде–бір элементі жоқ.

2. U – *универсалды жиын (универсум)*, қарастырылып отырған типтің (пәндік облыстың) барлық элементтерінен тұрады.

Теорияға байланысты универсум осы теорияда қарастырылатын барлық объектілерді элементтер ретінде қамтиды.

Мысалы, универсум болып мыналар табылады:

1) сандар теориясында – барлық бүтін сандар жиыны;

- 2) тілдер теориясында – әліпбидегі барлық сөздер жиыны;
- 3) геометрияда – кеңістіктегі барлық нүктелер жиыны.

1.1.1.2-анықтама. A және B екі жиыны берілсін, онда:

- 1) A және B жиындар элементтерінің *бірігуі*, яғни

$$A \cup B = \{x: x \in A \vee x \in B\}.$$

- 2) A және B жиындар элементтерінің *қиылысуы*, яғни

$$A \cap B = \{x: x \in A \& x \in B\}.$$

- 3) *Толықтыру* U универсум элементтерінен тұрады және A жиынының элементтерін қоспайды, яғни

$$\overline{A} = \{x: x \in U \& x \notin A\}.$$

- 4) *Айырым* A жиынының элементтерінен тұрады, B жиынының элементтерін қоспайды, яғни

$$A \setminus B = \{x: x \in A \& x \notin B\}.$$

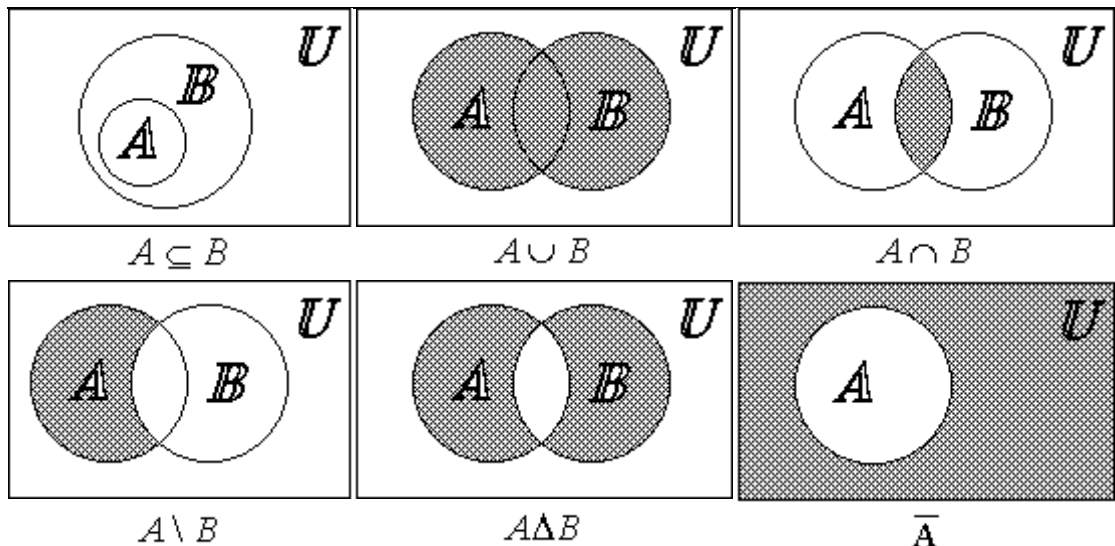
- 5) *Симметриялы айырым* тек қана A жиынының элементтерінен тұрады немесе тек B жиынының элементтерінен тұрады, яғни

$$A \Delta B = \{x: (x \in A \& x \notin B) \vee (x \in B \& x \notin A)\}$$

- 6) *Декартты (тура) көбейтінді* A және B элементтерінің барлық реттелген жұптарынан тұрады, яғни

$$A \times B = \{(a, b): a \in A \& b \in B\}$$

(1)–(3) амалдар Эйлер-Венна диаграммасымен бейнеленген, мұнда U универсум тіктөртбұрышпен, ал A және B жиындары – шеңбер ретінде көрсетілген, нәтижені белгілеу үшін штрихтау қолданылады (1.1.1-сурет).



I.1.1-сурет. Эйлер-Венна диаграммалары

Мұнда A және B жиындары U жиынының ішжиындары болып тұр және олар $A \subseteq U$ және $B \subseteq U$ түрінде жазылады (I.2.2 анықтаманы қараңыз).

(1) – (3) амалдарын тек екі жиында ғана емес, $A_1, A_2, \dots, A_n$ арқылы берілген n жиында да анықтауға болады, мұнда $n \in \mathbb{N}$ & $n > 2$.

$A_1, A_2, \dots, A_n$ жиындарының бірігуі келесі түрде көрсетіледі:

$$A_1 \cup A_2 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i.$$

$A_1, A_2, \dots, A_n$ жиындарының қиылысуы келесі түрде көрсетіледі:

$$A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i.$$

$A_1, A_2, \dots, A_n$ жиындарының тікелей көбейтілуі $(a_1, a_2, \dots, a_n)$ түріндегі кортеждер жиыны ретінде анықталады, мұнда элементтер $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$, яғни

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_1 \in A_1 \& a_2 \in A_2 \& \dots \& a_n \in A_n\}.$$

Мұнда, егер $A_1 = A_2 = \dots = A_n = A$ болса, онда $A \times A \times \dots \times A = A^n$ – дәреже болады.

$$A \times A \times \dots \times A = A^n$$

I.1.1.2-мысалдар.

1. $A = \{a, b, c, d, e, f\}, B = \{c, d\}$ болсын, онда:

2. $A \setminus B = \{a, b, e, f\};$

3. $A \Delta B = \{a, b, f\};$

4. $B \times B = \{(c, c), (c, d), (d, c), (d, d)\};$

5. $\bar{A}$ жиыны универсум U қандай болатындығына тәуелді.

Мысалы, егер $U = \{a, b, c, d, e, f, h\}$ болса, онда $\bar{A} = \{h\}$.

Енді жиындарды және оларда анықталған амалдарды кестелік тәсілмен беруді көрсетеміз. Мейлі $U, A \subseteq U$ және $x \in U$ берілсін.

I.1.1.3-анықтама. A жиыны үшін $I_A(x)$ индикаторлық (сипаттамалық) функция деп аталады, ол келесі түрде беріледі:

$$I_A(x) = \begin{cases} 1, & \text{егер } x \in A \\ 0, & \text{егер } x \notin A \end{cases}$$

Осылайша: $I_A: U \rightarrow \{0,1\}$.

$A \subseteq U$ және $B \subseteq U$ үшін келесі қасиеттер орындалады:

$$I_A(x) = I_B(x) \Leftrightarrow A = B;$$

$$I_A(x) \leq I_B(x) \Leftrightarrow A \subseteq B;$$

$$I_{\bar{A}}(x) = 1 - I_A(x);$$

$$I_{A \cup B}(x) = I_A(x) + I_B(x) - I_A(x) \cdot I_B(x);$$

$$I_{A \cap B}(x) = I_A(x) \cdot I_B(x);$$

$$I_{A \setminus B}(x) = I_A(x) - I_A(x) \cdot I_B(x);$$

$$I_{A \setminus B}(x) = I_A(x) - I_A(x) \cdot I_B(x);$$

$$I_{A \Delta B}(x) = I_A(x) + I_B(x) - 2I_A(x) \cdot I_B(x).$$

Индикаторларды *I.1.1*-кестенің көмегімен қойған ыңғайлы.

I.1.1-кесте. Индикаторлар.

$x \in A$	$x \in B$	$x \in A \cup B$	$x \in A \cap B$	$x \in A \setminus B$	$x \notin A$	$x \in A \Delta B$
0	0	0	0	0	1	0
0	1	1	0	0	1	1
1	0	1	0	1	0	1
1	1	1	1	0	0	0

Жиындарға қолданылатын операциялар келесі қасиеттерге ие:

I. Бірігу, қиылысу және айырым.

1) $A \cup \emptyset = A$ – нөлдік қасиет;

2) $A \cup A = A$ – идемпотенттілік;

3) $A \cup B = B$, егер A жиынының элементтері B жиынында бар болса;

4) $A \cup B = B \cup A$ – коммутативтілік;

5) $(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$ – ассоциативтілік;

6) $A \cap \emptyset = \emptyset$ – нөлдік қасиет;

7) $A \cap A = A$ – идемпотенттілік;

- 8) $A \cap B = A$, егер A жиынының элементтері B жиынында болса;
- 9) $A \cap B = B \cap A$ – коммутативтілік;
- 10) $(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$ – ассоциативтілік;
- 11) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ – дистрибутивтілік;
- 12) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ – дистрибутивтілік;
- 13) $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$ – дистрибутивтілік;
- 14) $A \cup \bar{A} = U$ – толықтыру қасиеті;
- 15) $A \cap \bar{A} = \emptyset$ – толықтыру қасиеті;
- 16) $\overline{A \cup B} = \bar{A} \cap \bar{B}$ – де Морган заңы
- 17) $\overline{A \cap B} = \bar{A} \cup \bar{B}$ – де Морган заңы
- 18) $\bar{\bar{A}} = A$ – инволютивтілік;
- 19) $A \setminus \emptyset = A$ – айырым қасиеті;
- 20) $A \setminus A = \emptyset$ – айырым қасиеті;
- 21) $A \setminus B = A \cap (\setminus B)$ – айырым қасиеті;
- 22) $B \setminus A = B \cap (\setminus A)$ – айырым қасиеті;
- 23) $A \cap (B \setminus A) = \emptyset$ – айырым қасиеті;

II. Симметриялы айырым және тура көбейтінді.

- 1) $A \Delta \emptyset = A$ – нөлдік қасиет;
- 2) $A \Delta A = \emptyset$ – идемпотенттілік;
- 3) $A \Delta B = (A \cup B) \setminus (A \cap B)$ – симметриялы айырым қасиеті;
- 4) $A \Delta B = B \Delta A$ – коммутативтілік;
- 5) $(A \Delta B) \Delta C = A \Delta (B \Delta C) = A \Delta B \Delta C$ – ассоциативтілік;
- 6) $(A \cup B) \times C = (A \times C) \cup (B \times C)$ – дистрибутивтілік;
- 7) $A \times (B \cup C) = (A \times B) \cup (A \times C)$ – дистрибутивтілік;
- 8) $(A \cap B) \times C = (A \times C) \cap (B \times C)$ – дистрибутивтілік;
- 9) $A \times (B \cap C) = (A \times B) \cap (A \times C)$ – дистрибутивтілік;
- 10) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$ – дистрибутивтілік;
- 11) $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$ – дистрибутивтілік;

I.1.1-мысал. $A=\{1, 2\}$, $B=\{a, b\}$, $C=\{+, -\}$ берілсін. Сонда дистрибутивтілік $(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$ былай анықталады $(\{1,2\} \cup \{a,b\}) \cup \{+,-\} = \{1,2\} \cup (\{a,b\}) \cup \{+,-\} = \{1,2\} \cup \{a,b\} \cup \{+,-\}$.

I.1.1-тапсырмалар. $A=\{1, 2, 4\}$, $B=\{3, 4, 5, 6\}$ үшін орындау керек:

- 1) $A \cup \bar{A}$;
- 2) $A \Delta B$;
- 3) $(A \cup B) \times C$;
- 4) $A \times (B \cup C)$;
- 5) $(A \setminus B) \times C$;
- 6) $(A \cap B) \times C$;
- 7) $(A \Delta B) \Delta C$.

I.1.1-сұрақтар.

- 1) Жиындарды тура көбейту қалай анықталады?
- 2) Жиындардың симметриялы айырымы қалай анықталады?
- 3) Жиындардың индикаторлы функциясы қалай анықталады?
- 4) Де – Морган заңы дегеніміз не?
- 5) Егер $A=\{1,2,3\}$, $B=\{3,4,5\}$, онда $\{1,2,3,4,5\}$ нені білдіреді?
- 6) Егер $A=\{1,2,4\}$, $B=\{4,3,2\}$, онда $\{2,4\}$ нені білдіреді?
- 7) Егер $A=\{1,2,3,4\}$, $B=\{3,4,5,6\}$, онда $\{2,4\}$ нені білдіреді?

I.1.1-тестілер.

1. Бірнеше типі бойынша ажыратылатын элементтерінің жиынтығы не деп аталады?

- A) жиындар;
- B) құрылымдар;
- C) жиымдар;
- D) қатарлар;
- E) сандар.

2. Бос жиын деп қандай жиынды айтады?

- A) бірде-бір элементі жоқ жиынды;
- B) қарастырылып отырған типтің (пәндік аймақтың) барлық элементтерінен тұратын жиынды;
- C) кез келген элементтерден тұратын жиынды;
- D) тек сандардан тұратын жиынды;

Е) тек әріптерден тұратын жиынды.

3. Универсалды жиын (универсум) деп қандай жиынды айтады?

А) қарастырылып отырған типтің (пәндік аймақтың) барлық элементтерінен тұратын жиынды;

В) бірде-бір элементі жоқ жиынды;

С) кез келген элементтерден тұратын жиынды;

Д) тек сандардан тұратын жиынды;

Е) тек әріптерден тұратын жиынды.

4. A және B элементтерінің бірігуінің жазылуы қалай болады?

А) $A \cup B = \{x | x \in A \vee x \in B\}$;

В) $A \cap B = \{x | x \in A \& x \in B\}$;

С) $\overline{A} = \{x | x \in U \& x \notin A\}$;

Д) $A \setminus B = \{x | x \in A \& x \notin B\}$;

Е) $A \Delta B = \{x | (x \in A \& x \notin B) \vee (x \in B \& x \notin A)\}$.

5. A және B элементтерінің қиылысуы қалай жазылады?

А) $A \cap B = \{x | x \in A \& x \in B\}$;

В) $A \cup B = \{x | x \in A \vee x \in B\}$;

С) $\overline{A} = \{x | x \in U \& x \notin A\}$;

Д) $A \setminus B = \{x | x \in A \& x \notin B\}$;

Е) $A \Delta B = \{x | (x \in A \& x \notin B) \vee (x \in B \& x \notin A)\}$.

1.1.2. Қатынастар және оларды бейнелеу тәсілдері

Бұл тақырыпта қатынастар ұғымы анықталады және оларды ұсыну түрлері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалын даярлау кезінде [1] дерек көзі пайдаланылды.

1.1.2.1-анықтама. Егер бір типті екі A және B жиындары берілсе, онда келесі қатынастарды енгізуге болады:

1) $A = B$: A жиыны B жиынына тең болады, егер A және B бірдей элементтен тұрса, яғни A және B жиындары бір-біріне ішжиын болып табылса;

2) $A \subseteq B$: A жиыны B жиынында бар болады, егер A -ның барлық элементтері B -ға тиісті болса немесе A жиыны B -ға тең болса, ол A жиыны B -ның ішжиыны болып табылатындығын білдіреді;

3) $A \subset B$: A жиыны қатаң B жиынында бар болады, егер A -ның барлық элементтері B -ға тиісті болса және A жиыны B -ға тең болмаса, яғни B -ның кейбір элементтері A -ға тиісті емес, ол A жиыны B -ның жеке ішжиыны болып табылатындығын білдіреді.

Осыған ұқсас қамтитын $A \supseteq B$ және қатаң қамтитын $A \supset B$ қатынастарын анықтауға болады.

Жоғарыдағы $=$, $\subseteq$ және $\subset$ қатынастары $A \times B$ тура көбейтудің ішжиындары болып табылатындығын байқауға болады.

Ендеше, кез келген қатынасты нақты ережемен бөлінетін тура көбейтіндінің қандай да бір ішжиындары деп санауға болады.

1.1.2.1-ескерту. $\emptyset$ бос жиыны кез келген ақырлы жиынның меншікті ішжиыны болады.

1.1.2.1-мысалдар.

1) егер $A=\{a,b,c\}$, $B=\{b,a,c\}$, онда $A=B$;

2) егер $A=\{1,2,3,4\}$, $B=\{3,1,4,2\}$, онда $A \subseteq B$;

3) егер $A=\{1,2,3\}$, $B=\{3,1,4,2\}$, онда $A \subset B$

1.1.2.2-анықтама. $A_1, A_2, \dots, A_n$ – әртүрлі болуы міндетті емес қандай да бір жиындар болсын, $n \geq 1$. Сонда $A_1, A_2, \dots, A_n$ жиындарында n -дік қатынас мынадай ішжиын болады

$$R^n \subseteq A_1 \times A_2 \times \dots \times A_n,$$

мұнда $R^1 - A_1$ –дің унарлы қатынасы, $R^2 - A_1, A_2$ –нің бинарлы қатынасы, $R^3 - A_1, A_2, A_3$ –тің тринарлы қатынасы және т.с.с.

A жиынындағы кез келген *унарлы қатынас* оның қандай да бір ішжиындарының сипаттамалық қасиеті болады. A -ға унарлы барлық қатынастар жиыны A жиынының барлық ішжиындарының жиынымен сәйкес келеді.

I.1.2.2-мысалдар.

1. $R_1^1 = \{n: n \in N \& n < 100\}$ унарлы қатынасы 100-ден кіші натурал сандардың жиынын анықтайды;

2. $R_2^1 = \{n: \forall k \in N (n = 2 * k)\}$ унарлы қатынасы жұп натуралды сандардың жиынын анықтайды;

3. $R_3^1 = \{n: \forall k \in Z (n = 2 * k + 1)\}$ унарлы қатынасы тақ бүтін сандардың жиынын анықтайды.

I.1.2.3-анықтама. Бинарлы қатынастар жиындар жұбында анықталады және үш тәсілдің бірімен ұсынылуы мүмкін:

1) *префикстік жазба* – қатынас белгісі бинарлы қатынасқа қатысушылардың алдына қойылады.

2) *инфикстік жазба* – қатынас белгісі бинарлы қатынасқа қатысушылардың ортасына қойылады.

3) *постфикстік жазба* – қатынас белгісі бинарлы қатынасқа қатысушылардың соңына қойылады.

Элементтер жұбымен бинарлы қатынастар көп жағдайда кестенің көмегімен ұсынылады: жолдар жұптың бірінші элементтеріне, бағандар – жұптың екінші элементтеріне сәйкес келеді, ал жолдар мен бағандардың нақты элементтер арасындағы қатынастардың бар болуы, арнайы таңбамен, мысалы «1» таңбасымен немесе т.б. таңбаланады.

I.1.2.3-мысалдар.

1. Егер $a \in A$ және $b \in B$, R бинарлы қатынас болса, онда:

Rab – префикстік жазба;

aRb – инфикстік жазба;

abR – постфикстік жазба.

2. Егер жиындар $A = \{a_1, \dots, a_r\}$ және $B = \{b_1, \dots, b_s\}$, R бинарлы қатынасында болса, онда оны I.2.2-кестесінің көмегімен көрсетуге болады, мұнда a_i элементтері – жолдармен, b_j элементтері – бағандармен ұсынылады, ал $a_i R b_j$ қатынасы “1” былай таңбаланады:

I.1.2-ресте. Екілік қатынас.

R	b_1	b_2	...	b_{s-1}	b_s
a_1	1		...	1	
a_2		1	...		
...	...	...	...	...	...
a_{r-1}	1		...	1	
a_r			...		1

I.1.2.4-анықтама. R бинарлы қатынасы S жиынында:

- 1) *рефлексивті* деп аталады, егер әрбір $s \in S$ үшін sRs бар болса;
- 2) *транзитивті* деп аталады, егер кез келген $s, t, u \in S$ үшін sRt және tRu -дан sRu шығатын болса;
- 3) *симметриялы* деп аталады, егер кез келген $s, t \in S$ үшін sRt -дан tRs шығатын болса;
- 4) *антисимметриялы* деп аталады, егер aRb және bRa -дан $a=b$ шығатын болса.

I.1.2.4-мысалдар.

1. Сандарда $>$ қатынасы рефлексивті болып табылмайды;
2. Сандарда $=, \geq, >$ қатынастары транзитивті болып саналады;
3. Сандарда $=$ қатынасы симметриялы болып саналады.

I.1.2.5-анықтама. R бинарлы қатынасы рефлексивтілік, транзитивтілік және симметриялық қасиеттеріне ие болса, ондай қатынас *эквивалентті* деп аталады.

S жиынында эквиваленттіліктің әр қатынасына осы жиынды көршілес сыныптарға жеке бөлу сәйкес келеді.

I.1.2.5-мысал. Бинарлық « $=$ » қатынасы кез келген сандық жиында эквивалентті қатынас болады, яғни, кез келген $k, m, n \in N$:

1. Рефлексивтілік: $n = n$.
2. Транзитивтілік: $k < m$ және $m < n$ болғанда $k < n$ алынады.
3. Симметриялылық: $k = m$ болғанда $m = k$ алынады.

I.1.2.6-анықтама. Рефлексивтілік, транзитивтілік және антисимметриялық қасиеттерін қанағаттандыратын қандай да бір S жиынындағы R бинарлы қатынасы *ішінара ретті* қатынас деп аталады.

I.1.2.6-мысалдар. Мына қатынастар ішінара ретті болады:

1. $\subseteq$ қатынасы қандай да бір жиынның ішжиындары үшін;
2. $\supseteq$ қатынасы қандай да бір жиынның ішжиындары үшін;
3. $=$ қатынасы бүтін сандар жиынында;
4. $\leq$ қатынасы натурал сандар жиынына;
5. $\geq$ қатынасы бүтін сандар жиынында.

I.1.2.7-анықтама.

1. *Ішінара реттелген жиын* деп A жиынында анықталған *ішінара ретті* қатынасы бар сол A жиынын айтамыз. Нақты айтқанда, реттелген жиын деп $\langle A, R \rangle$ жұбын айтады, мұнда A – жиын, ал R – A жиынындағы *ішінара ретті* қатынас.

2. *Ішінара реттелген A жиынының a және b элементтері ішінара ретті R -ға қатысты салыстырылатын* деп аталады, егер aRb немесе bRa болса.

3. A жиынындағы *ішінара ретті сызықты реттелу* деп аталады, егер A жиынының кез келген екі a және b элементтері, R ішінара реттелуге қатысты салыстырылатын болса.

4. *Сызықты-реттелген жиын* немесе *тізбе* деп әрбір жұп элементтері салыстырылатын ішінара реттелген жиынды айтады.

5. *Ішінара реттелген A жиынының a және b элементтерін салыстырылмайтын* деп атайды, егер олардың арасында бірде бір *ішінара ретті* қатынас орындалмаса. Салыстырылмайтын

элементтердің бар болу мүмкіндігін «ішінара реттелген жиын» терминінің мағынасы түсіндіреді.

1.1.2.7-мысалдар.

1. $\leq$ қатынасымен N натуралды сандар жиыны ішінара реттелген болады, барлық натуралды сандар $\leq$ қатынасына қатысты салыстырылатын және N тізбе болады.

2. $=$ қатынасымен барлық нақты сандар жиыны сызықты реттелген жиын болады, барлық нақты сандар $=$ қатынасына қатысты салыстырылатын болады.

3. Мейлі A ішінара ретті анықталған $<, =, >$ қатынасы бар $[0,1]$ аралығында анықталған нақты мәнді функциялар жиыны болсын, сонда $f(x) = x$ және $g(x) = 1-x$ элементтері салыстырылмайтын болады.

1.1.2.8-анықтама. Мейлі A – ішінара реттелген жиын, B – оның ішжиыны болсын, яғни $A \supseteq B$. Сонда:

1) A жиынындағы B жиынының *төменгі шегі* (жоғарғы шегі) деп кез келген $b \in B$ үшін $a \leq b$ ($b \leq a$) болатын $a \in A$ элементін айтады;

2) A жиынындағы $a \in A$ элементі *ең кіші* (ең үлкен) деп аталады, егер A -ның төменгі (жоғарғы) шегі a болса.

3) A жиынындағы $a \in A$ элементі *минималды* (максималды) деп аталады, егер $b < a$ ($a < b$) болатын $b \in A$ болмаса.

A жиынының ең кіші (ең үлкен) элементі оның жалғыз минималды (максималды) элементі болады.

1.1.2.8-мысалдар.

1. A жиынының барлық ішжиындарының жиынында ең кіші элемент $\emptyset$ және ең үлкен элемент A -ның өзі болады.

2. N натуралды сандар жиынында ең кіші элемент 1 және ең үлкен элементі жоқ болады.

3. Z барлық бүтін сандар жиынында ең кіші және ең үлкен, минималды және максималды элементі болмайды.

A және B жиынының элементтері *өзара-бір мәнді сәйкес* болады, егер әрбір $a \in A$ элементіне қандай да бір заң бойынша тек жалғыз

$b \in B$ элементі салыстырылса, сондай-ақ, әрбір $b \in B$ элементі тек қана жалғыз $a \in A$ элементімен салыстырылса.

A және B жиындары эквивалентті (тең қуатты) болады, егер олардың элементтері арасында өзара-бір мәнді сәйкестік орнатылатын болса.

1.1.2.2-ескерту. Бинарлы қатынастар үштікпен көрсетіле алады $\langle R, A, B \rangle$, мұнда $R \subseteq A \times B$ – қатынас графигі және ол $(a, b) \in R$ немесе aRb болып жазылады. Сонда мыналарды анықтауға болады:

Анықталу аймағы: $\text{Dom } R = \{x \in A: \exists y \in B (x, y) \in R\}$;

Мәндер аймағы: $\text{Run } R = \{y \in B: \exists x \in A (x, y) \in R\}$;

Кері қатынас: $R^{-1} = \{(y, x) \in B \times A: (x, y) \in R\}$;

Қатынас композициясы: $R \subseteq A \times B, S \subseteq B \times C$,

$R \cdot S = \{(x, z) \in A \times C: \exists y \in B [(x, y) \in R \& (y, z) \in S]\}$.

1.1.2.9-анықтама. $f \subseteq X \times Y$ бинарлы қатынасы X -мен Y -ке функция деп аталады, егер $\text{Dom } R = X$ және $(x, y) \in f, (x, z) \in f \Rightarrow y = z$.

Функция $f: X \rightarrow Y$:

1) *сюръективті* деп аталады, егер кез келген $y \in Y$ үшін $y = f(x)$ болатын $x \in X$ болса, яғни $\forall y \in Y \exists x \in X (y = f(x))$;

2) *инъективті* деп аталады, кез келген $x_1, x_2 \in X$ үшін, $x_1 \neq x_2$ одан $f(x_1) \neq f(x_2)$ шығады, яғни $\forall x_1 \in X \forall x_2 \in X (x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2))$;

3) *биективті* деп аталады, егер ол сюръективті және инъективті болса.

Кез келген бинарлы функцияны *тренарлы* қатынаспен байланыстыруға болады, мысалы, егер $f(x, y)$ бинарлы функциясы берілген болса, онда оны $z = f(x, y)$ болатындай $R^3(x, y, z)$ тренарлы қатынасымен байланыстыруға болады.

1.1.2.9-мысалдар. Мейлі $x, y, z \in N$ – натурал сандар болсын және $f(x, y) = z$ бинарлы функциясы берілсін. Сонда:

1) егер x 3-ке, y 5-ке, z 8-ге тең, ал f + қосу амалы болса, онда $f(x, y) = z$ жазбасының орнына $3+5=8$ деп жазуға болады және онымен $8 = 3+5$ жазбасын беретін $Add(3, 5, 8)$ тренарлы қатынасын байланыстыруға болады.

2) егер x 8-ге, y 2-ге, z 4-ке тең, ал f «:» бөлу амалы болса, онда $f(x, y) = z$ жазбасының орнына $8:2 = 4$ деп жазуға болады және

онымен $4 = 8:2$ жазбасын беретін $Dev(8, 2, 4)$ тренарлы қатынасын байланыстырады.

Кейбір $x, y \in N$ үшін бөлу амалын орындау нәтижесі бүтін сан болмайтындығы айқын және олар үшін $Dev(x, y, z)$ тренарлық қатынас орындалмайды. Сондықтан, осындай жағдайлар үшін нәтижелері туралы қосымша шарттар анықтау қажет.

1.1.2.10-анықтама. A жиынындағы R қатынасының *транзитивтік тұйықталуы* деп ішжиын ретінде құрамында R бар барлық транзитивтік қатынастардың қиылысуын айтады (басқаша айтқанда, ішжиын ретінде құрамында R бар минималды транзитивтік қатынас).

Транзитивтік тұйықталу кез келген қатынас үшін орындалады. Ол үшін айта кететін жағдай, транзитивтік қатынастардың кез келген жиындарының қиылысуы транзитивті болады. Сонымен қатар, ішжиын ретінде құрамында R бар транзитивті қатынас міндетті түрде бар болады.

Транзитивтік тұйықталу келесі қасиеттерге ие:

1) Рефлексивті қатынастың транзитивті тұйықталуы рефлексивті, себебі транзитивтік қатынас бастапқы қатынастан тұрады.

2) Симметриялық қатынастың транзитивті тұйықталуы симметриялы. Шынында, aRb транзитивтік қатынасы бар болсын, яғни $aRx_1, x_1Rx_2, \dots, x_nRb$ болатын $x_1, x_2, \dots, x_n$ бар болады. Бірақ, R қатынасының симметриялығынан $bRx_n, x_nRx_{n-1}, \dots, x_1Ra$ шығады, bRa .

3) Транзитивтік тұйықталу антисимметриялылықты сақтамайды, мысалы, $\{a, b, c\}$ жиынында $\{(a,b), (b,c), (c,a)\}$ қатынасы үшін.

4) Транзитивті қатынастың транзитивті тұйықталуы – ол өзі болады.

$R^* = R^+ \cup R^0$ қатынасын, мұнда $R^0 = \{(\varepsilon, \varepsilon) : \varepsilon \in A\}$ кейде рефлексивті-транзитивті тұйықталу деп атайды, бірақ көбінесе "транзитивті тұйықталу" деп нақты R^* айтады. Әдетте бұл

қатынастар арасындағы айырмашылықтар елеулі болып табылмайды.

I.1.2.10-мысалдар.

1. Кез келген $a, b, c \in N$ үшін $a < b$ және $b < c$ -дан $a < c$ шығады.

2. Кез келген $x, y \in N$ үшін және $x=y$ -тен $y=x$ шығады.

3. Егер A – қалалар жиыны болса және оларға "х-тен у-ке жететін маршрут бар" мағынасын беретін xRy қатынасы берілсе, онда бұл қатынастың транзитивті тұйықталуы "х-тен у-ке автобуспен жету мүмкіндігі бар" қатынасы болады.

I.1.2-тапсырмалар.

1. $A=\{a,b,c,d\}$, $B=\{b,d\}$, $C=\{c\}$ болғанда мына қатынастарды $A \supseteq C$ және $A \supseteq (B \cup C)$ анықтау керек;

2. Бинарлы қатынасқа анықтама беріңіз;

3. n -дік қатынасқа анықтама беріңіз;

4. A және B жиындарының арасындағы бинарлы қатынастың префиксті жазбасын көрсетіңіз;

5. A және B жиындарының арасындағы бинарлы қатынастың инфиксті жазбасын көрсетіңіз;

6. A және B жиындарының арасындағы бинарлы қатынастың постфиксті жазбасын көрсетіңіз;

7. Бинарлы қатынасты кесте көмегімен қарап көріңіз;

8. Эквивалентті қатынасқа анықтама беріңіз;

9. Симметриялық қатынастың транзитивті тұйықталуын табыңдар;

10. Транзитивті қатынастың транзитивті тұйықталуын табыңдар;

I.1.2-сұрақтар.

1. A жиындағы унарлы қатынасты не анықтайды?

2. Бинарлы қатынаста қанша жиын қолданылады?

3. Бинарлы қатынастың рефлексивтілігі қалай анықталады?

4. Бинарлы қатынастың транзитивтілігі қалай анықталады?

5. Бинарлы қатынастың симметриялылығы қалай анықталады?

6. Ішінара реттелген жиын неден тұрады?

7. Қатынас композициясы қалай анықталады?

8. Екі жиынның эквиваленттілігі қалай анықталады?
9. Ішінара ретті қатынас қалай анықталады?
10. Қандай жиындар салыстырылмайтын деп аталады?
11. Транзитивті тұйықталу дегеніміз не?

I.1.2-тестілер.

1. Егер A және B бірдей элементтен тұрса, яғни A және B жиындары бір-біріне ішкі жиын болып табылса, онда мына қатынастардың қайсысы рас болады?

- A) A жиыны B жиынына тең болады ($A = B$);
- B) A жиыны B -да бар болады ($A \subseteq B$);
- C) A жиыны қатаң B -да бар болады ($A \subset B$);
- D) B жиыны A -да бар болады ($B \subseteq A$);
- E) B жиыны қатаң A -да бар болады ($B \subseteq A$).

2. Егер A -ның барлық элементтері B -ға тиісті болса немесе A жиыны B -ға тең болса, онда мына қатынастардың қайсысы рас?

- A) A жиыны B -да бар болады ($A \subseteq B$);
- B) A жиыны B жиынына тең болады ($A = B$);
- C) A жиыны қатаң B -да бар болады ($A \subset B$);
- D) B жиыны A -да бар болады ($B \subseteq A$);
- E) B жиыны қатаң A -да бар болады ($B \subseteq A$).

3. Егер A -ның барлық элементтері B -ға тиісті болса және A жиыны B -ға тең болмаса, яғни B -ның кейбір элементтері A -ға тиісті емес болса, онда мына қатынастардың қайсысы рас болады?

- A) A жиыны қатаң B -да бар болады ($A \subset B$);
- B) A жиыны B жиынына тең болады ($A = B$);
- C) A жиыны B -да бар болады ($A \subseteq B$);
- D) B жиыны A -да бар болады ($B \subseteq A$);
- E) B жиыны қатаң A -да бар болады ($B \subseteq A$).

4. Егер әрбір $s \in S$ үшін sRs бар болса, онда R бинарлы қатынасы S жиынында не деп аталады?

- A) рефлексивті деп аталады;
- B) транзитивті деп аталады;
- C) симметриялы деп аталады;
- D) антисимметриялы деп аталады;

Е) эквивалентті деп аталады.

5. Егер кез келген $s, t, u \in S$ үшін sRt және tRu -дан sRu алынса, онда R бинарлы қатынасы S жиынында қалай аталады?

А) транзитивті деп аталады;

В) рефлексивті деп аталады;

С) симметриялы деп аталады;

Д) антисимметриялы деп аталады;

Е) эквивалентті деп аталады.

1.2. Функциялар теориясының элементтері

1.2.1. Функцияның анықтамасы, мәні және туындысы

Бұл тақырыпта қарапайым функциялар, күрделі функциялар және кері функциялар түсініктері, олардың берілу және мәндерін анықтау тәсілдері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау үшін мына дерек көз пайдаланылды [3].

X, Y – кейбір мәндердің жиыны (*сандық, логикалық, символдық* немесе басқа) болсын. Сонда X жиынының айнымалы x шамасының әрбір мәніне анықталған f ереже бойынша Y жиынында y шамасының жалғыз мәні сәйкес келеді, яғни $y = f(x)$ қарапайым функциясы беріледі. Қарапайым функциялар үшін X *берілу аймағы* немесе *анықталу аймағы* деп және *тәуелсіз айнымалысы* $x \in X$ *аргумент* деп аталады, ал Y жиыны – *мәндер аймағы* немесе *өзгеру аймағы* деп және *тәуелді айнымалысы* $y \in Y$ – *мәні* деп аталады.

Жалпы жағдайда, қарапайым f функциясы үшін анықталу аймағын $D(f)$ деп белгілеп, ал мәндер аймағын $E(f)$ деп белгілейді, яғни, оларды $D(f) = X, E(f) = Y$ түрінде жазуға болады.

Қарапайым функцияны тәуелсіз айнымалылардың мәндерінде орындалатын және нәтижесінде тәуелді айнымалының мәні алынатын амалдар және олардың орындалу тізбегін көрсететін кейбір алгебралық өрнектер түрінде аналитикалық беруге болады.

Функцияны тәуелсіз x айнымалысының мәндерін белгілі бір $x_1, x_2, \dots, x_n$ ретімен, ал олардың тұсында тәуелді y айнымалысының $y_1, y_2, \dots, y_n$ мәндерін көрсетіп, кесте түрінде беруге де болады:

Аргумент мәні	x_1	x_2	...	x_n
Функция мәні	y_1	y_2	...	y_n

Функцияның берілуін кестелік тәсілмен көрсету функцияның нақты мәнін ешқандай қосымша есептеулерсіз бірден анықтауға мүмкіндік береді, бірақ бұл аргументке байланысты функцияның өзгеруін көрнекі сурет түрінде бермейді.

Енді берілген қарапайым функциядан *күрделі функцияның* қалай жасалатындығын көрсетейік.

Мейлі қарапайым $u = \varphi(x)$ функциясы $D(\varphi) = X$ анықталу аймағымен және $E(\varphi) = D_1$ мәндер аймағымен, ал $y = \psi(u)$ функциясы $D(\psi) = D_1$ анықталу аймағымен және $E(\psi) = Y$ мәндер аймағымен берілсін. Сонда әрбір x мәніне φ және ψ сәйкестіктерін қолданып y мәнін анықтауға болады. Ол келесі түрде жазылады:

$$y = \psi(\varphi(x)).$$

$u = \varphi(x)$ функциясының мәнін *аралық айнымалы* немесе *аралық аргумент* деп атаймыз, $y = \psi(\varphi(x))$ функциясының функциясы тәуелсіз x айнымалысының *күрделі функциясы* немесе *суперпозиция* немесе ішкі φ және сыртқы ψ функцияларының *композициясы* деп аталады. Мұнда сыртқы функцияның анықталу аймағы ішкі функцияның мәндер аймағын қамтуы тиіс, яғни,

$$D(\psi) \supset E(\varphi).$$

Әдетте суперпозиция үшін $y = \psi(u)$ жазбасы қолданылады, онда $u = \varphi(x)$, ал композиция $\psi \circ \varphi$ түрінде жазылады, $\circ$ –композиция амалы.

Кері функция түсінігі өте маңызды болады, оны келесі түрде анықтауға болады.

Егер $y = g(x)$ функциясы берілсе, онда кері функциясы g^{-1} функциясы болады, $E(g)$ мәндер аймағынан әрбір y элементіне сәйкес $D(g)$ анықталу аймағының барлық x элементтерін қояды, олар үшін $g(x) = y$ орындалады, яғни, $g^{-1}(y) = \{x: g(x) = y\}$ және g және g^{-1} функциялары өзара кері деп аталады. Өзара кері

функциялар үшін $D(g) = E(g^{-1})$ және $D(g^{-1}) = E(g)$ байланыстары орындалады.

Егер $y = g(x)$ функциясы берілсе, оған кері g^{-1} функциясын табу үшін y айнымалысына қатысты теңдеуді шешу қажет, ол үшін x және y айнымалыларының орнын ауыстырамыз.

Функцияның шегі Коши бойынша былай анықталады:

$A \in E(f)$ мәні $f(x)$ функциясының $x_0 \in D(f)$ нүктесіндегі *Коши бойынша шегі* деп аталады, егер кез келген алдын ала алынған оң сан $\varepsilon > 0$ үшін оған жауап беретін $\delta(\varepsilon) > 0$ табылып, $0 < |x - x_0| < \delta$ шартын қанағаттандыратын барлық x аргументтерінде $|f(x) - A| < \varepsilon$ теңсіздігі орындалса.

$$\lim_{x \rightarrow x_0} f(x) = A \Leftrightarrow \forall \varepsilon > 0 \exists \delta(\varepsilon) > 0 \forall x: 0 < |x - x_0| < \delta \Rightarrow |f(x) - A| < \varepsilon.$$

$f(x)$ функциясы $x_0 \in D(f)$ нүктесінде *үздіксіз* болады, егер кез келген $\varepsilon > 0$ үшін $\delta(\varepsilon) > 0$ бар болса және кез келген $x \in D(f)$ үшін $0 < |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$ орындалса.

Егер f функциясы берілген E жиынының әрбір нүктесінде *үздіксіз* болса, онда ол барлық E жиынында *үздіксіз* болады.

Мейлі $f(x)$ мына $[a; b]$ интервалында анықталған кейбір сандық функция және x_0 осы интервалдағы кейбір бекітілген нүкте (кейбір тұрақты мән) болсын. Осы интервалда қандай да бір x мәні үшін $x - x_0 = \Delta x$ айырмасын құрамыз және оны *тәуелсіз айнымалының (аргументтің) өсімі* деп атаймыз. Сонда $f(x)$ функциясының x_0 нүктесіндегі *өсімі* деп оның $x = x_0 + \Delta x$ нүктесіндегі мәнімен x_0 нүктесіндегі мәнінің айырмасын айтамыз және оны мына өрнекпен белгілейміз.

$$\Delta f(x_0) = f(x_0 + \Delta x) - f(x_0),$$

мұнда x_0 бекітілген (тұрақты) деп саналғандықтан, функцияның $\Delta f(x_0)$ өсімі Δx аргумент өсімінің функциясы болады. Келесі қатынасты құрамыз

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{\Delta f(x_0)}{\Delta x},$$

бұл да Δx аргумент өсімінің функциясы болады және осы өрнектің Δx нөлге ұмтылғандағы шегін қарастырамыз. Егер осы шек болса, онда $f(x)$ функциясының x_0 нүктесінде туындысы бар дейміз және оны былай жазамыз

$$\lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{\Delta f(x_0)}{\Delta x} = f'(x_0).$$

$f'(x_0)$ санын x_0 нүктесіндегі функцияның туындысы дейді.

$f(x)$ функциясының x_0 нүктесіндегі туындысын былайда белгілейді.

$$f'(x)_{x=x_0} \text{ немесе } \frac{df(x_0)}{dx}.$$

Егер $f(x)$ функциясының x_0 нүктесінде туындысы болса, онда ол x_0 нүктесінде дифференциалданады дейді.

Егер $f(x)$ функциясы $[a; b]$ интервалының әрбір нүктесінде дифференциалданатын болса, онда ол $[a; b]$ интервалында дифференциалданады деп айтылады.

$f(x)$ функциясының $[a; b]$ интервалында дифференциалданатын туындысының өзі x -тің функциясы болады.

Мейлі $u(x)$ және $v(x)$ функциялары берілсін. Сонда функциялардың қосындысы, айырмасы, көбейтіндісі және бөліндісінің туындыларын есептеу ережелерін беруге болады:

$$1. (u \pm v)' = u' \pm v';$$

$$2. (u * v)' = u'v + v'u;$$

$$3. \left(\frac{1}{u} \right)' = -\frac{u'}{u^2};$$

$$4. \left(\frac{u}{v} \right)' = \frac{u'v - v'u}{v^2}.$$

Мейлі $y = f(g(x))$ күрделі функциясы берілсін. Сонда күрделі функция туындысы мына формула арқылы есептеледі

$$y' = (f(g(x)))' = f'(g(x)) \cdot g'(x).$$

$x_0 \in D(f)$ аргумент мәні $f(x)$ функциясының *локалды минимум нүктесі* (немесе *локалды максимум*) деп аталады, егер x_0 мәні немесе қандай да бір $\delta > 0$ шамасы үшін $[x_0 - \delta, x_0 + \delta] \subset D(f)$ аймағы табылып, осы аймақтағы барлық x мәндері үшін $f(x_0) < f(x)$ (немесе $f(x_0) > f(x)$) теңсіздігі орындалса.

Функция мәні локалды минимум нүктесінде (немесе локалды максимум) $f(x)$ функциясының x_0 *локалды минимумы* (немесе *локалды максимумы*) деп аталады.

$y = f(x)$ функциясы $x = x_0$ болғанда локалды минимумға ие, егер $f(x_0)$ мәні оның x_0 нүктесінің жанында қабылданатын басқа барлық мәнінен кіші болса, яғни, $f(x_0) < f(x)$ үшін $x \in [x_0 - \delta, x_0 + \delta]$ & $\delta > 0$.

$y = f(x)$ функциясы $x = x_0$ болғанда локалды максимумға ие, егер $f(x_0)$ мәні оның x_0 нүктесінің жанында қабылданатын басқа барлық мәнінен үлкен болса, яғни, $f(x_0) > f(x)$ үшін $x \in [x_0 - \delta, x_0 + \delta]$ & $\delta > 0$.

Функцияның минимумы және максимумы *экстремум* деп аталады. Сонымен қатар, $f(x_0) \leq f(x)$ (немесе $f(x_0) \geq f(x)$) шарты орындалса, онда экстремум (минимум немесе максимум) *қатаң емес* деп аталады, яғни, көрсетілген δ аймағында функцияның тең мәндері болуы мүмкін.

Мейлі x_0 нүктесі $f(x)$ функциясының x_0 нүктесінің кейбір аймағында анықталған экстремум нүктесі болсын делік. Сонда $f'(x_0)$ туындысы жоқ болады немесе $f'(x_0) = 0$.

Мейлі $f(x)$ функциясы x_0 нүктесінде екі рет дифференциалданған болсын. Сонда $f'(x_0) = 0$ және $f''(x_0) > 0$ шарттары бойынша x_0 нүктесі локалды минимум нүктесі болады. Ал егер $f'(x_0) = 0$ және $f''(x_0) < 0$ болса, онда x_0 нүктесі локалды максимум нүктесі болады.

I.2.1-тапсырмалар.

1. $y = \log_2 x$ функциясының кері функциясын табыңыз.
2. $y = \arctg^2 x$ функциясының мәнін және анықталу аймағын табыңыз.
3. $y = x^2$ функциясының минимумын табыңыз.

I.2.1-сұрақтар.

1. Күрделі функциялар қалай анықталады?
2. Кері функциялар қалай анықталады?
3. Функцияның максимумы және минимумы дегеніміз не?
4. x_0 нүктесінде функцияның дифференциалдануы дегеніміз не?

I.2.1-тестілер.

1. Экстремум дегеніміз не?
А) максимум және минимум функциялар;
В) дәрежелік және көрсеткіштік функция;
С) нақты және бүтін функция;
D) симметриялы және симметриялы емес функция;
Е) қарапайым және күрделі функция.
2. $f(x)$ функциясы x_0 нүктесінде қалай дифференциалданады?
А) егер оның x_0 нүктесінде туындысы болса;
В) егер оның x_0 нүктесінде интегралы болса;
С) егер оның x_0 нүктесінде 0-ге тең болса;
D) егер оның x_0 нүктесінде 1-ге тең болса;
Е) егер оның x_0 нүктесінде мәні жоқ болса.
3. g және g^{-1} функциялары өзара қалай аталады?
А) өзара кері;
В) өзара қарама-қарсы;
С) өзара тең;
D) өзара тең дәрежелі;
Е) өзара теңбе тең.
4. f функциясы E жиынында қай уақытта үздіксіз болады?
А) егер ол берілген жиынның әрбір нүктесінде үздіксіз болса;
В) егер ол берілген жиынның әрбір нүктесінде үздікті болса;
С) егер ол берілген жиынның әрбір нүктесінде кемімелі болса;
D) егер ол берілген жиынның әрбір нүктесінде өспелі болса;
Е) егер ол берілген жиынның әрбір нүктесінде анықталған болса.
5. $f(x)$ функциясы x_0 нүктесінде екі рет дифференциалданған болса, онда $f'(x_0)=0$ және $f''(x_0)<0$ шарттары бойынша x_0 нүктесі қандай нүкте болады?

- A) локалды максимум нүктесі;
- B) локалды минимум нүктесі;
- C) локалды экстремум нүктесі;
- D) глобалды максимум нүктесі;
- E) глобалды минимум нүктесі.

1.2.2. Сандық функциялар және олардың қасиеттері

Бұл тақырыпта қарапайым функция, күрделі функция және кері функциялар сияқты сандық функциялардың түрлері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау үшін мына дерек көз пайдаланылды [3].

Функция түрлерін оқу және анықтау кезінде $y = f(x)$ функциясы үшін $D(f)$ анықталу аймағы және $E(f)$ мәндер аймағы сандар жиынымен берілу жағдайы маңызды болады.

$D(f)$ және $E(f)$ – сандар жиыны болсын, онда $D(f)$ жиынында x сандық аргументінен және $f(x)$ сандық мәнінен *сандық функция* қойылған деп айтылады.

Сандық функциялар арасынан $D(f)$ және $E(f)$ нақты R сандардың ішжиындарының жиыны, яғни $D(f) \subseteq R$ және $E(f) \subseteq R$ болып табылатын *нақты функциялар* қызығушылық тудырады,.

Нақты функциялардың көптеген түрлері бар, бірақ оның ішінде *негізгі элементар функциялар* көп қарастырылады.

Негізгі элементар функцияларға мыналар жатады:

1. Дәрежелік функция:

$$y=x^a \quad (a \in R - \text{тұрақты шама});$$

2. Көрсеткіштік функция:

$$y=a^x \quad (a \in R - \text{тұрақты шама және } a > 0, a \neq 1);$$

3. Логарифмдік функция:

$$y=\log_a x \quad (a \in R - \text{тұрақты шама және } a > 0, a \neq 1);$$

4. Тригонометриялық функциялар:

$$y= \sin x, y= \cos x, y= \operatorname{tg} x, y= \operatorname{ctg} x, y= \sec x, y= \operatorname{cosec} x;$$

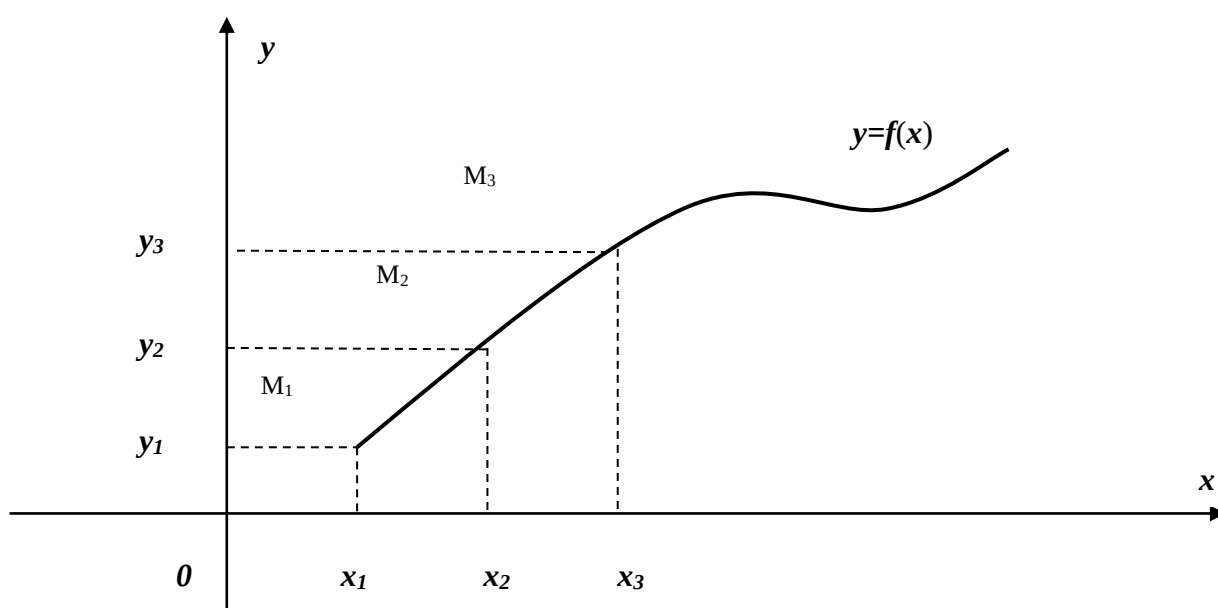
5. Кері тригонометриялық функциялар:

$$\begin{aligned} y &= \arcsin x, y = \arccos x, \\ y &= \operatorname{arctg} x, y = \operatorname{arcctg} x, \\ y &= \operatorname{arcsec} x, y = \operatorname{arccosec} x. \end{aligned}$$

Элементар функциялар деп арифметикалық амалдардың және суперпозицияның ақырлы санын қолдана отырып негізгі элементар функциялар арқылы өрнектеуге болатын функцияларды айтамыз.

Сандар жиынын түзумен бейнелеуге болады, ол үшін түзудегі нүктелерді белгілі сандық мәндермен белгілеу керек. Мұнда, горизонтал–көлденең (вертикал–тік) түзудегі 0 мәнін беретін нүктенің *сол (төмен)* жағындағы барлық нүктелер *теріс сандарды*, ал *оң (жоғары)* жағындағы нүктелер *оң сандарды* бейнелейді.

Берілген функцияны графигтік бейнелеу үшін абсцисса осіне x аргументінің $x_1, x_2, x_3, \dots$ мәндер қатарын белгілейміз және y функциясының мәніне сәйкес $y_1, y_2, y_3, \dots$ ординаталарын құрамыз. Нәтижесінде $M_1(x_1, y_1), M_2(x_2, y_2), M_3(x_3, y_3), \dots$ нүктелер қатарын аламыз. Оларды қисық сызықпен біріктіре отырып, I.2.2.1-суретте көрсетілгендей осы функционалдық тәуелділіктің графигін аламыз.



I.2.2.1-сурет. Функционалдық тәуелділіктің графигі

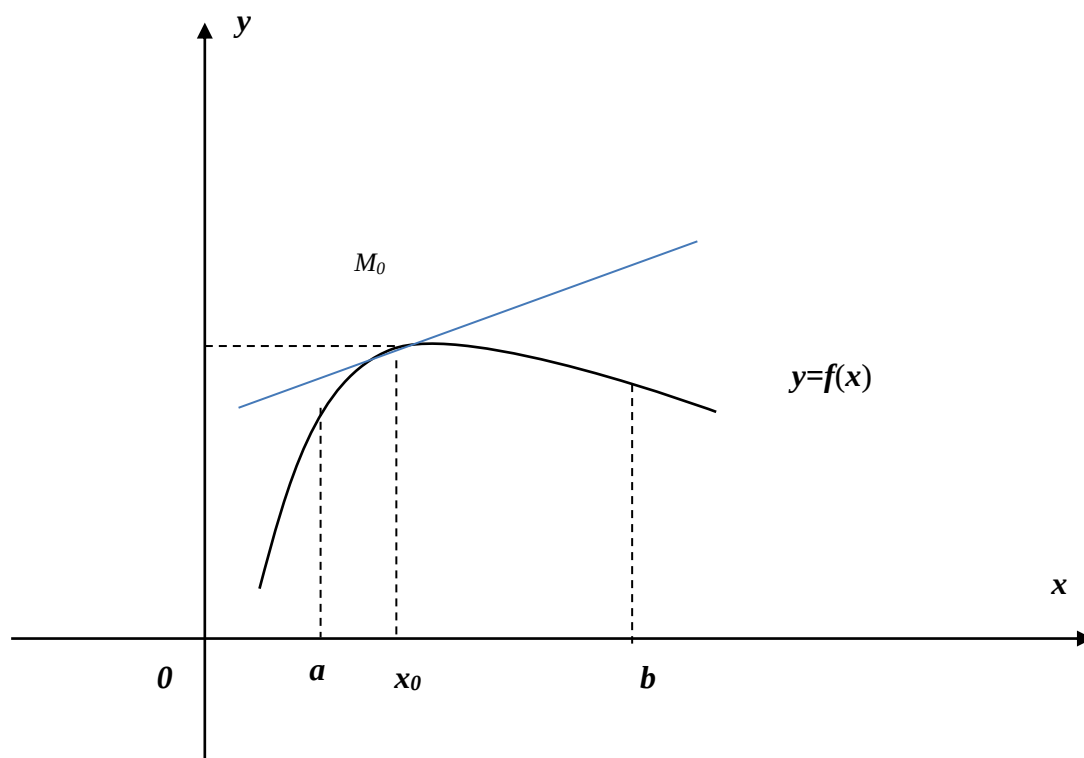
Кестеге қарағанда графикалық көріністің артықшылығы оның көрнекілігі және көзбен жеңіл шолу болады; ал кемшілігі – дәлдік дәрежесі төмен болады. Үлкен практикалық маңыздылығы масштабтарды таңдау мүмкіндігіне ие.

I.2.2-мысалдар.

1. Негізгі элементар функциялардың арасынан $\ln x$, $\cos x$, $\arcsin x$, $\arccos x$, $\arctg x$, $\operatorname{arcctg} x$, $\operatorname{arcsec} x$, $\operatorname{arccosec} x$ шектеулі болады, мысалы, $|\cos x| \leq 1$.

2. Барлық негізгі элементар функциялар барлық анықтау аймағында да, оның жеке бөліктерінде де монотонды.

3. $y = f(x)$ функция графигінің $x_0 \in [a; b]$ абсцисасымен кез келген M_0 нүктесін алайық және M_0 нүктесі арқылы теңдеуі $y = f'(x_0)(x - x_0) + f(x_0)$ болатын жанама жүргіземіз. Функцияның $[a, b]$ аралығындағы графигі бұл жанамадан төменде жататындығын (дөңес болатындығын) көрсетуіміз керек, яғни x -тің қандай да бір мәнінде $y = f(x)$ қисығының ординатасы I.2.2.2-суретте көрсетілгендей жанама ординатасынан кіші болады.

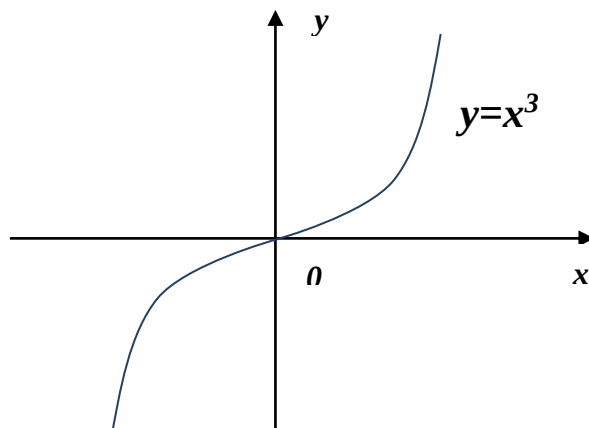


I.2.2.2-сурет. Дөңес график

4. $[0; 2; \pi]$ аралығындағы $y = \ln x$ функциясының графигі $[0; \pi]$ интервалында дөңес және $[\pi; 2\pi]$ интервалында ойық.

5. $y = x^3$ функциясының графигі $[-\infty; +\infty]$ шексіз интервалында дөңес те және ойық та бола алады, себебі $[-\infty; 0]$ – интервалында

дөңес, ал $[0; +\infty]$ – интервалында ойық, сонымен қатар $x=0$ нүктесі графикте $y=x^3$ функциясының дөңес бөлігінен ойық бөлігін ажыратады.



I.2.2-тапсырмалар.

1. Нақты функциялардың анықталу аймағын көрсетіңіз.
2. Нақты функциялардың мәндер аймағын көрсетіңіз.
3. Негізгі қарапайым функцияларды атап көрсетіңіз.

I.2.2-сұрақтар.

1. Дәрежелі және көрсеткіштік функциялар қалай анықталады?
2. Логарифмдік функциялар қалай аталады?
3. Тригонометриялық функциялар қалай анықталады?

I.2.2-тестілер.

1. $D(f)$ және $E(f)$ – сандар жиыны болсын, онда $D(f)$ жиынында x сандық аргументінен және $f(x)$ сандық мәнінен қойылған функция қалай аталады?

- А) сандық функция;
- В) дәрежелік функция;
- С) нақты функция;
- Д) симметриялы функция;
- Е) элементар функция.

2. Арифметикалық амалдардың және суперпозицияның ақырлы санын қолдана отырып негізгі элементар функциялар арқылы өрнектеуге болатын функция қалай аталады?

- A) элементар функция;
- B) дәрежелік функция;
- C) нақты функция;
- D) симметриялы функция;
- E) сандық функция.

3. Мыналардың қайсысы дәрежелік функция болады?

- A) $y=x^a$ ($a \in R$ – тұрақты шама);
- B) $y=\log_a x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$);
- C) $y=\arcsin x$;
- D) $y=\sin x$;
- E) $y=a^x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$).

4. Мыналардың қайсысы тригонометриялық функция болады?

- A) $y=\arcsin x$;
- B) $y=\log_a x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$);
- C) $y=x^a$ ($a \in R$ – тұрақты шама);
- D) $y=kax$;
- E) $y=a^x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$).

5. Мыналардың қайсысы көрсеткіштік функция болады?

- A) $y=a^x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$);
- B) $y=\log_a x$ ($a \in R$ – тұрақты шама және $a>0$, $a \neq 1$);
- C) $y=x^a$ ($a \in R$ – тұрақты шама);
- D) $y=\sin x$;
- E) $y=\arcsin x$.

1.2.3. Нақты функциялар.

Бұл тақырыпта нақты функция түрлері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [3,4].

Нақты функциялар қасиеттеріне қарай келесі түрлерге бөлінеді:

- шектелген және шектелмеген функциялар;
- өспелі және кемімелі функциялар;
- сатылы функциялар;
- тақ және жұп функциялар;
- периодты функциялар;
- дөңес және ойыс функциялар.

Шектелген және шектелмеген функциялар.

$y = f(x)$ функциясы *шектелген* деп аталады, егер кез келген $x \in D(f)$ үшін $|f(x)| \leq M$ шарты орындалатын M нақты оң санын көрсетуге болатын болса.

$y = f(x)$ функциясы *шектелмеген* деп аталады, егер M нақты оң саны үшін $|f(x)| > M$ теңсіздігін қанағаттандыратын $x \in D(f)$ нүктесі табылатын болса.

Шектелген функция графигі $y = -M$ және $y = M$ параллел түзулері арасында жатады.

Өспелі және кемімелі функциялар.

Мейлі кейбір сандық X жиыны функция анықталған $D(f)$ жиынында жататын болсын, яғни $X \subseteq D(f)$. Сонда:

1) X сандар жиынында $y = f(x)$ функциясы *өспелі* (*қатаң өспелі*) деп аталады, егер X жиынындағы кез келген екі $x_1 < x_2$ саны үшін $f(x_1) < f(x_2)$ теңсіздігі орындалса.

2) X сандар жиынында $y = f(x)$ функциясы *кемімелі* (*қатаң кемімелі*) деп аталады, егер $x_1 < x_2$ шарты үшін $f(x_1) > f(x_2)$ теңсіздігі орындалса.

3) $y = f(x)$ функциясы барлық анықталу аймағы $D(f)$ -те өспелі немесе кемімелі болса, онда функция қатаң монотонды деп аталады.

4) X жиынында $y = f(x)$ функциясы кемімелі емес деп аталады, егер X жиынының кез келген $x_1 < x_2$ саны үшін $f(x_1) \leq f(x_2)$ шарты орындалса.

5) $y = f(x)$ функциясы өспелі емес деп аталады, егер $x_1 < x_2$ шарты үшін $f(x_1) \geq f(x_2)$ теңсіздігі орындалса.

6) Функция, өзінің анықталу аймағының әрбір байланысқан ішжиындарында өспелі емес немесе кемімелі емес болса, монотонды деп аталды.

7) Екі өспелі (кемімелі) функцияның қосындысы өспелі (кемімелі) болады.

8) Екі оң өспелі (кемімелі) функция көбейтіндісі өспелі (кемімелі) функция болады.

9) Егер $y = f(x)$ функциясы өспелі және $f(x) \neq 0$ болса, онда $1/f(x)$ функциясы кемімелі және керісінше.

10) Егер $y = f(x)$ функциясы қатаң монотонды болса, онда оның $x = f^{-1}(y)$ кері функциясы да *қатаң монотонды*.

11) Егер $y = f(x)$ функциясы X жиынында *бөлікті-монотонды* деп аталады, егер бұл жиын әрбір берілген функция монотонды болатындай ішжиындарға бөлінетін болса.

12) Егер $y = f(x)$ функциясы X жиынында өспелі (немесе кемімелі) болса, онда ол сол жиында *қайтарымды* болады, яғни $x_1 \neq x_2$ шартынан $f(x_1) \neq f(x_2)$ шығады.

Сатылы функциялар.

$D(f)$ анықталу аймағын өзара қиылыспайтын $D_1, D_2, \dots$, ішаймақтарға бөлуге болатын болса, онда функцияны аналитикалық бергенде бір $x \in D_1$ аргументінің мәні үшін функция бір өрнекпен беріле алады, ал аргументтің басқа мәндері $x \in D_2$ үшін – екінші өрнек және т.с.с. Кейде осындай функциялар *сатылы функциялар* деп аталады.

Тақ және жұп функциялар.

Егер барлық $x \in D(f)$ аргумент мәндері үшін $f(-x) = f(x)$ теңдігі орындалса, онда $f(x)$ функциясы *тақ* болады.

Егер барлық $x \in D(f)$ аргумент мәндері үшін $f(-x) = -f(x)$ теңдігі орындалса, онда $f(x)$ функциясы *жұп* болады.

Тақ және жұп функциялар келесі қасиеттерге ие болады:

1) Екі тақ (жұп) функцияның қосындысы тақ (жұп) функция болады.

2) Екі тақ (жұп) функцияның көбейтіндісі тақ (жұп) функция болады, ал жұп және тақ функция көбейтіндісі тақ функция болады.

3) Егер $y = \psi(u)$ бірінші функция және $u = \varphi(x)$ екінші функция тақ болса, онда $y = \psi(\varphi(x))$ күрделі функциясы да тақ болады.

4) Егер $y = \psi(u)$ бірінші функция жұп және $u = \varphi(x)$ екінші функция тақ болса, онда $y = \psi(\varphi(x))$ күрделі функциясы жұп болады.

5) Егер $f(x)$ функциясы жұп және $f(x) \neq 0$ шарты орындалса, онда $1/f(x)$ түрінде берілген өрнек те жұп функция болады.

Жұп функция графигі ордината осіне қатысты симметриялы, ал тақ функциянікі – координата бас нүктесіне қатысты симметриялы. Сонымен қатар тақ және жұп функциялардың анықталу аймағы координатаның бас нүктесіне қатысты симметриялы. Бұл функцияның тақ немесе жұп болуының қажетті (бірақ жеткілікті емес) шарты.

Периодты функциялар.

$f(x)$ функциясы периодты деп аталады, егер u мәні $x \in D(f)$ аргумент мәніне нөл емес T санын қосқанда өзгермесе, яғни әрбір $x \in D(f)$ және $T \neq 0$ үшін $f(x) = f(x+T)$ теңдігі орындалуы қажет. Мұнда T саны $f(x)$ функциясының *периоды* деп аталады. Периодты функцияның ең аз оң периоды оның *негізгі периоды* болады. Егер периодты функцияның негізгі периоды T_0 бар болса, онда оның басқа периодтары сол T_0 периодына еселенген болады.

Дөңес және ойыс функциялар.

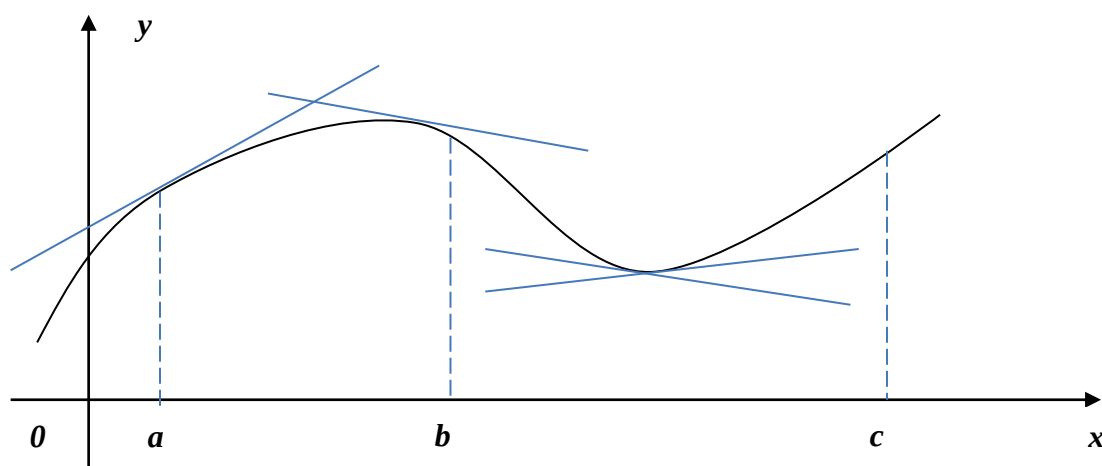
Функция дөңес деп аталады, егер оның $y = f(x)$ графигі (қисығы)

$[a;b]$ интервалында дөңес (жоғары қараған дөңес) болса және ол берілген интервалда өзінің жанамасынан төмен орналасса.

$y = f(x)$ функция графигі (қисығы) $[a;b]$ интервалында ойыс (төмен қараған ойыс) деп аталады, егер ол берілген интервалда өзінің жанамасынан жоғары орналасса.

Егер нүкте бойында жанама бар болса, онда ол иілу нүктесі деп аталады.

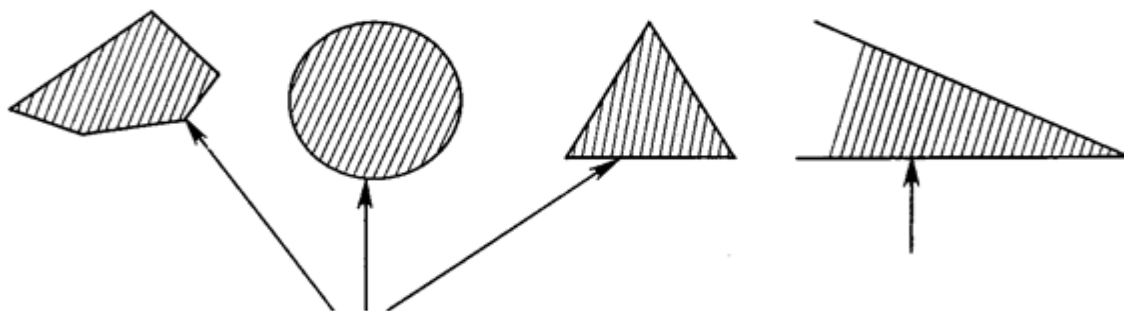
1.2.3.1- суретте функция графигі берілген, $[a;b]$ интервалында дөңес және $[b;c]$ –да ойыс.



1.2.3.1-сурет. Дөңес және ойыс график

Аймақ дөңес деп аталады, егер оның кез келген екі нүктесін байланыстыратын бөлігі толығымен сол аймақта жататын болса. Аймақ шектелген деп аталады, егер оны шеңберге алуға болса.

Шектелмеген аймақты шеңбер ішіне (мысалы, екі параллел түзудің арасына) алуға болмайды. 1.2.3.2-суретте шектелген және шектелмеген дөңес аймақтардың мысалдары келтірілген.



1.2.3.2-сурет. Шектелген және шектелмеген дөңестер

1.2.3-мысалдар.

1. Кез келген функция тақ немесе жұп болмайды, мысалы,

$y = f(x) = x+1$ функциясы тақ та, жұп та емес, себебі

$$f(-x) = -x+1 \neq f(x) \text{ и } f(-x) = -x+1 \neq -f(x).$$

2. Кейбір негізгі элементар функциялар жұп болады, ал басқалары тақ: $\cos x$ – жұп функция, $\sin x$ – тақ функция.

1.2.3-тапсырмалар.

1. Шектелген және шектелмеген функция анықтамалары.

2. Келесі $y = x^2$, $y = x^5$ функциялардан жұп функцияны анықтаңыз:

3. Бөлікті-монотонды функцияның шартын анықтаңыз.

4. -1, 0, 1 мәндері арқылы сатылы функцияны құрыңыз.

5. Периодты функцияның шартын анықтаңыз.

1.2.3-сұрақтар.

1. Қандай функциялар сатылы деп аталады?

2. Өспелі және кемімелі функцияның айырмашылығы қандай?

3. Жұп және тақ функциялардың айырмашылығы?

4. Периодты функциялар қалай анықталады?

5. Функцияның ойыс және дөңес графиктері қалай анықталады?

1.2.3-тестілер.

1. Егер M нақты оң санын көрсетуге болатын болса және кез келген $x \in D(f)$ үшін $|f(x)| \leq M$ шарты орындалса $y = f(x)$ функциясы қалай аталады?

А) шектелген;

В) шектелмеген;

С) өспелі;

Д) сатылы;

Е) тақ.

2. Егер M нақты оң саны үшін $|f(x)| > M$ теңсіздігін қанағаттандыратын $x \in D(f)$ нүктесі табылатын болса, онда $y = f(x)$ функциясы қалай аталады?

- A) шектелмеген;
- B) шектелген;
- C) өспелі;
- D) сатылы;
- E) тақ.

3. Егер X жиынының кез келген $x_1 < x_2$ саны үшін $f(x_1) < f(x_2)$ теңсіздігі орындалса, X сандар жиынында $y = f(x)$ функциясы қалай аталады?

- A) өспелі;
- B) шектелген;
- C) шектелмеген;
- D) сатылы;
- E) тақ.

4. $y = f(x)$ функциясы $D(f)$ -тің барлық анықталу аймағында өспелі немесе кемімелі болса, онда функция қалай аталады?

- A) қатаң монотонды;
- B) шектелген;
- C) шектелмеген;
- D) сатылы;
- E) тақ.

5. Егер y мәні $x \in D(f)$ мәндеріне T санын тіркегенде өзгермесе, яғни $f(x) = f(x+T)$ болса, ол функция қалай аталады?

- A) периодты;
- B) шектелген;
- C) шектелмеген;
- D) сатылы;
- E) тақ.

II. НЕЙРОНДЫҚ ЖЕЛІЛЕР ТЕОРИЯСЫНЫҢ ЭЛЕМЕНТТЕРІ

II.1 Негізгі түсініктер

II.1.1 Табиғи нейрондар

Бұл тақырыпта табиғи нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

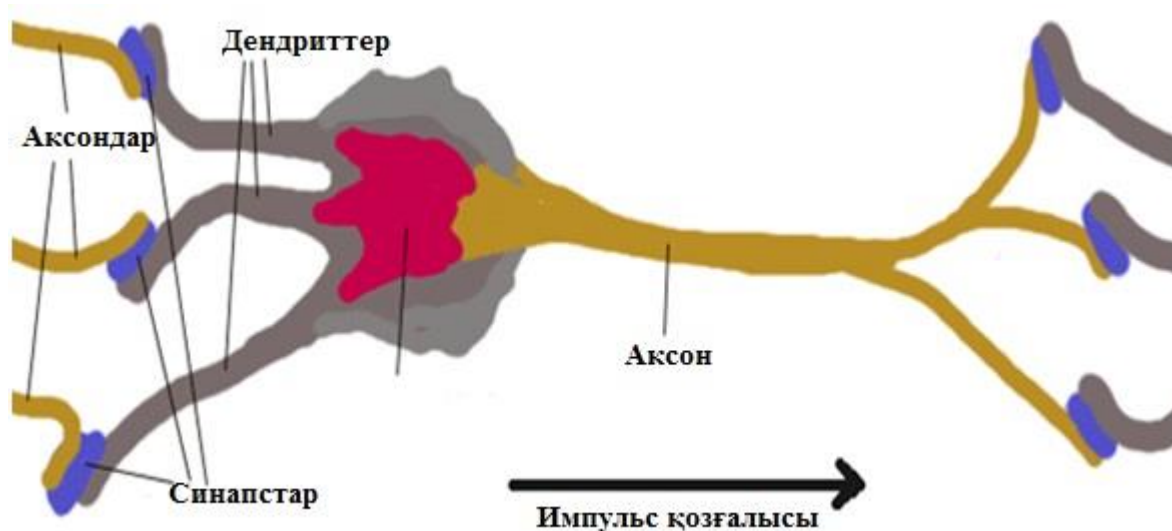
Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5-38].

Адамдарды өз ойлау қабілеті қызықтырған кезден бастап адам миын зерттеу басталды деп санауға болады. Өзің туралы ойлау адам миының басты ерекшелігі болып табылады. Ойлау тегі туралы руханидан бастап анатомиялыққа дейінгі көптеген толғаулар бар. Философтар мен теологтардың физиологтар мен анатомдардың әсерлі дауындағы бұл сұрақтарды талқылау пәнді оқып үйрену қиындығына байланысты аз пайда әкелді. Өздігінен талдау мен толғау жүргізгендер нақты ғылымдардың қатаң деңгейіне жауап бермейтін қорытындылар шығарды. Тәжірибе жүргізушілер миды бақылаудың қиындығын және оның құрылысы тығырыққа алып келетіндігін айтты. Басқа сөзбен айтқанда, біздің нақты өмірге деген көзқарасымызды өзгерткен қуатты ғылыми зерттеу әдістері, адамды түсінуге келгенде әлсіз болып қалады.

Адамның миы бүгінгі таңда ең күрделі ақпарат өңдеу жүйесі болады және өзара байланысқан 100 миллиардқа (10^{11}) жуық *нейрондардан* құралған. Әрбір нейронда орта есеппен 10 мың (10^4) байланыс бар, олар 10^{15} өзара байланыс тудырады. Дегенмен, күн сайын көптеген нейрондардың өлуіне қарамастан, адам миы өте сенімді, ал ми сөз өз жұмысын жалғастыруда.

Нейрон әсер уақыты бірнеше миллисекунд болатын жай қимылдайтын элемент екеніне қарамастан, ми үлкен көлемді ақпаратты өңдеуді өте жылдам, секундтың бөлігі ішінде, жүзеге асырады. Мидың сондай жоғары сенімділігі мен жылдамығының үйлесімділігі әзірше түсініксіз. Заманауи ғылым мен техника жеке нейрондардың құрылымы мен функцияларын жақсы зерттеген, мидың кейбір құрылымдық бөлігіндегі нейрондардың өзара ішкі

және сыртқы байланыстары туралы деректер бар, ал әртүрлі құрылымдардың ақпаратты өңдеу үрдісіне қатысуы туралы деректер өте аз. Әрбір нейронда сома – нейрон денесі, дендриттер – кіріс байланыстар жиыны, аксон – соңында тағы да тарамдалатын жалғыз шығыс байланысы, синапстар – аксондар мен басқа нейронның дендриттерімен байланыс орнататын дәнекер бар. Синпас екі нейрон арасындағы элементар құрылым және функцияналдық түйін болады. II.1.1-суретте табиғи нейронның құрылымы көрсетілген.



II.1.1-сурет. Табиғи нейронның құрылымы

Нейрон кіріс ақпаратын өзінің дендриты арқылы басқа нейрон аксондарынан алады. Дендриттер нейрон маңында мейлінше үлкен кеңістікке тармақталады. Дендрит ұзындығы 1мм-ге дейін барады.

Алынған ақпарат сома арқылы шығыс ақпаратын тудыру үшін өңделеді. Соманың көлденең өлшемі бірнеше ондаған микрон. Шығыс ақпараты, соңында көптеген синапстарға тармақталатын аксондар арқылы басқа нейрондарға беріледі. Аксон ұзындығы жүздеген миллиметрге жетуі мүмкін. Синапстың көлденең өлшемдері, әдетте, бірнеше микроннан аспайды, көбінесе бұл мөлшер шамамен 1 мкм. Синапс арқылы жіберу үнемі бір бағытты. Импульсті жіберу бағытына байланысты пресинапсты және постсинапсты жасушалар деп бөлінеді.

Импульс синапстық аяқталуға жеткенде, нейротрансмиттер деп аталатын химиялық заттар босап шығады. Нейротрансмиттер синапстың түріне қарай нейрон-қабылдағыштың электрлік импульстарды тудыру қабілеттілігін үдете немесе тежей отырып синапстық саңылау арқылы қозғалады. Жеке нейрон ақпаратты өңдеудің элементар бірлігі бола алмайды, ол нервтік орталықтың қызметін атқарады. Дендриттер мен аксондар желі құру арқылы басқа нейронның мембрана аймақтарымен байланысқа түсе алады. Бұл желілер ақпаратты өңдеу жүйесінің қызметін атқарады.

Биологиялық (табиғи) нейронның жұмыс істеу алгоритмі келесіден тұрады. Синапс арқылы өткен электрлік сигнал өзінің амплитудасын өзгертеді: арттырады немесе кемітеді. Мұны сигнал амплитудасын салмақтық (синапстық) коэффициентке көбейту деп түсінуге болады. Дендриттің дарақтарда ілінген кіру сигналдары жасуша денесінде қосылады, және содан соң аксондық шығыста шығыс импульсы (спайк) немесе импульстар бумасы тудырылады. Шығу сигналы аксон тармағы бойынша жүріп өтіп, аксонды басқа нейрондардың дендрит дарақтарымен қосатын синапстарға жетеді. Синапстар арқылы сигнал іргелес нейрондар үшін жаңа кірістік сигналға өзгереді. Бұл сигнал синапстың түріне байланысты оң немесе теріс (қоздырушы немесе тезеуші) болуы мүмкін. Синапстың шығысында түрлендіру сигналының шамасы синапстың қалыптасу кезінде өзгере алатын синапстық коэффициентпен (синапс салмағымен) анықталады.

Нейрондарды үш үлкен топқа бөлуге болады: рецепторлық, эффекторлық және аралық. Рецепторлық нейрондар миға сенсорлы ақпаратты енгізуді жүзеге асырады. Олар сезім органдарына түсетін (көз торындағы оптикалық, құлақ қалқанындағы акустикалық немесе мұрын хеморецепторларындағы иіс) сигналдарды өздерінің аксондарының электрлік импульстарына түрлендіреді. Эффекторлық нейрондар оларға келіп түскен сигналдарды атқарушы органдарға жібереді. Олардың аксондарының ұштарында атқарушы органдармен байланыстыратын арнайы синапстық қосылыстар бар, мысалы,

бұлшықеттермен, оларда нейрондардың қозуы бұлшықеттің жиырылуына алып келеді. Аралық нейрондар рецепторлардан алынған ақпараты өңдеуді жүзеге асырады және эффекторлар үшін басқару сигналын қалыптастырады. Нақты осылар орталық жүйке жүйесін құрайды.

Дендрит дарактарының кіру сигналдары салмақталады және сомада қосылады, егер оның нәтижесі белгілі бір шектен аспайтын болса, онда шығу сигналы мүлдем қалыптаспайды—нейрон «жұмыс істемейді». Шығу сигналы аксон тармақтары арқылы өтеді және синапстарға жетеді, олар аксонды басқа нейронның дендриттік дарактарымен байланыстырады. Синапстар арқылы сигналдар іргелес нейрондар үшін жаңа кірістік сигналға өзгереді. Бұл кірістік сигнал синапстың түріне байланысты оң немесе теріс, яғни қоздырушы немесе тежеуші болуы мүмкін. Синапс тудыратын кірістік сигнал шамасы синапс арқылы өтетін сигнал шамасының бірдей болған жағдайында да әр түрлі болуы мүмкін. Бұл ерекшеліктер синапс тиімділігімен немесе салмағымен анықталады, мұнымен қоса синапстың салмағы синапстың жұмыс істеу үдерісінде де өзгере береді. Көптеген ғалымдар осындай өзгерісті нейрофизиологиялық коррелят, яғни жады) іші деп санайды. Мұндай жағдайда молекулярлы жады механизмінің рөлі осы іздердің ұзақ уақытқа бекітілуінде болады.

Табиғи нейронның функционалдық жұмысы былай болады:

- басқа нейрон аксондарынан қабылданған кіру сигналдары синапстар арқылы өтеді де, синапстар салмағына пропорционалды түрде өзгереді;
- сомаға бір мезгілде бірнеше дендриттар бойынша келіп түскен өзгертілген кіру сигналдары қосындыланады;
- егер қосынды импульс қандай бір межеден асатын болса, онда нейрон қозады және сомадан аксондарға кететін және тармақталатын шығу сигналын қалыптастырады.
- шығу сигналдарының тармақтары синапстар арқылы басқа нейрондардың дендриттеріне беріледі.

Мұнымен қоса, синапстық байланыс салмақтары уақыт өте өзгеруі мүмкін, ендеше нейрондардың қылығы да өзгеруі мүмкін.

Тұтастай алғанда табиғи (биологиялық) нейрондық желілер келесі қасиеттерге ие:

- ақпаратты және есептеуді таратылған түрде бейнелеуі; – ақпаратты өңдеудің параллелділігі;
- нейрондар арасындағы байланыстың салмағы бойынша өзгеруі;
- оқып үйренуге қабілеттілігі және жалпылауға қабілеттілігі;
- ақпаратты контекстік өңдеу қасиеті;
- бейімділігі және қателерге төзімділігі.

II.1.1-тапсырмалар.

1. Табиғи нейронның құрамын атап шығыңыздар.
2. Табиғи нейронның құрамын және құрылысын салыңыз.
3. Табиғи нейронның функционалдық жұмысын сипаттаңыз.
4. Рецепторлық нейронның қызметін сипаттаңыз.
5. Эффлекторлық нейронның қызметін сипаттаңыз.
6. Аралық нейронның қызметін сипаттаңыз.
7. Табиғи нейронның функционалдық жұмысын сипаттаңыз.

II.1.1-сұрақтар.

1. Тірі организмнің миы нелерден тұрады?
2. Адам миында қанша нейрон бар?
3. Әрбір нейронда қанша байланыс бар?
4. Адам миында барлығы қанша байланыс болады?
5. Нейрондар не арқылы өзара әрекеттеседі?
6. Нейрондар ақпаратты не арқылы қабылдайды? 7. Табиғи нейрондар қандай қасиеттерге ие?

II.1.1-тестілер.

1. Әрбір нейрондағы нейрон денесі қалай аталады?
А) сома;
В) дендриттер;

- C) аксондар;
- D) синапстар;
- E) нейрондар.

2. Әрбір нейрондағы кіріс байланыстары қалай аталады?

- A) дендриттер;
- B) сома;
- C) аксондар;
- D) синапстар;
- E) нейрондар.

3. Әрбір нейрондағы жалғыз шығыс байланысы қалай аталады?

- A) аксондар;
- B) сома;
- C) дендриттер;
- D) синапстар;
- E) нейрондар.

4. Нейрондардың аксондары мен дендриттерінің арасындағы байланысы үшін не қолданылады?

- A) синапстар;
- B) сома;
- C) дендриттер;
- D) аксондар;
- E) нейрондар.

5. Импульс синапстық аяқталуға жеткенде, қандай химиялық заттар босап шығады?

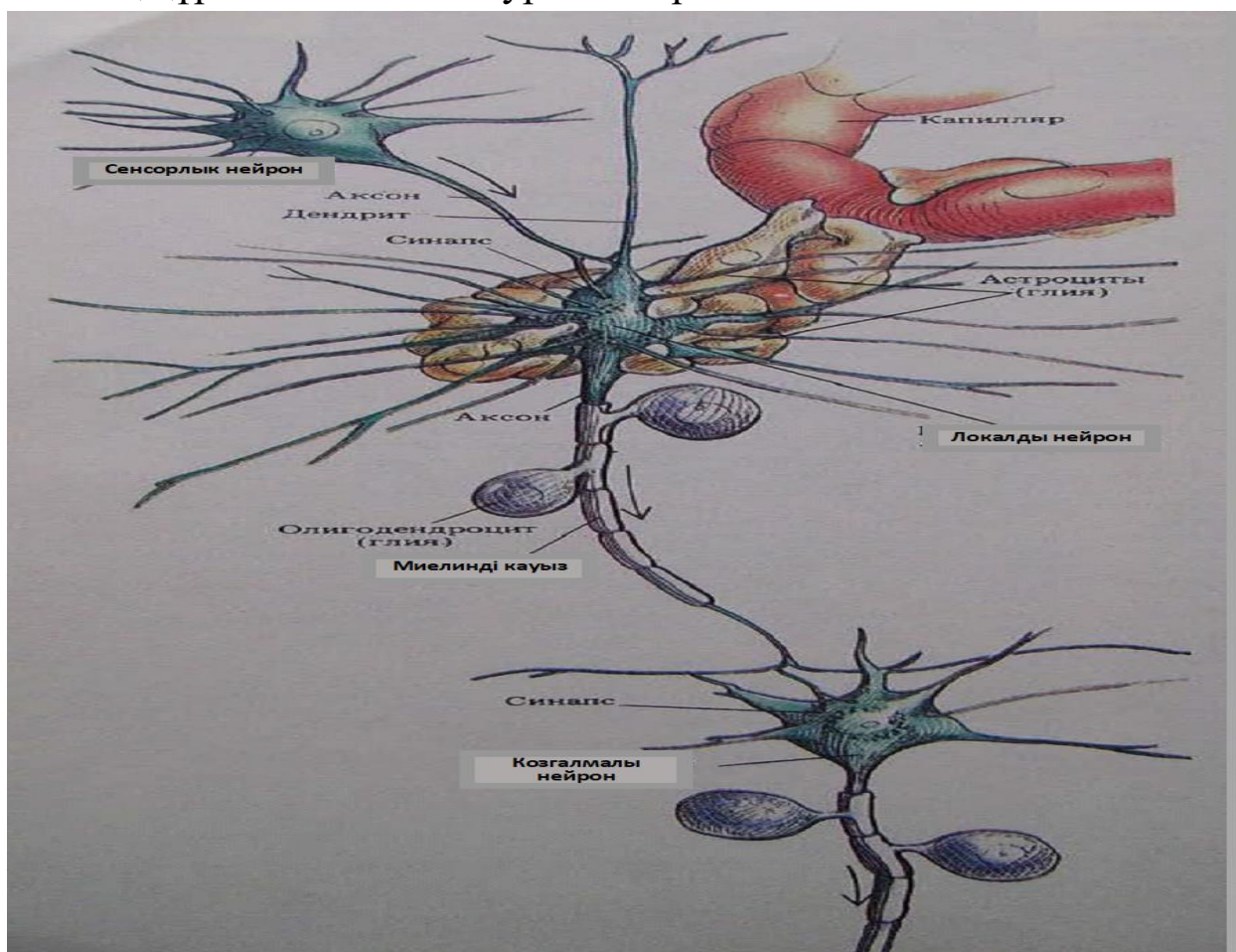
- A) нейротрансмиттер;
- B) сомалар;
- C) дендриттер;
- D) аксондар;
- E) нейрондар

II.1.2. Табиғи нейрондық желілер

Бұл тақырыпта табиғи нейрондық желілер қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Нейрондар синапстар арқылы бір-бірімен байланыса отырып, бірегей күрделі нейрондық желіні құрайды. Табиғи нейрондық желінің құрылымы I.1.2.1-суретте көрсетілген.



II.1.2.1 -сурет. Табиғи нейрондық желінің құрылымы

Тірі организмдердегі нейрондар электрлік белсенділігі бар жасушалардың ерекше түрі. Олардың негізгі функциялары организмді басқару болады.

Мидағы нейрофизиологиялық үдерістің нәтижесінде пайда болған нейрондық желідегі нейрондар – өзара электрохимиялық импульстардың сериялары арқылы әрекеттеседі. Әрбір импульс жиілігі бірнеше бірліктен жүздеген герцке дейін жететін жиілік

сигналы болады. Егер кез келген нейрофизиологиялық үдеріс бірден өзара байланысқан нейрондар жиынын белсендіретін ескерсек, онда адам миында туындайтын ақпараттар немесе сигналдар санын елестетуге болады.

I.1.2.2-суретте нейрондық желідегі электрохимиялық импульстар бейнеленген.



I.1.2.2-сурет. Нейрондық желідегі электрохимиялық импульстар

Электрохимиялық импульстар бірнеше миллисекундқа созылады да, әрбір импульсте жиілігі бірнеше бірліктен жүздеген герцке дейін жететін жиілік сигналы болады. Қазіргі заман компьютерлерімен салыстырғанда бұл өте баяу, бірақ адамның миы аналогты ақпаратты машинаға қарағанда әлдеқайда жылдам өңдейді, мысалы: бейнелерді тану, дәм сезу, дыбысты тану, бөтен жазуды ажырату. Жоғарыдағылардың барлығы да өзара синапстармен байланысқан нейрондық желі арқылы жүзеге асады. Басқаша айтқанда, ми – тізбекті есептеулерге қарағанда, әлдеқайда тиімдірек жұмыс істейтін параллелді процессорлар жүйесі. Әрқайсысы жүздеген немесе мыңдаған басқа нейрондармен байланысқан жүздеген миллиардтай нейрондардың өзі ең қуатты суперкомпьютерлерден асып түсетін жүйені құрайды. Дегенмен, адам миы өз мүмкіншіліктерінің тек қана 2-3% ғана қолданады. Осындай артықшылықтың арқасында, адамның миы зақымдалу мен жоғалтуларға қарамастан, жұмыс істеуге мүмкіндік беретін үлкен

беріктік қорға ие. Қазіргі заманғы компьютерлер осы қабілеттен айырылған. Заманауи компьютерлер тізбекті есептеулерді шешуде өзінің техникалық мүмкіншіліктерінің шегіне жетті, сондықтан қазіргі уақыттағы мәселе параллелді программалау әдістерін дамыту мен параллелді компьютерлерді құру болады. Осыған байланысты, нейрондық желілер бұл бағыттың кезекті қадамы болады. Алайда, нейрондық желілерді моделдейтін компьютерлердің табиғи нейрондық желіден айырмашылығы – моделденетін желілер белгілі бір тәртіппен реттелген және элементтері белгіленген байланыстар санына ие болады, ал табиғи желілер хаосты, құрылымы жағынан динамикалық және нейрондар арасындағы байланыс саны реттелмеген болып келеді. Табиғи нейрондар өздігінен ұйымдастырылған, бұл олардың моделденетіндерге қарағанда айтарлықтай тиімдігін арттырады. Сонымен қатар оларға мақсат қоюдың қажеті жоқ, олардың соңғысы жасушаның және бүкіл колонияның гомеостазымен анықталады.

II.1.2-тапсырмалар.

1. Нейрондық желілерді құру әдісін сипаттаңыз.
2. Адам миындағы үлкен беріктік қорының бар болу себебін түсіндіріңіз.
3. Компьютерлік технологияның дамуындағы өзекті мәселені көрсетіңіз.

II.1.2-сұрақтар.

1. Табиғи нейрондық желіде нейрондар өзара қалай байланысады?
2. Әрбір электрохимиялық импульс нені білдіреді?
3. Нейрондық желілердің даму перспективасы неде?

II.1.2-тестілер.

1. Мидағы нейрофизиологиялық үдерістің нәтижесінде пайда болған нейрондық желідегі нейрондар өзара не арқылы әрекеттеседі?

- А) электрохимиялық импульстардың сериялары;
- В) электрофизикалық импульстардың сериялары;

- С) электрлік импульстардың сериялары;
- Д) қарапайым импульстардың сериялары;
- Е) нейрондар импульстары.

2. Адам миы өз мүмкіншіліктерінің қандай пайызын ғана қолданады?

- А) тек қана 2-3% ғана;
- В) тек қана 12-13% ғана;
- С) тек қана 22-23% ғана;
- Д) тек қана 32-33% ғана;
- Е) тек қана 92-93% ғана.

3. Электрохимиялық импульстар қанша уақытқа созылады?

- А) бірнеше миллисекундқа;
- В) бірнеше миллиминутқа;
- С) бірнеше сағатқа;
- Д) бірнеше секундқа;
- Е) бірнеше минутқа.

4. Әрбір импульсте не болады?

А) жиілігі бірнеше бірліктен жүздеген герцке дейін жететін жиілік сигналы;

В) жиілігі бірнеше мыңдаған мегагерцке дейін жететін қарапайым сигналы;

С) жиілігі бірнеше бірліктен ғана тұратын жиілік сигналы;

Д) жиілігі өте төмен болатын жиілік сигналы;

Е) жиілігі болмайтын сигналдар.

5. Тізбектелген есептеулерге қарағанда, әлдеқайда тиімдірек жұмыс істейтін параллелді үдерістер жүйесі?

- А) ми;
- В) капилляр;
- С) дендрит;
- Д) аксон;
- Е) синапс.

II.1.3 Жасанды нейрондар және нейрондық желілер тарихы

Бұл тақырыпта жасанды нейрондар және нейрондық желілердің пайда болуы мен дамуының қысқаша тарихы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

1943 жылы нейро-физиолог У. Мак-Каллок және математик У. Питтс «Жүйке белсенділігіне қатысты идеяларды логикалық есептеу» мақаласында *жасанды нейрондық желі* ұғымын енгізді және жасанды нейронның формалды моделін ұсынды. 1948 жылы Н. Винер өзінің әріптестерімен бірге жарыққа шығарған кибернетика туралы жұмысында күрделі биологиялық үдерістерді математикалық моделдермен бейнелеуді ұсынған. 1949 жылы Д. Хебб «Іс-қимылды ұйымдастыру» жұмысында нейрондарды оқып үйренудің негізгі принциптерін және алғашқы алгоритмін сипаттаған.

1950 жылдары жасанды нейрондық желілердің программалық моделдері пайда болды. Бірінші жұмыстар Н. Рочестрдің басқаруымен IBM зерттеу зертханасында өткізілді. Бірақ әрі қарай жүзеге асуы жақсы болғанымен бұл жоба сәтсіз аяқталды, себебі дәстүрлі есептеу жылдам артқандықтан нейрондарды зерттеу көлеңкеде қалып қойды.

1956 жылы Дартмуттың жасанды интелекті зерттеу институты жасанды интелект, дербес жағдайда нейрондық желілер аймағындағы зерттеулердің өсуін қамтамасыз етті. Жасанды интелекті зерттеуді ынталандыру екі бағытқа: *жасанды интелект жүйесін өнеркәсіпте қолдануға* (эксперттік жүйелерге) және *миды моделдеуге бөлінді*.

1957 жылы Ф. Розенблатт мидың персептрон (латын тілінен *perception–қабылдау*) деп аталатын компьютерлік моделін ұсынды және осы жылы Корнель Аэронавтар Зертханасында IBM 704 компьютерінде персептрон жұмысын моделдеу жұмысы сәтті аяқталды. Персептрон жадының электромеханикалық ұяшықтар

блогында сенсорлық өрісті беретін фотоэлементтерден сигналдар жіберген. Бұл ұяшықтар өзара коннективизм қағидасы бойынша кездейсоқ жолмен біріктірілген.

1958 жылы Ф.Розенблатт «Персептрон: Бас миында ақпараттарды сақтаудың және ұйымдастырудың ықтималдық моделі» мақаласын жариялады, онда бір қабатты пресептрон енгізілген және оның сыныптау есептерін шешу қабілеттілігі көрсетілген. Персептрон әйгілі болды және оны бейнелерді тану, ауа райын болжау және т.б үшін қолдана бастады. Осы жылы Джон фон Нейман вакуумдық құбырды пайдаланып нейрондардың қарапайым функциясын имитациялауды ұсынды.

1959-60 жылдары Б.Уидроу және М.Хофф ADALINE және MADALINE (Жиынтық Бейімді Сызықтық Элементтер) моделдерін әзірледі. ADALINE (бейімді қосқыш) алдымен болжау есептері және бейімделген басқару үшін қолданылған, одан кейін жүйе сигналдарды өңдеудің стандартты элементі болды. MADALINE телефон желісіндегі жаңғырықты тазалайтын бейімделетін фильтр ретінде жұмыс істеді және әлі де коммерциялық қолдану үстінде.

1960 жылы Корнель университетінде ағылшын алфавитіндегі әріптерін тани алатын «Марк-1» алғашқы нейрокомпьютері көрсетілді.

1962 жылы Ф.Розенблатт өзінің теориясын және қабылдау мен персептрондар үдерісіне қатысты болжамын «Нейродинамика қағидасы: персептрондар және ми механизмі теориясы» кітабында толық сипаттап көрсетті. Ол дайын бір жасырын қабатты персептрон моделін көрсетіп қана қоймай қиылысқан және кері байланысқан көп қабатты персептронды қарастырған, сонымен қатар, персептронның қиылысу теоремасын дәлелдейді. Ол кезде бір қабатты пресептрон классикалық нейрондар жүйесінен шыққан деп саналды және екі сыныптың біреуіне кіретін сигнал сыныптауы үшін қолданылды. Бірінші рет персептрон ағылшын тілінің әріптерін тану үшін қолданылған. Ол үшін қателерді түзету әдісі арқылы – адамды оқыту үдерісін еске түсіретін, сынама-лар мен қателіктер негізінде арнайы итерациялық оқыту әдісі ойлап

табылған болатын. Сонымен қатар, персептрон әріпті тану барысында оның ерекшеліктерін де бөліп көрсететін. Осылайша персептрон әр түрлі (жазбамен) жазылған жазбаларды бір жалпы үлгіге жинақтай алатын әріптермен жалпылау қабілетіне ие болған.

1963 жылы КСРО ҒА ақпаратты жіберу мәселесі Институтында А.П. Петров персептрондар үшін «қиын» тапсырмаларды толық зерттеу жұмыстарын жүргізді. КСРО-да жасанды нейрондық желіні (ЖНЖ) моделдеу саласындағы бұл жаңашыл жұмыс белгілі ғалым М.М. Бонгардтың – «салыстырмалы түрде персептронға шағын өзгертулер жасау арқылы оның кемшіліктерін түзету» сияқты ойлар кешені үшін бастама нүктесі болды.

Дамып келе жатқан академиялық және техникалық әлемде нейрондық желілердің аса қатты дамуына көңіл бөлінуі, сол заманның әдебиетін де таң қалдырды. «Ойлайтын машиналар» адамдар үшін қаупі бар екендігін жазушылар үнемі айтып отырған. Алайда, персептронның мүмкіндіктері тым шектеулі еді. Машина жартылай жабық әріптерді, сондай-ақ өзге мөлшермен берілген әріптерді, бұрылған әріптерді немесе басқаша орналасқан әріптерді анық тани алмады.

1969 жылы М.Минский және С. Пейперт персептронның көптеген қарапайым тапсырмаларды, соның ішінде бейнелеудің инварианттылығымен байланысқан «жұптық» немесе «блокта жалғыз» проблемасын, «ерекшелейтін NEMECSE – eXclusive OR (XOR)» бинарлы функциясын жүзеге асыра алмайтындығын теориялық тұрғыдан дәлелдеді. Бұл факт және уәдесін орындай алмаумен байланысқан қауіптену, нейрондық желілерді зерттеуді сынаған көптеген мамандардың көңілін қалдырды. Нәтижесінде, нейрондық желілерді зерттеуге қызығушылық күрт төмендеп, оның ары қарай дамуына қаражат бөлу де тоқтатылды.

1972 жылы Т.Кохонен мен Дж.Андерсен жады ретінде жұмыс істеу қабілеті бар жаңа нейрондық желілер типін ұсынды.

1973 жылы Б.В.Хакимов сплайндар негізінде құрылған сызықты емес синапс моделін ұсынып, медицина, геология және экология саласындағы тапсырмалады шешу үшін оны жетілдірді.

1974 жылы Пол Дж.Вербос және А. И. Галушкин көп қабатты персептрондарды үйрету үшін қателерді кері тарату алгоритмін бір мезетте ойлап табады.

1975 жылы Фукусима бейнелерді инвариантты тану үшін арналған, өздігінен жинақталатын желі болып табылатын когнитронды ұсынды. Дегенмен, бұл желі тек барлық бейнелерді есте сақтау көмегімен жүзеге асқан болатын.

1982 жылы бірнеше оқиғалар нейрондық желілерге қызығушылық туғызды. Дж. Хопфилд АҚШ ұлттық Ғылым Академиясында өзінің мақаласын таныстырды. Ол кері байланысы бар нейрондық желі энергияны азайтайтын өзіндік жүйе құрайтынын көрсетті, яғни жаңа құрылым негізінде нейрондық желі моделінің мүмкіндіктерін көрсетті және оларды Хопфилд желісі деп атады. Осы сәтте, Киотода (Жапония) нейрондық желілер бойынша Біріккен америкалық-жапондық конференция болды, онда желілер бесінші тудырудың жетістігі болып жарияланды. Америкалық басылымдар осы аталған оқиғаны көтерді, олардың ойынша АҚШ бұл салада артта қалу мүмкіндігі бар екенін айтқаннан кейін ғана, нейрожелілер саласының әрі қарай зерттелуіне қаражат бөлініп, дамуына алып келді. Содан кейін Кохонен «Кохоненнің нейрондық желісі» деп аталатын желі моделін ұсынды. Бұл желі кластерлеу есептерін, деректерді визуализациялау және деректерді талдаудың басқа да есептерін шешуді мұғалімсіз ақ үйретеді.

1985 жылы Американдық Физика Институты – "Нейрондық желілер есептеулер үшін " атты жыл сайынғы кездесуін бастады.

1986 жылы Д. Румельхарт, Дж. Хинтон және Р. Вильямс, сонымен қатар, олардан бөлек әрі бір уақытта С. И. Барцев және В. А. Охонин қателерді кері тарату әдісін біршама дамытып, қайтадан ашты. Үйрететін нейрондық желілерге қызығушылық кенеттен арта түсті.

1989 жылы "Нейрондық желілер қорғаныс үшін" атты кездесуінде Б. Уидроу ұрыс алаңы ретінде әлемдік нарық және

өндіріс болып табылатын төртінші дүние жүзілік соғыс басталғаны туралы хабарлады.

1990 жылы Шағын бизнесті қорғау бағдарламаларының инновациялық зерттеулер департаменті нейрондық желілерді қолданатын 16 негізгі және 13 қосымша тақырыптарды атады.

1996 жылы Accurate Automation Corp, Chattanooga, TN фирмасы NASA мен Air Force тапсырмасы бойынша тәжірибелі автопилотты гипердыбыстық LoFLYTE (Low-Observable Flight Test Experiment) барлау ұшағын жасады. Ұшақ ұзындығы 2,5 м., салмағы 32 кг. және ұшудың жаңа принципіні зерттеуге арналған. LoFLYTE ұшқыштың ұшу тәсілдерін көшіру арқылы автопилотты оқытатын нейрондық желілерді пайдаланды. Ұшақ 4-5 мах жылдамдықпен ұшуға арналған болғандықтан, ұшу режимін дәлмедәл өзгертуге ұшқыш-адамның реакциясының жылдамдығы жеткіліксіз болуы мүмкін.

Бұл жағдайда ұшқыштың басқару тәжірибесін қабылдайтын және ақпаратты жылдам өңдейтіндіктен апатты және экстремалды жағдайлардан тез шығатын нейрондық желілер көмекке келеді.

2007 жылы Торонто университетінде Джеффри Хинтон көп қабатты нейрондық желілерді тереңдей үйрететін алгоритмді құрды. Хинтонның төменгі қабат желілерін үйретуде, Больцманның шектелген машинасын (RBM – Restricted Boltzmann Machine) қолдануы үлкен жетістікке алып келді. Хинтон бойынша терең үйрету дегеніміз – бұл өте баяу үдеріс. Бейнелерді тану үшін көптеген үлгілер қажет (мысалы, әр түрлі фондағы адам кескінің бейнелері). Үйретуден кейін нақты тапсырманы (мысалы, суреттегі тұлғалардың кескінін іздеуді) орындай алатын, жылдам жұмыс істейтін қосымша дайын болады. Қазіргі таңда адам бейнесін іздеу функциясы стандартқа айналып, барлық заманауи цифрлық фотоаппараттарға ендірілген. Терең оқыту технологиясы суреттерді олардың құрамындағы үлгілер бойынша сыныптау барысында интернет-іздістірушілермен белсенді қолданылады. Бейнелерді тану үшін қолданылатын жасанды нейрондық желілерде – 9 қабат

нейрон болуы мүмкін. Оларды оқыту миллиондаған суреттерде бейнелерді іздеу арқылы жүргізіледі.

Қазіргі кезде нейрондық желілер барлық жерде талқыланады. Оларды пайдалану перспективасы өте жарқын, себебі қалыпты емес проблемаларды шешуде өте ыңғайлы. Қазіргі уақытта көптеген нейрондық желі әзірлемелері жұмыс істейді, бірақ өзіндік үдерістік шектеулер бар. Нейрожелілердің зерттеу бағыттары программалық және аппараттық жүзеге асыруларға бағытталған. Компаниялар жақын болашақтың тренді болатын үш түрлі нейрочип құруға ұмтылуда: цифрлық, аналогты және оптикалық.

Тұтастай алғанда, нейрондарды функцияналдау түсінігі зерттеушілер үшін өздерінің теориясын тексеретін математикалық модел құруға мүмкіндік берді. Ендігі кезекте тәжірибелер, адамдар немесе жануарлардың көмегінсіз, цифрлық компьютерлерге жасалынады. Бұл көптеген практикалық және моралды-этикалық проблемаларды шешеді. Алғашқы жұмыстардан ақ, бұл моделдердің тек қана мидың функцияларын қайталап қоймай, сондай-ақ өзіндік құндылығы бар қызметтерді де орындауға қабілетті екені айқындалды.

Нейрондық моделдеудің екі мақсаты бар:

- 1) біріншісі—физиологиялық және психологиялық деңгейде адам жүйке жүйесінің жұмыс істеуін ұғыну;
- 2) екіншісі—ми функциясына ұқсас есептеу жүйесін (жасанды нейрондық желілер) құру.

II.1.3-тапсырмалар.

1. Жасанды нейрон ұғымын енгізген және оның формалды моделін ұсынған авторды атап шығыңыз.
2. Нейрондарды оқып үйренудің негізгі қағидаларын және алғашқы алгоритмдерін сипаттаған авторды атап шығыңыз.
3. Персептрон көмегімен шешілген тапсырмаларды атаңыз.
4. XOR функциясын жүзеге асыра алмаған персептрондардың шектеулігін дәлелдеген авторды атаңыз.

5. Бейнелерді инвариантты тану тапсырмаларын шешу үшін когнитронды ұсынған авторды атаңыз.

6 Көп қабатты нейрондық желілер үшін терең оқыту алгоритмін құрастырған авторды атаңыз.

II.1.3-сұрақтар.

1. Бейімделген сызықты элементтердің моделдері қалай әрекет етеді?

2. Не себептен және қай уақытқа дейін жасанды нейрондық желілерге қызығушылық жоғалып кетті?

3. Нейрондық желілерге деген қайта қызығушылық тудыруға не әсер етті?

4. Көп қабатты персептрондарды үйрету үшін қателерді кері жіберу алгоритмін кім ойлап тапты?

5. Кім және қашан қателерді кері жіберу әдісін дамытып, қайта ашты?

6. Көп қабатты персептрондарды терең үйрету алгоритмі қалай жұмыс істейді?

7. Жұмыс істеп тұрған нейрондарға қандай шектеулер қойылады?

II.1.3-тестілер.

1. 1973 жылы жасанды нейрондық желі ұғымын енгізген және жасанды нейронның формалды моделін ұсынған ғалымдар?

А) У. Мак-Каллок және У. Питтс;

В) Н. Винер және Д. Хебб;

С) Н. Рочестр және Ф. Розенблатт;

Д) Б.Уидроу және М.Хофф;

Е) А.П. Петров және М.М. Бонгард.

2. Алғашқы жұмыстары Н. Рочестрдің басқаруымен IBM зерттеу зертханасында өткізілген жасанды нейрондық желілердің программалық моделдері қай жылдары пайда болды?

А) 1950 жылдары;

- B) 1940 жылдары;
- C) 1930 жылдары;
- D) 1920 жылдары;
- E) 1970 жылдары.

3. Мидың персептрон деп аталатын компьютерлік моделін қай ғалым, қай жылы ұсынды?

- A) 1957 жылы Ф. Розенблатт;
- B) 1956 жылы Дартмут;
- C) 1950 жылы Н. Рочестр;
- D) 1949 жылы Н. Винер;
- E) 1943 жылы У. Мак-Каллок.

4. 1972 жылы жады ретінде жұмыс істеу қабілеті бар жаңа нейрондық желілер типін ұсынған ғалымдар?

- A) Т.Кохонен мен Дж.Андерсен;
- B) Пол Дж.Вербос және А. И. Галушкин;
- C) Дж. Хопфилд және Н. Рочестр;
- D) Дж. Хинтон және Р. Вильямс;
- E) С. И. Барцев және В. А. Охонин.

5. 2007 жылы Торонто университетінде көп қабатты нейрондық желілерді тереңдей үйрететін алгоритмді құрған кім?

- A) Джеффри Хинтон;
- B) А. И. Галушкин;
- C) Дж. Хопфилд;
- D) Р. Вильямс;
- E) С. И. Барцев.

II.1.4. Жасанды нейронның формалды моделі

Бұл тақырыпта алғашқы жасанды нейронның құрылымдық сұлбасы мен жұмыс алгоритмі, формалды (математикалық) моделі қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

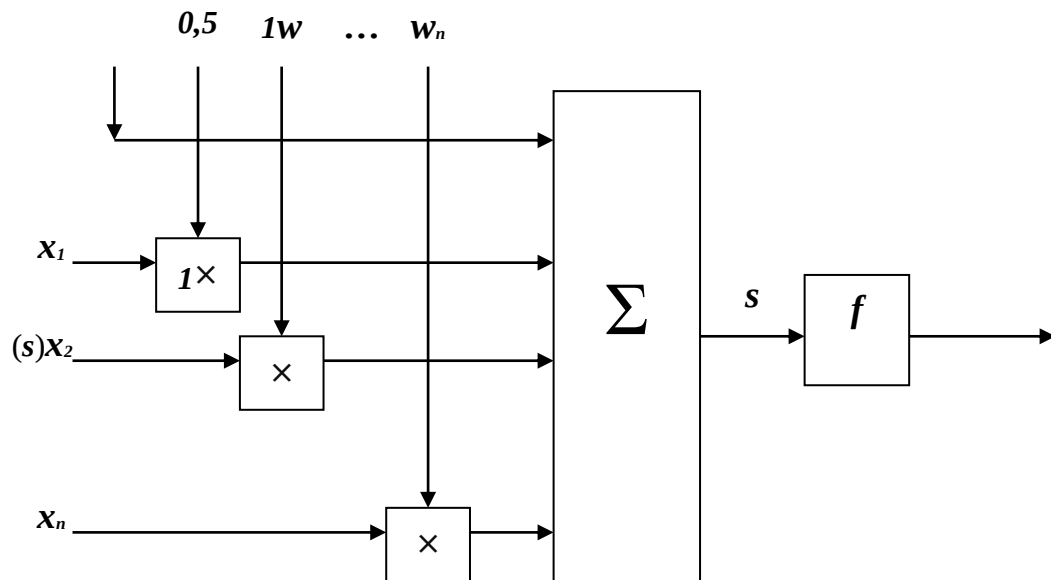
Өткен тақырыпта көрсетілгендей, табиғи (биологиялық) нейронның жұмыс істеу алгоритмін орындауды ұқсата алатын (имитациялайтын) жасанды нейронның электрондық сұлбадағы бірінші моделін ғалымдар Уоррен Мак-Каллок (Warren McCulloch) пен Уолтер Питтс (Walter Pitts) ұсынған.

Жасанды нейрон дегеніміз – үш (*көбейткіш, қосқыш, түрлендіргіш*) блоктан тұратын және кіру сигналдарын алғашқы күйдің мәндерін ескере отырып синапстық байланыстар салмағы мен белсендіру функциясына сәйкес келетін шығу сигналдарына өңдеуге мүмкіндік беретін абстрактылы компьютер.

Жалпы жағдайда жасанды нейронда $n \geq 1$ кірістер бар және бұл кірістер синапстарына кіру сигналдарының векторы $X = (x_1, x_2, \dots, x_n)$ келеді, сигналдар синапстан өткеннен кейін синапстың салмақ векторы $W = (w_1, w_2, \dots, w_n)$ мәндеріне пропорционалды түрде өзгереді, яғни, $W \cdot X = (w_1 \cdot x_1, w_2 \cdot x_2, \dots, w_n \cdot x_n)$ болады. Өзгерген кіру сигналдары дендриттер арқылы сомаға келеді, онда *белсендіру деңгейі* деп аталатын қосынды импульс $s = w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n$ пайда болады. Белсендіруді *белсендіру деңгейі* деп аталатын арнайы функция f шақырады. Белсендіру функциясы әсер еткеннен кейін белсендіру деңгейі кейбір *меже (ығысу)* w_0 мәнімен салыстырылады. Егер белсендіру деңгейінің мәні осы меже мәнінен артып кетсе, онда нейрон қозады және y шығу сигналын қалыптастырады.

Жалпы жағдайда кіру сигналының мәні, салмағы және ығысуы нақты сан болып келеді. Шығу сигналының мәні белсендіру функциясының түрімен анықталады және нақты сан да, бүтін сан да бола алады.

Мак-Каллок пен Питтстің жасанды нейронның құрылымдық сұлбасы II.1.4-суретте көрсетілген.



II.1.4-сурет. МакКаллок және Питтстің жасанды нейронының құрылымдық сұлбасы

Мак-Каллок пен Питтстің жасанды нейроны төмендегі алгоритм бойынша жұмыс істейді:

1. Жұмыс басталмас бұрын қосқыш блогына бастапқы күй сигналының мәні (ығыстырылу – bias) w_0 беріледі;

2. Көбейткіш блогындағы әрбір x_I кіру сигналының мәні сәйкес келетін w_I синапстың салмағына көбейтіледі, $I=1, 2, \dots, n$;

3. Қосқыштағы барлық салмақталған кіру сигналдарының мәндері (кіру сигналдарының мәндері мен салмақтардың көбейтіндісі) және бастапқы күй сигналының мәні қосылып, нейронның белсендіру деңгейін $\sum_{i=1}^n w_i x_i + w_0 = W \cdot X + w_0 = s$ анықтайды.

4. Егер белсендіру функциясы f бар болса, онда s салмақталған қосынды берілген меже мәні ескеріліп шығу сигналының мәні өндірілетін f функционалды түрлендіру блогына беріледі және $y = f(x)$ болады, ал кері жағдайда, шығу сигналы s алдыңғы қадамда алынған салмақталған қосындыға тең, яғни $y = s$.

Сонда Мак-Каллок пен Питтстің жасанды нейронының математикалық моделін былай жазуға болады:

$$y = f\left(\sum_{i=1}^n w_i x_i + w_0\right) = f(W \cdot X + w_0) = f(s),$$

мұнда x_i – кіру сигналы, w_0 – бастапқы күйдің сигналы, w_i – синапс салмағы, s – қосынды, y – шығу сигналы, f – белсендіру функциясы, n – кіріс (салмақ) саны, $i=1,2,\dots, n$.

Егер салмақ мәні w_i оң (*теріс*) болса, онда i -інші синапстық байланыс *қоздырғыш* (*тежегіш*) болады.

Жасанды нейронның көмегімен нақты есепті шығару үшін кіру деректерінің мәндерін *нормалдау* керек.

Кіру деректерінің мәндерін *нормалдау* – бұл, барлық кіру деректерін «теңелдіру» үдерісінен өткізу, яғни $[0,1]$ немесе $[-1,1]$ интервалына келтіру.

Егер нормалдауға келтірілмесе, онда кіру деректері нейронға қосымша ықпал етіп, шешімнің дұрыс шықпауына әкеліп соғады. Басқаша айтқанда, әртүрлі реттегі шамаларды қалай салыстыруға болады?

Жалпы айтқанда нормалдау формуласы келесі түрдегідей:

$$x_{norm} = \frac{(x - x_{min}) \cdot (d_2 - d_1)}{x_{max} - x_{min}} + d_1 \quad (\text{II.1.})$$

мұнда x – қалпына келтіретін мәні, $[x_{min}, x_{max}]$ – x мәндерінің интервалы, $[d_1, d_2]$ – x мәндері келтірілетін интервал.

Айтылғанды мысал арқылы дәлелдейік. $[0,10]$ интервалынан n кіру деректері бар делік. Сонда $x_{min} = 0, x_{max} = 10$.

Деректерді $[0,1]$ интервалына әкелеміз, онда $d_1 = 0, d_2 = 1$ болады. Барлық деректердің мәндерін (II.1.) формуласына салып, n – кіру деректерінен тұратын кез келген x үшін x_{norm} нормалданған мәндерді есептеуге болады.

Алдымен шығу мәні үшін шеткі мәндерін көрсетеміз, сонан кейін индикатордың минималды және максималды мәндерін аламыз (индикатордан деректер көшіру көрсетілмеген, мысалы, ол жерде соңғы 10 мәні болуы мүмкін). Соңғы қадамда әрбір кіру элементін (әртүрлі бардағы индикатордың мәнін) циклде

нормалдаймыз және ары қарай жұмыс істеу үшін оны массивте сақтаймыз.

Мак-Каллок-Питтс моделінде кіру сигналдарының уақытша тежелуі жоқ, сондықтан s -тің мәні нейрон қабылдаған сыртқы қозуды толық анықтайды. Нейронның әрекеті әрі қарай «бәрі немесе ештеңе» принципі бойынша сипатталады, яғни, айнымалы сызықсыз межелік түрленуге ұшырайды. Мұнда, егер салмақталған қосынды s межелі мәнінен көп болса, онда шығыс (нейронның белсену жағдайы) y бірге тең болып орнатылады, ал кері жағдайда шығыс y нөлге тең болады. Меженің мәні (көбінесе нөлге тең) γ локалды жадыда сақталады. Сөйтіп, Мак-Каллок-Питтс моделінде мынадай межелі белсендіру функциясы қолданылады:

$$y = f(s) = \begin{cases} 1, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}$$

Салмақ коэффициенттерін талдауға негізделген нейрондарды оқыту әдістерін таңдау үшін белсендіру функциясының қасиеттері шешуші ықпал етеді. Оқытудың екі жолын ерекше бөліп көрсетеді: *мұғаліммен оқыту (supervised learning)* және *мұғалімсіз оқыту (unsupervised learning)*.

Мұғаліммен оқыту кезінде x_i кіру сигналынан басқа, нейронның күтілетін d_i шығу сигналы да мәлім болуы мүмкін деп есептеледі. Салмақты таңдау нейронының нақты шығу сигналы y_i мәні нейронының күтілетін шығу сигналы d_i мәніне неғұрлым жақын болуы ретінде ұйымдастырылады.

Егер осындай тәсіл мүмкін болмаса мұғалімсіз оқыту стратегиясы қабылданады. Бұл жағдайда салмақ коэффициенттерін таңдау нейрондардың өзара бәсекелестігі немесе оқушылар мен шығу сигналдарының корреляциясын ескеру негізінде жүргізіледі. Мұнда нейрондардың бейімделу кезеңінде оның шығу сигналдарын болжау мүмкін емес.

II.1.4-суретте келтірілген құрылым жасанды нейронның стандартты моделі болады. Математикалық моделдердің көпшілігі биологиялық нейрондардың қызмет ету принципіне

сәйкес жасалған және Мак-Каллок-Питтс стандартты моделіне негізделген, мысалы, *персептрон*, *сигмоидалды нейрон*.

Кейінірек жасанды нейрон әзірлеушілер жасанды нейронның бірнеше стандартты емес детерминделген моделін (*наде-нейрон*, *квадрат қосқышты нейрон*, *бейімделген сызықты нейрон*, *сигма – пи нейрон*, *WTA*, *Хебб*, *Горссберг*, *есептегіші сәйкесті нейрон* және т.б.) ұсынды, сондай-ақ, олар нейронның стохастикалық (ықтималдық) моделін ұсынды.

Одан басқа, егер кіру сигналының, синапс салмағының, ығысуының мәндері мынадай болса:

1) кез келген нақты сан және оған нақты арифметика қолданылса, онда *нақты нейрон* туралы айтқанымыз;

2) логикалық мән 0 немесе 1 болса және оған логикалық амалдар қолданылса, онда *логикалық нейрон* туралы айтқанымыз;

3) $[0,1]$ интервалындағы сан болса және оған бұлдыр логикасының (*fuzzy logic*) амалдар қолданылса, онда *бұлдыр нейрон* туралы айтқанымыз болады.

Иллюстрация үшін нақты мысал қарастырайық. Мейлі нейрон өзінің кірісінде деңгейі X векторына тең сигнал алсын. Оған сәйкес синапстық байланыс салмағы W векторына тең болсын. Сонда жартылай сызықты $\alpha = 3$ нейроны үшін j нейронының шығу деректерінің мәнін есептеу керек.

Берілгені:

$X = (1,3; 2,3; -3); W = (0,2; 0,6; -0,02), w_0 = -1,2$

Белсендіру функциясы:

$$y = f(s) = \begin{cases} \alpha s, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}$$

Шешуі:

Нейронның жұмыс істеуін төмендегі қосақ теңдеумен сипаттауға болады:

$$s = \sum_{i=1}^n w_i x_i, \quad y = f(s + w_0),$$

мұнда $X = (x_1, x_2, \dots, x_n)$ – кіру сигналдары; $W = (w_1, w_2, \dots, w_n)$ – нейронның синапстық байланыстар салмағы; w_0 – меже; s – нейронның салмақталған сомасы; $f(s)$ – нейронның белсендіру функциясы; y – нейронның шығу сигналы.

Бізде бары:

$$x_1 = -1,3; x_2 = 2,3; x_3 = -3; w_1 = 0,2; w_2 = 0,6; w_3 = -0,02; w_0 = b = -1,2.$$

$$s = \sum_{i=1}^n w_i x_i = x_1 w_1 + x_2 w_2 + x_3 w_3 = -1,3 \cdot 0,2 + 2,3 \cdot 0,6 + (-3) \cdot (-0,02) = 1,18$$

$$y = f(s+b) = f(1,18-1,2) = f(-0,02) = 0.$$

Жауабы: $y = 0$

II.1.4-тапсырмалар.

1. Жасанды нейронның құрамындағы блоктарын көрсетіңіз.
2. Жасанды нейронның жұмысын сипаттаңыз.
3. Мак-Каллок және Питтс нейронының жұмыс алгоритмін сипаттаңыз.
4. Мак-Каллок және Питтс нейронының математикалық моделін жазыңыз.
5. Кіру деректерінің нормалдау формуласын жазыңыз.
6. Жасанды нейронның жұмыс істеу формуласын жазыңыз.
7. Белсендіру функциясының мысалын келтіріңіз.

II.1.4-сұрақтар.

1. Кіру сигналдары, салмақ пен ығысу қандай мәнге ие болуы мүмкін?
2. Кіру сигналдарын нормалдау деген не?
3. Қосқыш блогына қызметін бастамас бұрын не беріледі?
4. Қандай жағдайда синапстық байланыс қозған немесе тежелген күйде болады?
5. Белсенділік қызметі қандай рөл атқарады?
6. Қалыпқа келтірілмеген ену деректерін беру неге әкеліп соғады?

7. Жасанды нейронның стандартты емес детерминделген моделді қайсысы әзірлеушілерге кейінірек ұсынылды?

II.1.4-тестілер.

1. Табиғи (биологиялық) нейронның жұмысын имитациялайтын жасанды нейронның электрондық сұлбадағы алғашқы моделін ұсынған ғалымдар кім?

A) У. Мак-Каллок (Warren McCulloch) және У. Питтс (Walter Pitts);

B) Н. Винер және Д. Хебб;

C) Н. Рочестр және Ф. Розенблатт;

D) Б. Уидроу және М. Хофф;

E) А.П. Петров және М.М. Бонгард.

2. Мак-Каллок пен Питтстің жасанды нейронының математикалық моделін көрсететін формула?

A) $y = f\left(\sum_{i=1}^n w_i x_i + w_0\right) = f(W \cdot X + w_0) = f(s);$

B) $y = f\left(\sum_{i=1}^n w_i x_i\right) = f(s);$

C) $y = f\left(\sum_{i=1}^n x_i w_0\right) = f(X + w_0) = f(s);$

D) $y = f(W \cdot X);$

E) $y = f\left(\sum_{i=1}^n w_i x_i + w_i x_i\right).$

3. Мыналардың қайсысы нормалдау формуласы болады?

A) $x_{norm} = \frac{(x - x_{min})(d_2 - d_1)}{x_{max} - x_{min}} + d_1;$

B) $x_{norm} = \frac{(x - x_{min})}{x_{max} - x_{min}} + d_1;$

$$C) x_{norm} = \frac{(d_2 - d_1)}{x_{max} - x_{min}} + d_1;$$

$$D) x_{norm} = \frac{(x - x_{min}) \cdot (d_2 - d_1)}{x_{max} - x_{min}};$$

$$E) x_{norm} = \frac{(x - x_{min})}{x_{min}} + d_1.$$

4. Нейрондарды оқытудың қандай екі жолы бар?

- A) мұғаліммен және мұғалімсіз оқыту;
- B) пайдаланушымен және пайдаланушысыз оқыту;
- C) бағдарламамен және бағдарламасыз оқыту;
- D) аппаратпен және аппаратсыз оқыту;
- E) қызметшімен және қызметшісіз оқыту.

5. Кіру деректерінің мәнін нормалдау деген не?

- A) барлық кіру деректерін «теңелдіру» үдерісінен өткізу үрдісі;
- B) барлық кіру деректерін «жою» үдерісінен өткізу үрдісі;
- C) барлық кіру деректерін «арттыру» үдерісінен өткізу үрдісі;
- D) барлық кіру деректерін «кеміту» үдерісінен өткізу үрдісі;
- E) барлық кіру деректерін «терістеу» үдерісінен өткізу үрдісі.

II.1.5. Жасанды нейрондарды сыныптау

Бұл тақырыпта әртүрлі критерийлеріге байланысты жасанды нейрондарды сыныптау қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Жасанды нейрондарды жүйедегі орнына, қолданылатын деректердің типіне, нейрондардың аппараттық жүзеге асыру тәсіліне, нейрон моделін сипаттау тәсіліне, салмағының өзгеруіне және шығу сигналының ықтималдық мәніне байланысты сыныптауға болады.

Жасанды нейрондар жүйедегі орнына (рөліне) байланысты *рецепторлық (кіріс) нейрондар, аралық нейрондар және эффекторлық (шығыс) нейрондар* болып бөлінеді.

Рецепторлық нейрондар бастапқы сигналдарды кодтайтын кіріс векторларын қабылдайды және олармен амалдар орындамайды, ол жай ғана кодталған бастапқы сигналды күшейтіп немесе бәсеңдетіп шығу жолына жібереді.

Аралық нейрондар бастапқы және аралық сигналдарға байланысты негізгі амалдарды орындайды.

Эффекторлық нейрондары желінің шығу сигналдарын ұсынады. Шығу нейрондарында шығу сигналдарымен қандай да бір амалдар орындалуы мүмкін.

Жасанды нейрондар өңделген деректердің типіне байланысты *сандық нейрондар және логикалық нейрондар* болып бөлінеді.

Сандық нейрондар нақты сандар түрінде берілген сандық деректермен амалдарды орындайды. Үлгілерді танудың қандай да бір есептерін шешу кезінде бүтін сандарды қолдануға рұқсат беріледі.

Логикалық нейрондар логикалық деректермен амалдар орындайды. Егер логикалық деректердің мәні 0 немесе 1 болса, онда классикалық (бульдік) логика амалдары қолданылатын *екілік логикалық нейрон* туралы айтылып отыр. Егер логикалық деректердің мәні $[0,1]$ интервалында болған жағдайда, онда біз

бұлдыр логиканың амалдарын пайдаланатын *бұлдыр логикалық нейрондармен жұмыс жасаймыз.*

Жасанды нейрондар олардың аппараттық жүзеге асырылу тәсіліне байланысты *цифрлық және аналогтық* болып бөлінеді.

Цифрлық нейрондар интеграциясы орташа және жоғары деңгейлі микросұлбалармен жүзеге асырылады, бұл кезде кристалда тікелей нейрондық желі түзіледі, яғни жеке нейрондар өзбетінше бірлік ретінде ерекшеленбейді, себебі цифрлық модел нақты есепті шешу үшін әзірленеді, ал бір нейрон көмегімен оны шешу мүмкін емес.

Аналогтық нейрон қоздырушы сигналдардың кеңістіктік және уақыттық қосындысын, абсолюттік және салыстырмалы қасарыс (қабылдамаушылық) қасиеттерін, жеке нейронмен және нейрондардың әрекеттесуімен ақпараттарды өңдеу үдерісін жаңғыртуға мүмкіндік береді. Аналогтық модел күйі толық тыныштықтан қозудың ең жоғарғы деңгейіне шейін үздіксіз өзгере алатын құрылғы ретінде қарастырылады.

Жасанды нейрондар олардың моделдерін сипаттау тәсілдеріне қарай нейронның *формалды моделі, физиологиялық моделі және феноменологиялық моделі* болып бөлінеді.

Нейронның формалды моделі жақсы әзірленген математикалық аппаратымен ерекшеленеді, бірақ өзінің биологиялық прототипінің көптеген қасиеттерін жоққа шығарады.

Нейронның физиологиялық моделі нейронның мінез құлқын сандық тұрғыдан сипаттаумен ерекшеленеді, олар биологиялық нейрондарға тәжірибе жасау арқылы пайда болған.

Нейронның феноменологиялық моделінің қатаң математикалық және тәжірибелік базасы жоқ.

Нейронның формалды моделі, өз кезегінде, кіру салмағының өзгергіштігіне байланысты, сонымен қатар белсендіру функциясының түріне және шығу сигналының мүмкін болатын мәніне байланысты бөлінуі де мүмкін.

Салмағының өзгеруіне байланысты нейрондар *кіріс салмағы өзгермеген нейрондар, кіріс салмағын басқаруға болатын*

нейрондар, берілген арна бойынша кіру сигналына байланысты қайта құрылатын салмағы бар нейрондар, шығу сигналының әсерімен өзгермелі салмағы бар нейрондар болып бөлінеді.

Шығу сигналының ықтималдық мәніне байланысты нейрондар *нейронның детерминделген моделі және нейронның стохастикалық моделі* болып бөлінеді.

Әрі қарай біз тек нейронның формалды моделін қарастырамыз.

II.1.5-тапсырмалар.

1. Желідегі жағдайына байланысты нейрондық желілердің сыныпталуын атап шығыңыз.

2. Жүзеге асырылатын (өңделетін) деректердің типі бойынша нейрондық желілердің сыныпталуын атап шығыңыз.

3. Аппараттық жүзеге асырылуына байланысты нейрондық желілердің сыныпталуын атап шығыңыз.

4. Олардың моделдерін сипаттау тәсілдеріне байланысты нейрондық желілердің сыныпталуын атап шығыңыз.

5. Кірістік салмағының өзгеруіне, белсендіру функциясының түріне және шығу сигналының ықтималдық мәніне байланысты нейрондық желілердің сыныпталуын атап шығыңыз.

II.1.5-сұрақтар.

1. Кірістік нейрондары, аралық және шығыстық нейрондары қандай критерийлері бойынша бөлінеді?

2. Кірістік нейрондар не істейді?

3. Аралық нейрондар не істейді?

4. Шығыстық нейрондар не істейді?

5. Сандық нейрон мен логикалық нейрондары қандай критерийлері бойынша бөлінеді?

6. Сандық нейрондар қандай сандармен жұмыс істейді?

7. Логикалық нейрондар қандай шамалармен жұмыс істейді?

8. Формалды модел, физиологиялық модел және феноменологиялық модел қандай критерийлері бойынша бөлінеді?

9. Аппараттық жүзеге асырылуына байланысты жасанды нейрондар қалай бөлінеді?

II.1.5-тестілер.

1. Жасанды нейрондар жүйедегі орнына (рөліне) байланысты қалай бөлінеді?

- А) кірістік нейрондар, аралық нейрондар және шығыстық нейрондары;
- В) сандық нейрондар және логикалық нейрондар;
- С) цифрлық және аналогтық;
- Д) детерминделген және стохастикалық;
- Е) өзгермелі және өзгермейтін.

2. Жасанды нейрондар өңделетін деректердің типіне байланысты қалай бөлінеді?

- А) сандық нейрондар және логикалық нейрондар;
- В) кірістік нейрондар, аралық нейрондар және шығыстық нейрондары;
- С) цифрлық және аналогтық;
- Д) детерминделген және стохастикалық;
- Е) өзгермелі және өзгермейтін.

3. Жасанды нейрондар олардың аппараттық жүзеге асырылуына байланысты қалай бөлінеді?

- А) цифрлық және аналогтық;
- В) кірістік нейрондар, аралық нейрондар және шығыстық нейрондары;
- С) сандық нейрондар және логикалық нейрондар;
- Д) детерминделген және стохастикалық;
- Е) өзгермелі және өзгермейтін.

4. Жақсы әзірленген математикалық аппаратымен ерекшеленетін, бірақ өзінің биологиялық прототипінің көптеген

қасиеттерін жоққа шығаратын жасанды нейронның моделі қалай аталады?

- A) формалды;
- B) физиологиялық;
- C) феноменологиялық;
- D) детерминделген;
- E) стохастикалық.

5. Биологиялық нейрондарға тәжірибе жасау арқылы пайда болған, нейронның мінез құлқын сандық тұрғыдан сипаттаумен ерекшеленетін нейронның моделі қалай аталады?

- A) физиологиялық;
- B) формалды;
- C) феноменологиялық;
- D) детерминделген;
- E) стохастикалық.

II.1.6. Нейрондық желілерді сыныптау

Бұл тақырыпта нейрондық желілерді сыныптау жайында жалпы мағлұмат қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Жасанды нейрондары желісі (ЖНЖ, нейрожелі) – өзара бір-бірімен байланысқан нейрондар жиынтығы. Қағида бойынша, барлық нейрондардың белсендіру функциясы желіде бекітілген, ал салмақтары өзгере алатын желі параметрі болады. Нейрондардың кейбір кірістері желінің сыртқы кірістері ретінде, ал кейбір шығыстары сыртқы шығыстары ретінде белгіленген. Нейрондық желілердің кірістеріне кез келген санды (сигналды) беріп, желі шығысында біз қандай-да бір сандар (сигналдар) жиынтығын аламыз.

Осылайша, нейрондық желілердің жұмысы кіру векторының шығу векторына түрленуінен тұрады. Мұнда түрлену желі салмағымен беріледі.

Нейрондық желілердің сыныптауын келесі сипаттармен сыныптауға болады:

1. Нейрондық желілердегі сигнал типтері. Кіріс не шығыста қолданылатын сигналдар мәндерінің сипатына байланысты нейрондық желілерді сандық (бинарлы) және аналогты (нақты) желілер деп бөледі. Бинарлы нейрондық желілер тек екілік сигналдармен жұмыс істейді және нейронның әр шығысы логикалық бір мәнін (қозған күйін), не логикалық нөл мәнін (тежеу күйін) қабылдайды. Ал аналогты нейрондық желілердегі шығыс және кіріс сигналдарының мәндері нақты сандар болады.

2. Белсендіру функциясының сипаты. Егер желінің барлық нейрондары үшін белсендіру функциясы біреу ғана болса, онда бұл желі *біртекті (гомогенді) желі* деп аталады. Ал егер белсендіру функциясы нейроннан нейронға мәні өзгертін кейбір

параметрлерге тәуелді болса, онда желі *біртекті емес (гетерогенді) желі* деп аталады.

3. Белсендіру функциясының типі. Нейрондық желінің әртүрлі қабаттарында қолданылатын белсендіру функциясының типіне байланысты *сызықтық және сызықтық емес желі* болып бөлінеді.

4. Нейрондық желілердің моделдік уақыты. Нейрондық желілер моделдік (үздіксіз не дискретті) уақытқа байланысты *үздіксіз не дискретті* уақытты желілер болып бөлінеді.

5. Синапстерді баптау сипаты. Синапстерді баптау сипаты бойынша нейрондық желілер *бекітілген байланысты нейрондық желілер* және *динамикалық байланысты нейрондық желілер* болып бөлінеді.

Бекітілген байланысты нейрондық желілерде синапстық байланыстардың салмақтар жиыны W есептің шартынан бірден таңдалады, мұнда $\frac{dW}{dt}=0$, ал динамикалық байланысты нейрондық желілерде синапстық байланыстардың салмақтар жиыны оқыту үдерісінде бапталады, яғни $\frac{dW}{dt} \neq 0$.

Сигналдарды жіберу уақыты. Сигналдарды жіберу уақыты (нейрондардың күйлерін өзгерту әдісі) бойынша нейрондық желілер *синхронды нейрондық желілер* және *асинхронды нейрондық желілер* болып бөлінеді. Синхронды нейрондық желілерде әр уақыт моментінде тек 1 нейрон өзінің күйін өзгертеді. Асинхронды нейрондық желілерде бір уақытта жалпы нейрондық топтың күйі өзгереді, яғни, әдетте бүкіл қабаттың күйі өзгереді. Нейрондық желілердегі уақыттың алгоритмдік жүрісі біртиптік әрекеттерді нейрондарда итерациялық орындалуымен беріледі.

Синхронизациялық әдіс. Бірқатар нейрондық желілерде белсендіру функциясы тек ω_{ij} салмақтық байланыс

коэффициенттері ғана емес, сонымен қатар, байланыс арнасы бойынша сигналдарды τ_{ij} тасымалдау уақытына да тәуелді болады. Сонда, c_{ij} байланысты белсендіру функциясы u_i элементінен u_j элементіне дейін былай болады: $c_{ij}^* = f(w_{ij}(t), u_{ij}^*(t - \tau_{ij}))$. Егер әр байланыстың тасымалдау уақыты $\tau_{ij} = 0$ не басқа нақты тұрақтыға $\tau_{ij} = \tau$ болса, онда бұл желі *синхронды нейрондық желі* деп аталады. Ал егер u_i және u_j элементтері арасындағы әр байланыстың тасымалдау уақыты өзіндік және тұрақты болса, онда бұл желі *асинхронды нейрондық желі* деп аталады. Нейрондық желілерді программалық орындауларда синхронизациялау сұрағы көкейкесті емес. Ал басқаша жүзеге асырылу тәсілінде ол өте маңызды.

6. Шешілетін есептер түрлері. Шешілетін есептер түрлеріне қарай нейрондық желілер *қалыптасатын желілер, байланыс матрицасы қалыптасатын желілер, құрамдас нейрондық желілер, оқытылатын нейрондық желілер* болып бөлінеді.

Қалыптасатын нейрондық желілер нейрожелілік базисте нақты формалданатын шешу алгоритмі бар есептер үшін жобаланады. Ал *байланыс матрицасы қалыптасатын нейрондық желілер* қиын формалданатын есептерде қолданылады. Әдетте мұндай нейрондық желілердің құрылымы бірдей болады, тек байланыс матрицасымен ерекшеленеді (Хопфилд желісі). Осындай желілер жұмыста көрнекті көрсетілген.

Құрамдас нейрондық желілер негізгі желілердің екі не үш түрлерін қамтиды және бірнеше қабатты болып келеді. Құрамдас нейрондық желілердің әр қабаты түрлі топологиямен көрсетіледі және арнайы алгоритіммен оқытылады. Құрамдас нейрондық желілердің тобы әзірлеушілерге кең ауқымды мүмкіндіктер береді.

7. Нейрон құрылымының типтері. Нейрондардың құрылым типтеріне байланысты нейрондық желілер *гомогенді* (біртекті) және

гетерогенді (әртекті) желілер болып бөлінеді. Гомогенді нейрондық желілер бірыңғай белсендіру функциясы бар, бір типті нейрондардан тұрады. Ал, гетерогенді нейрондық желіге әртүрлі белсендіру функциясы бар нейрондар кіреді.

8. Топология. Топология түріне байланысты нейрондық желілер *бір қабатты* және *көп қабатты* болып бөлінеді. Нейрондық желілердің топологиясы – нейрондардың бір-бірімен байланысуын салмақталған байланыстары бар, бағытталған граф түрінде суреттеу, мұндағы нейрондар түйіні (төбесі) болады.

Нейрондық желілердің құрылымын таңдау есептің ерекшелігі мен күрделілігіне сәйкес болады. Есептердің жеке сыныптарын шешу үшін нейрондық желілердің оңтайлы конфигурациясы бар. Егер де шешілетін есеп ешбір белгілі есептер сыныбына келтірілмесе, онда жаңа конфигурацияны синтездеу үшін күрделі мәселені шешу қажет болады. Ал нейрондық желіні синтездеу шығарылатын есепке қатты тәуелді болғандықтан, белгілі бір әдісті ұсыну қиынға соғады. Көп жағдайда оңтайлы шешу жолы интуитивік таңдаумен алынады.

Қазіргі таңда келесі нейрондық желілер бар:

- Розенблатт персептроны;
- Көп қабатты персептрон;
- Джордан желісі;
- Элман желісі;
- Хэмминг желісі;
- Ворд желісі;
- Хопфилд желісі;
- Кохонен желісі;
- Нейрондық газ;
- Когнитрон;
- Неокогнитрон;
- Хоатикалық нейрондық желі;
- Осциляторлық нейрондық желі;
- Қарама-қарсы тарау желісі;
- Радиал базистік функция желісі;

- Жалпылау регрессия желісі;
- Ықтималдылық желі;
- Сиамдық нейрондық желі;
- Адаптивтік резонанс желісі.

II.1.6-тапсырмалар.

1. Белсендіру функциясының сипаттары, топология және байланыс типі бойынша нейрондық желілердің сыныпталуын тізіп, анықтап көрсетіңіз.

2. Синапстарды баптау сипаты бойынша нейрондық желілердің сыныпталуын тізіп, анықтап көрсетіңіз.

3. Толық байланысты нейрондық желілердегі кіру сигналдарының таратылуын сипаттаңыз.

4. Толық байланысты нейрондық желілердегі шығу сигналдарының таратылуын сипаттаңыз.

II.1.6-сұрақтар.

1. Нейрондық желілер қандай критерий бойынша сандық нейрондық желілерге және аналогтық нейрондық желілерге бөлінеді?

2. Нейрондық желілер қандай критерий бойынша біртекті нейрондық желілерге және әртекті нейрондық желілерге бөлінеді?

3. Нейрондық желілер қандай критерий бойынша бір қабатты нейрондық желілерге және көп қабатты нейрондық желілерге бөлінеді?

4. Жүйелі түрдегі нейрондық желілердегі нейрондардың байланыс саны неге тәуелді болады?

II.1.6-тестілер.

1. Жасанды нейрондары желісі деген не?

А) өзара бір-бірімен байланысқан нейрондар жиынтығы;

В) өзара бір-бірімен байланыспаған нейрондар жиынтығы;

С) басқа нейрондармен байланысқан нейрондар жиынтығы;

Д) басқа нейрондармен байланыспаған нейрондар жиынтығы;

Е) барлық нейрондар жиынтығы.

2. Кіріс не шығыста қолданылатын сигналдар мәндерінің сипатына байланысты нейрондық желілерді қалай бөледі?

А) сандық (бинарлы) және аналогты (нақты);

- B) логикалық және нақты;
- C) үздікті және үздіксіз;
- D) дискретті және үздіксіз;
- E) унарлы және бинарлы.

3. Нейрондық желілердің әр түрлі қабаттарында қолданылатын белсендіру функциясына байланысты желі қалай бөледі?

- A) сызықтық және бейсызықтық желі;
- B) логикалық және нақты желі;
- C) үздікті және үздіксіз желі;
- D) дискретті және үздіксіз желі;
- E) унарлы және бинарлы желі.

4. Нейрондық желілер моделдік уақытқа байланысты қалай бөледі?

- A) дискретті және үздіксіз уақытты желі;
- B) логикалық және нақты желі;
- C) үздікті және үздіксіз желі;
- D) сызықтық және бейсызықтық желі;
- E) унарлы және бинарлы желі.

5. Нейрон құрылымының типтеріне байланысты нейрондық желілер қалай бөледі?

- A) гомогенді (біртекті) және гетерогенді (әртекті) желілер;
- B) логикалық және нақты желі;
- C) үздікті және үздіксіз желі;
- D) сызықтық және бейсызықтық желі;
- E) дискретті және үздіксіз уақытты желі.

II.2. Белсендіру функциясы

II.2.1. Белсендіру функциясы туралы жалпы мағлұмат

Бұл тақырыпта белсендіру функциясы жайында жалпы мағлұмат қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Белсендіру функциясы (табыстама функциясы) – аргумент ретінде қосқыштың шығысында алынатын салмақталған қосынды s қабылдайтын, y жасанды нейронның шығу сигналын есептейтін, яғни $y=f(s)$ болатын, f -функциясы.

Белсендіру функциясының түрі жасанды нейронның күшейткіш сипаттамасы болады және оның функционалды мүмкіндіктері мен оны оқыту әдістерін анықтайды.

Негізінен сызықты емес белсендіру функциясын қолданады, себебі сызықтық функциялар шектелген және олардың шығысы кірісіне пропорционал болады. Сызықты белсендіру функцияларын қолдану нейрондық желілердің бұрынғы моделдерінде проблема тудырды және олардың шектілігі мен мақсатқа сай еместігі Мински мен Пейперттің «Персептрондар» кітабында дәлелденді.

Бүгінгі таңда белсендіру функциясының көптеген түрі белгілі. Белсендіру функциясының мысалдары II.2.1-кестеде көрсетілген.

II.2.1-кесте. Белсендіру функциясының мысалдары.

№	Аталуы	Түрі	Анықталу аймағы мен мәні
1	Сызықтық белсендіру функциясы (PurelIn)	$f(s) = ks$	$D(f) = (-\infty, +\infty)$ $E(f) = (-\infty, +\infty)$
2	Жартылай сызықтық белсендіру функциясы (PoslIn)	$f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}$	$D(f) = (-\infty, +\infty),$ $ k > 0$ $E(f) = (-\infty, +\infty)$
3	Ығысқан қанықтырғыш сызықтық – қадамдық белсендіру функциясы (SatlIn)	$f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s-T}{\Delta}, & \text{егер } T < s < T + \Delta \\ t, & \text{егер } s \geq T + \Delta \end{cases}$	$D(f) = (-\infty, +\infty)$ $E(f) = (0, +t)$

4	Симметриялық қанықтырғыш сызықтық – қадамдық белсендіру функциясы (satlIns)	$f(s) = \begin{cases} -t, & \text{егер } s \leq -T \\ s \cdot \frac{t}{T}, & \text{егер } s < T \\ t, & \text{егер } s \geq T \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= (-t, +t) \end{aligned}$
5	Үшбұрышты белсендіру функциясы (Trlbas)	$f(s) = \begin{cases} 0, & \text{егер } s < -1 \\ 1 - s , & \text{егер } -1 \leq s \leq 1 \\ 0, & \text{егер } s > 1 \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= (0, +1) \end{aligned}$
6	Ығысқан сатылы – межелік белсендіру функциясы (HardlIm)	$f(s) = \begin{cases} t, & \text{егер } s \geq T \\ 0, & \text{егер } s < T \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= \{0\} \cup \{+t\} \end{aligned}$
7	Симметриялы сатылы – межелік белсендіру функциясы (HardlImS)	$f(s) = \begin{cases} 1, & \text{егер } s \geq 0 \\ -1, & \text{егер } s < 0 \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ E(f) &= \{-1\} \cup \{+1\} \end{aligned}$
8	Модулдық белсендіру функциясы (Modul)	$f(s) = s $	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= (0, +\infty) \end{aligned}$
9	Ығысқан сигмоидальқ-логистикалық белсендіру функциясы (Logslg)	$f(s) = \frac{1}{1 + e^{-\alpha s}}$	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= (0, +1) \end{aligned}$
10	Симметриялық сигмоидальқ белсендіру функциясы – гиперболалық тангенс (Tanslg)	$f(s) = \frac{e^{2s} - 1}{e^{2s} + 1}$	$\begin{aligned} D(f) &= (-\infty, +\infty) \\ E(f) &= (-1, +1) \end{aligned}$
11	Рационалды сигмоидальқ белсендіру функциясы (RsIgmId)	$f(s) = \frac{s}{k + s }, \quad f(s) = \frac{s}{k + s}$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ & k > 0 \\ E(f) &= (-1, +1) \end{aligned}$
12	Ығысқан ара тәріздес белсендіру функциясы (Sawto)	$f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } 0 < s < 1 \\ 0, & \text{егер } s \leq -1 \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ E(f) &= (0, +1) \end{aligned}$
13	Симметриялық ара тәріздес белсендіру функциясы (Sawtos)	$f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } -1 < s < 1 \\ -1, & \text{егер } s \leq -1 \end{cases}$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ E(f) &= (-1, +1) \end{aligned}$
14	Синусоидалық белсендіру функциясы (SIn)	$f(s) = \sin(s)$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ E(f) &= (-1, +1) \end{aligned}$
15	Дәрежелі белсендіру функциясы (Deg)	$f(s) = s^n$	$\begin{aligned} D(f) &= (-\infty, +\infty), \\ E(f) &= (0, +\infty), \end{aligned}$

			$E(f) = (-\infty, +\infty)$
16	Квадраттық түбір (Square) – белсендіру функциясы	$f(s) = \sqrt{s}$	$D(f) = (0, +\infty),$ $E(f) = (0, +\infty)$
17	Экспоненциальды белсендіру функциясы (Exp)	$f(s) = e^{-ks}$	$D(f) = (-\infty, +\infty),$ $E(f) = (0, +\infty)$
18	Кіріс мәндерін азайтатын белсендіру функциясы (Softmax)	$f(s_i) = \frac{e^{s_i}}{\sum_j e^{s_j}},$	$D(f) = (-\infty, +\infty),$ $E(f) = (0, +1)$
19	Радиалды базистік белсендіру функциясы (Radbas)	$f(s) = \varphi\left(\frac{s}{\sigma^2}\right)$	$D(f) = (-\infty, +\infty),$ $E(f) = (0, +1)$
20	Бәсекелесетін белсендіру функциясы (Compet)	$f(s) = \begin{cases} 1, & \text{егер } s = \max_i \sum_{i=1}^n x_i w_i \\ 0, & \text{егер } s \neq \max_i \sum_{i=1}^n x_i w_i \end{cases}$	$D(f) = (-\infty, +\infty),$ $E(f) = (0, +1)$

II.2.1-тапсырмалар.

1. Сызықтық белсендіру функциясының формуласын жазыңыз.
2. Жартылай сызықтық белсендіру функциясының формуласын жазыңыз.
3. Ығысқан қанықтырғыш сызықтық–қадамдық белсендіру функциясының формуласын жазыңыз.
4. Үшбұрышты белсендіру функциясының формуласын жазыңыз.
5. Модулдық белсендіру функциясының формуласын жазыңыз.
6. Ығысқан сатылы белсендіру функциясының формуласын жазыңыз.
7. Симметриялы сатылы белсендіру функциясының формуласын жазыңыз.
8. Логистикалық белсендіру функциясының формуласын жазыңыз.
9. Гиперболалық белсендіру функциясының формуласын жазыңыз.

10. Экспоненциалды белсендіру функциясының формуласын жазыңыз.

11. Синусоидалық белсендіру функциясының формуласын жазыңыз.

12. Рационалды белсендіру функциясының формуласын жазыңыз.

13. Қадамдық белсендіру функциясының формуласын жазыңыз.

14. Межелік белсендіру функциясының формуласын жазыңыз.

15. Сигнатурды белсендіру функциясының формуласын жазыңыз.

16. Квадраттық белсендіру функциясының формуласын жазыңыз.

II.2.1-сұрақтар.

1. Синусоидалық белсендіру функциясы қалай анықталады?
2. Радиалды базисті белсендіру функциясы қалай анықталады?
3. Ара тәріздес белсендіру функциясы қалай анықталады?
4. Гиперболалық тангенс функциясы қалай анықталады?
5. Рационалды-сигмоидалық функциясы қалай анықталады?
6. Логистика-базистік функциясы қалай анықталады?
7. Модулдық белсендіру функциясы қалай анықталады?
8. Сигмоидалық белсендіру функциясы қалай анықталады?
9. Бәсекелесуші белсендіру функциясы қалай анықталады?

II.2.1-тестілер.

1. Аргументі салмақталған қосынды s болатын жасанды нейронның шығу сигналын есептейтін функция қалай аталады?

- А) белсендіру функциясы;
- В) межелік функция;
- С) қарапайым функция;
- Д) логарифмдік функция;
- Е) дискретті функция.

2. Модулдық белсендіру функциясының (Modul) түрі?

- A) $f(s) = |s|$;
- B) $f(s) = ks$;
- C) $f(s) = \frac{1}{1 + e^{-\alpha s}}$;
- D) $f(s) = s^n$;
- E) $f(s) = s^n$.

3. Ығысқан сигмоидалық-логистикалық белсендіру функциясының (Logsig) түрі?

- A) $f(s) = \frac{1}{1 + e^{-\alpha s}}$;
- B) $f(s) = |s|$;
- C) $f(s) = ks$;
- D) $f(s) = s^n$;
- E) $f(s) = s^n$.

4. Дәрежелі белсендіру функциясының (Deg) түрі?

- A) $f(s) = s^n$;
- B) $f(s) = |s|$;
- C) $f(s) = \frac{1}{1 + e^{-\alpha s}}$;
- D) $f(s) = ks$;
- E) $f(s) = s^n$.

5. Кіріс мәндерін азайтатын белсендіру функциясының (Softmax) түрі?

- $$f(s_i) = \frac{e^{s_i}}{\sum_j e^{s_j}}$$
- A) ;
- B) $f(s) = |s|$;
- C) $f(s) = \frac{1}{1 + e^{-\alpha s}}$;
- D) $f(s) = s^n$;
- E) $f(s) = ks$.

II.2.2. Сызықтық белсендіру функциясы

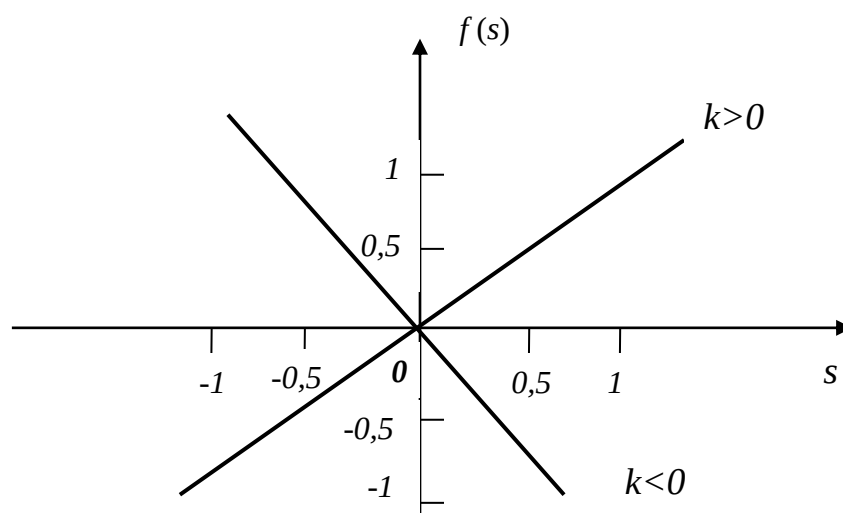
Бұл тақырыпта сызықтық белсендіру функциясы мен олардың жетілдірілуі жайында жалпы мағлұмат қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Сызықтық белсендіру функциясы s нейронның потенциал мәніне пропорционалды y шығу сигналы туындалады және $f(s) = ks$ өрнегімен көрсетіледі.

Сызықтық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$ болады, ал мәндер аймағы $E(f) = (-\infty, +\infty)$ болады.

Сызықтық белсендіру функциясының графигі II.2.2.1-суретте көрсетілген.



II.2.2.1-сурет. Сызықтық белсендіру функциясының графигі

Сызықтық белсендіру функциясы әр түрлі типтегі нейрондық желілерде қолданылады, соның ішінде сызықтық белсендіру функциясында және радиалды базистік белсендіру функциясының шығыс қабаттарында қолданылады.

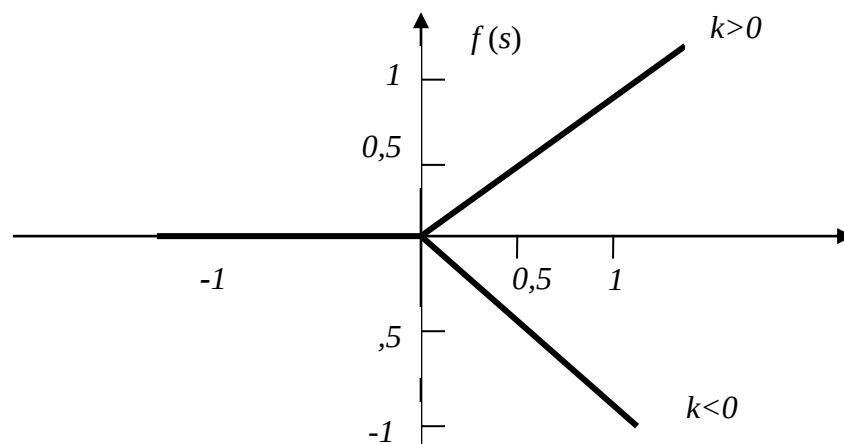
Сызықтық белсендіру функциясы келесі жетілдірулерден тұрады: *жартылай сызықтық, қадамдық және үшбұрышты*.

Жартылай сызықтық белсендіру функциясы: егер нейрон потенциалының мәні оң болса, s нейронның потенциалына пропорционалды u шығу сигналы туындалады, ал теріс болса шығу сигналы туындалмай келесі өрнекпен сипатталады:

$$f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}$$

Коэффициентінің мәні оң немесе теріс болуына байланысты, сызықтық белсендіру функциясының мәні де сәйкесінше оң немесе теріс болады. Сол себепті сызықтық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәнінің аймағы $E(f) = (-\infty, +\infty)$ болады.

Жартылай сызықтық белсендіру функциясының графигі II.2.2.2-суретте көрсетілген:



II.2.2.2-сурет. Жартылай сызықтық белсендіру функциясының графигі

Жартылай сызықтық белсендіру функциясы қарапайым желілерді моделдеуде қолданылады, мысалы, персептронды моделдеуде.

Қадамдық белсендіру (қанығуы бар сызықтық белсендіру) функциясының екі сызықтық аумағы бар, мұнда оның мәні мүмкін болатын минималды $-t$ ға және мүмкін болатын максималды t ға тең, және функцияның қатаң монотонды өсетін аумағы да бар. Сонымен қатар екі ось бойынша да функцияның ауытқуы болуы мүмкін: t – абцисса осі бойымен, Δ – ордината осі бойымен.

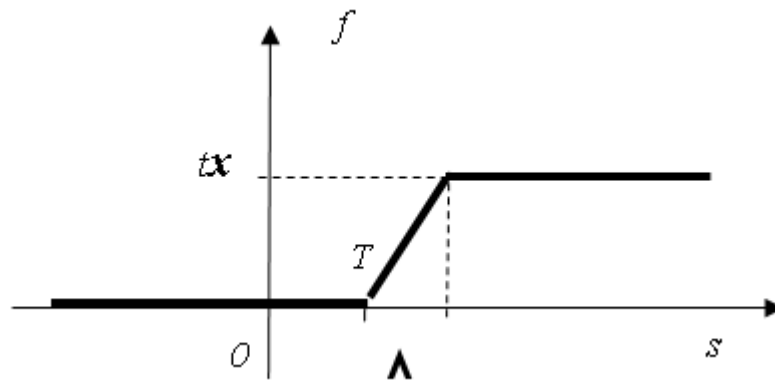
Қадамдық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәнінің аймағы $-E(f) = (-t, +t)$ болады.

Қадамдық белсендіру функциясының екі түрі бар: ығыстырылған қадамдық белсендіру функциясы, симметриялы қадамдық белсендіру функциясы.

Ығыстырылған қадамдық белсендіру функциясы келесідей өрнектеледі:

$$f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s - T}{\Delta}, & \text{егер } T < s < T + \Delta \\ t, & \text{егер } s \geq T + \Delta \end{cases}$$

Ығыстырылған қадамдық функциясының графигі II.2.2.3-суретте көрсетілген.

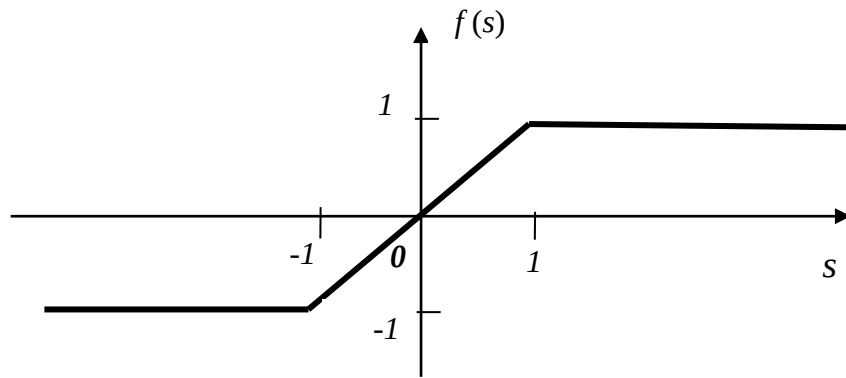


II.2.2.3-сурет. Ығыстырылған қадамдық белсендіру функциясының графигі

Симметриялы қадамдық белсендіру функциясы келесідей өрнектеледі:

$$f(s) = \begin{cases} -t, & \text{егер } s \leq -T \\ s * \frac{t}{T}, & \text{егер } |s| < T \\ t, & \text{егер } s \geq T \end{cases}$$

Мына $t=1$ және $T=1$ шарттары орындалғанда симметриялы қадамдық функциясының графигі II.2.2.4-суреттегідей болады.



II.2.2.4-сурет. Симметриялы қадамдық белсендіру функциясы

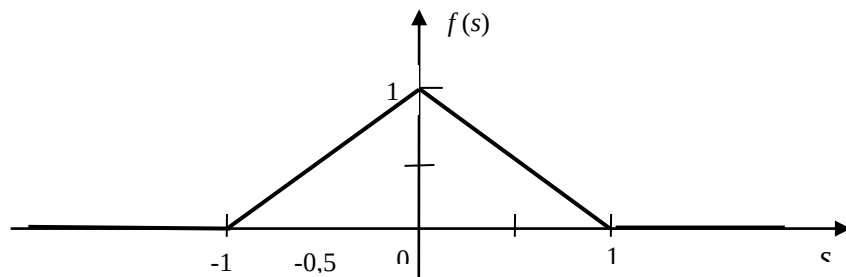
Қадамдық белсендіру функциясы қарапайым желілерді моделдеуде қолданылады, мысалы, персептрондарды моделдеуде.

Жартылай сызықтық және қадамдық белсендіру функцияларының сызықтыққа қарағанда кемшілігі, олар барлық сандық осте дифференциалданбайды, демек белсендіру функциясының дифференциалдануын талап ететін кейбір алгоритмдер бойынша оқытуда пайдалану мүмкін болмайды, мысалы, қателерді қайта тарату алгоритмі.

Үшбұрышты (шектеулі кесек-сызықты) белсендіру функциясы екі сызықтық телімге ие, мұнда оның мәні нөлге тепе-тең болады және екі телім болады: біріншісінде функция қатаң монотонды өседі, екіншісінде функция қатаң монотонды кемиді.

$$f(s) = \begin{cases} 0, & \text{егер } s \leq -1 \\ 1 - |s|, & \text{егер } -1 \leq s \leq 1 \\ 0, & \text{егер } s > 1 \end{cases}$$

Сызықтық белсендіру функциясының графигі II.2.2.5-суретте көрсетілген.



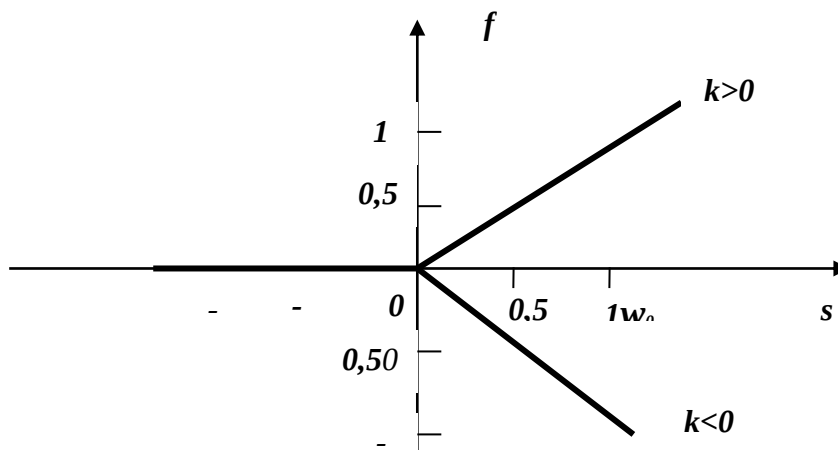
II.2.2.5-сурет. Сызықтық белсендіру функциясының графигі

II.2.2-тапсырмалар.

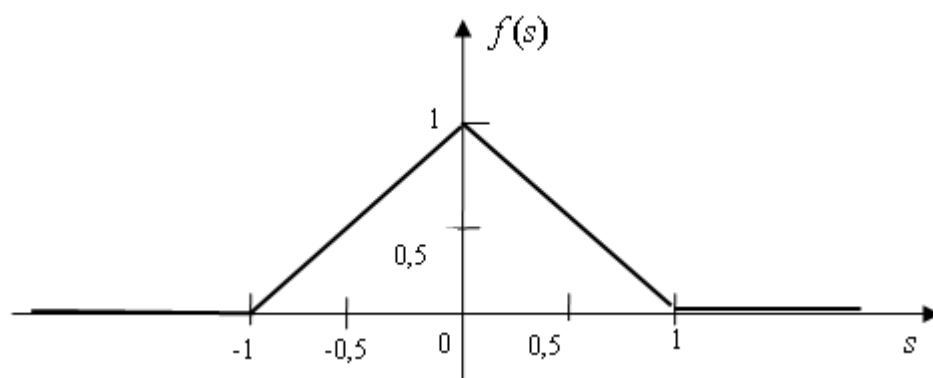
1. Сызықтық белсендіру функциясының формуласын жазыңыз.
2. Жартылай сызықтық белсендіру функциясының формуласын жазыңыз.
3. Ығыстырылған қадамды белсендіру функциясының формуласын жазыңыз.
4. Симметриялы қадамды белсендіру функциясының формуласын жазыңыз.
5. Үшбұрышты белсендіру функциясының формуласын жазыңыз

II.2.2-сұрақтар.

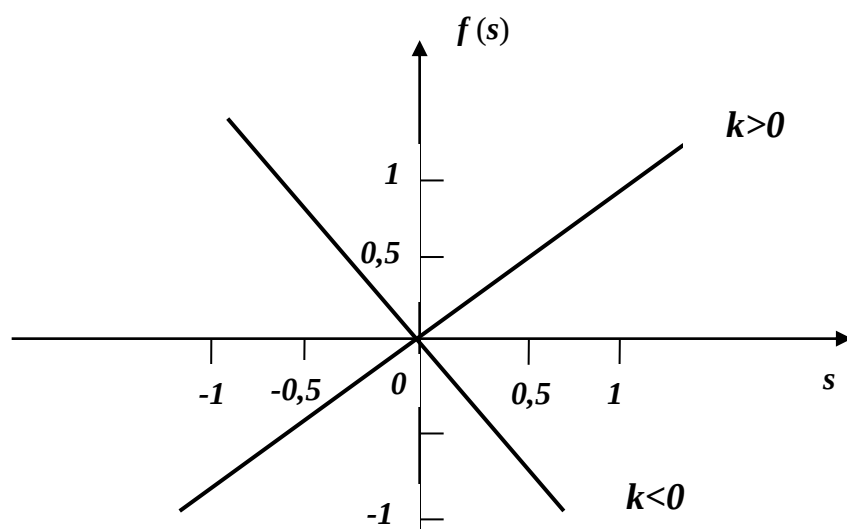
1. Сызықтық белсендіру функциясы мен жартылай сызықтық белсендіру функциясының айырмашылығын түсіндіріңіз?
2. Сызықтық белсендіру функциясы мен үшбұрышты белсендіру функциясының айырмашылығын түсіндіріңіз ?
3. Төмендегі сурет қандай функцияның графигі болады?



4. Төмендегі сурет қандай функцияның графигі болады?



5. Төмендегі сурет қандай функцияның графигі болады?



II.2.2-тестілер.

1. Мына s нейрон потенциалының (салмақталған сома) мәніне пропорционалды y шығу сигналы туындалатын сызықтық белсендіру функциясы қандай өрнекпен беріледі?

A) $f(s) = ks$;

B) $f(s) = s^n$;

C) $f(s) = |s|$;

D) $f(s) = e^{-ks}$;

E) $f(s) = \sqrt{s}$.

2. Жартылай сызықтық белсендіру функциясы (PosIn) қандай өрнекпен сипатталады?

$$A) f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases};$$

$$B) f(s) = s^n;$$

$$C) f(s) = |s|;$$

$$D) f(s) = e^{-ks};$$

$$E) f(s) = \sqrt{s}.$$

4. Ығыстырылған қадамдық белсендіру функциясы қалай өрнектеледі?

$$A) f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s-T}{\Delta}, & \text{егер } T < s < T + \Delta \\ \frac{\Delta}{t}, & \text{егер } s \geq T + \Delta \end{cases}$$

$$B) f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}$$

$$C) f(s) = |s|;$$

$$D) f(s) = e^{-2x}$$

$$E) f(s) = \sqrt{s}.$$

4. Симметриялы қадамдық белсендіру функциясы қалай өрнектеледі?

$$A) f(s) = \begin{cases} -t, & \text{егер } s \leq -T \\ s * \frac{t}{T}, & \text{егер } |s| < T \\ t, & \text{егер } s \geq T \end{cases};$$

$$B) f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases};$$

$$C) f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s-t}{\Delta}, & \text{егер } T < s < T + \Delta \\ t, & \text{егер } s \geq T + \Delta \end{cases}$$

$$D) f(s) = e^{-ks};$$

$$E) f(s) = \sqrt{s}.$$

5. Ығыстырылған қадамдық белсендіру функциясы қалай өрнектеледі?

$$A) f(s) = \begin{cases} 0, & \text{егер } s \leq -1 \\ 1 - |s|, & \text{егер } -1 \leq s \leq 1; \\ 0, & \text{егер } s > 1 \end{cases}$$

$$B) f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases};$$

$$C) f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s-t}{\Delta}, & \text{егер } T < s < T + \Delta \\ t, & \text{егер } s \geq T + \Delta \end{cases}$$

$$D) f(s) = \begin{cases} -t, & \text{егер } s \leq -T \\ s * \frac{t}{T}, & \text{егер } |s| < T; \\ t, & \text{егер } s \geq T \end{cases}$$

$$E) f(s) = \sqrt{s}.$$

II.2.3. Межелік белсендіру функциясы

Бұл тақырыпта межелік белсендіру функциясы жайында жалпы мағлұмат қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Межелік белсендіру функциясы айырымнан тұрады және нейрондардың белсенділігіне байланысты a мен b шеткі мәні бар шығу сигналын тудырады. Сонымен қатар, егер нейрон потенциалының мәні белгілі бір T деңгейіне жетпесе, шығу сигналы 0-ге тең болады, кері жағдайда шығыс сигналы қадам-қадам t -ға өзгереді.

Ығысқан межелік функция келесі өрнекпен ұсынылады:

$$f(s) = \begin{cases} t, & \text{егер } s \geq T \\ 0, & \text{керісінше} \end{cases},$$

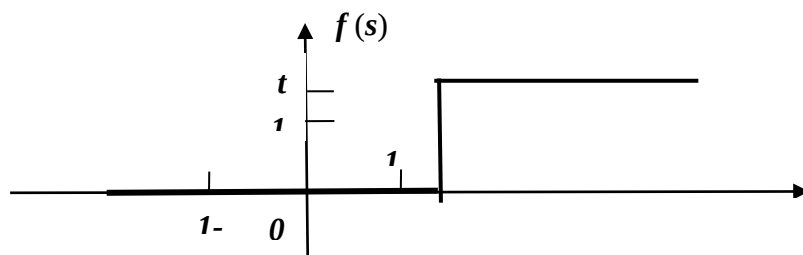
мұнда T – белсендіру функциясының көлденең оське қатысты ығысуы және $T = -w_0 \cdot x_0$, мұнда s қосылғышты есептемегендегі кірістегі нейрондар сигналдарының салмақталған сомасы болады.

Межелік белсендіру функциясын кейде *сатылы белсендіру функциясы* деп атайды.

Межелік белсендіру функциясы *ығысқан межелік белсендіру функциясына* және *симметриялық межелік белсендіру функциясына* бөлінеді.

Ығысқан межелік белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = \{-t\} \cup \{+t\}$ болады.

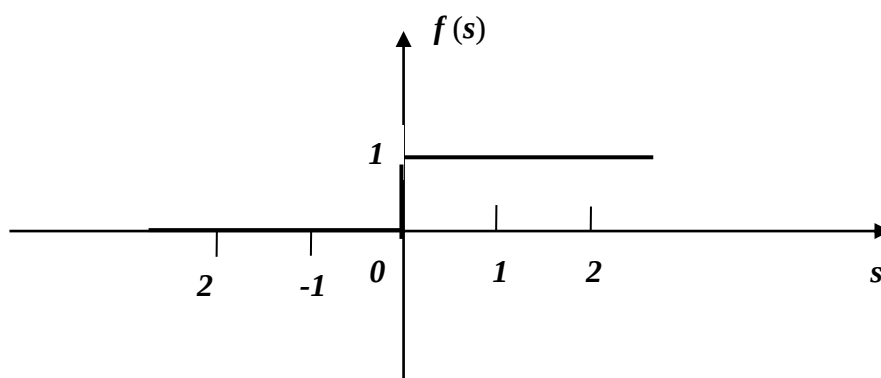
Ығысқан межелік белсендіру функциясының графигі II.2.3.1-суретте көрсетілген:



II.2.3.1-сурет. Ығысқан межелік функция графигі

Егер ығысқан межелік белсендіру функциясында $T=0$ және $t=0$ болса, онда ол *бір секіру функциясына (Хэвисайд функциясына)* айналады, ол келесі өрнекпен беріледі және II.2.3.2-суретте көрсетілген:

$$f(s) = \begin{cases} 1, & \text{егер } s \geq 0 \\ 0, & \text{егер } s < 0 \end{cases}$$



II.2.3.2-сурет. Бір секіруді белмендіру функциясының графигі

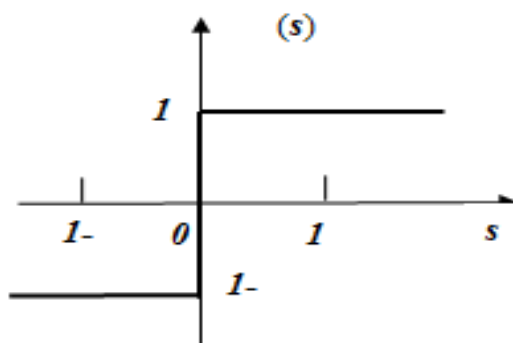
Симметриялы межелік (таңбалық) белсендіру функциясы келесі өрнекпен сипатталады:

$$f(s) = \begin{cases} 1, & \text{егер } s \geq 0 \\ -1, & \text{егер } s < 0 \end{cases}$$

Симметриялы межелік белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = \{-1\} \cup \{1\}$ болады.

Симметриялы межелік функцияны кейде *биполяры белсендіру функциясы* немесе *қатаң таңбалық функция* деп аталады.

Симметриялы межелік белсендіру функциясының графигі
II.2.3.3-суретте көрсетілген.



II.2.3.3-сурет. Симметриялы межелік функция графигі

Егер нейронның шығу сигналы a және b межелік мәндернің ішінде $m-1$ шегімен анықталатын m мәндерінің бірін қабылдайтын болса көп межелік белсендіру функциясын анықтауға болады,

Бұл функция барлық абцисса осінде дифференциалданбайтындықтан оны қателерді кері тарату алгоритмі бойынша және белсендіру функциясының дифференциалдануын талап ететін басқа алгоритмдер бойынша оқып үйрететін желілерде қолдануға болмайды.

Межелік функцияларды шешетін есепке баптауға және оқытуға икемсіз болатын нейрондық желілерді моделдеу кезінде қолдана алады, себебі, егер қосынды мәні межеге жетпейтін болса, онда шығу сигналы қалыптаспайды да, желі жұмыс істемейді. Нейронның шығу сигналының белсенділігі жоғалады және нейронның келесі қабатында салмақталатын кіріс деңгейі төмен қалыптасады.

II.2.3-тапсырмалар.

1. Межелік белсендіру функциясының түрлерін атап көрсетіңіз.
2. T меженің мәнін анықтайтын өрнекті жазыңыз.
3. Ығысқан межелік белсендіру функциясының формуласын жазыңыз.
4. Межелік функцияның графигін салыңыз.

5. Симметриялы межелік белсендіру функциясының формуласын жазыңыз.

6. Ығысқан межелік функцияның графигін салыңыз.

7. Симметриялы межелік функцияның графигін салыңыз.

II.2.3-сұрақтар.

1. Межелік функцияны басқаша қалай атайды?

2. Көлденең оське қатысты белсендіру функциясының ығысуы қалай аталады?

3. Межелік функцияның мәндер аймағы қандай?

4. Ығысқан межелік функцияның симметриялы межелік функциядан айырмашылықтары неде?

5. Қателерді қайта тарату алгоритмі бойынша нейрондық желілерді оқыту үдерісінде межелік белсендіру функциясы қолданылады ма?

6. Қателерді кері тарату әдісін сызықтық функцияны және оның жетілдіруін оқытуда қолдануға болады ма?

II.2.3-тестілер.

1. Ығысқан межелік функция қандай өрнекпен ұсынылады?

$$A) f(s) = \begin{cases} t, & \text{егер } s \geq T \\ 0, & \text{айтпесе} \end{cases} ;$$

$$B) f(s) = \begin{cases} ks, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases} ;$$

$$C) f(s) = \begin{cases} 0, & \text{егер } s \leq T \\ \frac{s-T}{\Delta}, & \text{егер } T < s < T + \Delta \\ t, & \text{егер } s \geq T + \Delta \end{cases}$$

$$D) f(s) = \begin{cases} -t, & \text{егер } s \leq -T \\ s * \frac{t}{T}, & \text{егер } |s| < T \\ t, & \text{егер } s \geq T \end{cases} ;$$

$$E) f(s) = \sqrt{s}.$$

2. Межелік белсендіру функциясы қалай бөлінеді?

А) ығысқан межелік және симметриялы межелік белсендіру функциясы;

В) сатылы және бір секіру белсендіру функциясы;

С) биполярлы белсендіру функциясы және қатаң таңбалық функция;

Д) биполярсыз белсендіру функциясы және таңбалық функция;

Е) сатылы және екі рет секіру белсендіру функциясы.

3. Симметриялы межелік функцияны кейде қалай атайды?

А) биполярлы белсендіру функциясы немесе қатаң таңбалық функция;

В) сатылы және бір секіру белсендіру функциясы;

С) ығысқан межелік және симметриялық межелік белсендіру функциясы;

Д) биполярсыз белсендіру функциясы және таңбалық функция;

Е) сатылы және екі рет секіру белсендіру функциясы.

4. Егер нейронның шығу сигналы a және b межелік мәндердің ішінде $m-1$ шегімен анықталатын m мәндерінің бірін қабылдайтын болса қандай белсендіру функциясын анықтауға болады?

А) көп межелік;

В) сатылы межелік;

С) ығысқан межелік;

Д) таңбалық межелік;

Е) қатаң межелік.

5. Симметриялы межелік белсендіру функциясының мәндер аймағы қандай?

А) $E(f) = \{-1\} \cup \{1\}$;

В) $E(f) = \{-2\} \cup \{2\}$;

С) $E(f) = \{-1\} \cup \{0\}$;

Д) $E(f) = \{0\} \cup \{1\}$;

Е) $E(f) = \{-10\} \cup \{10\}$.

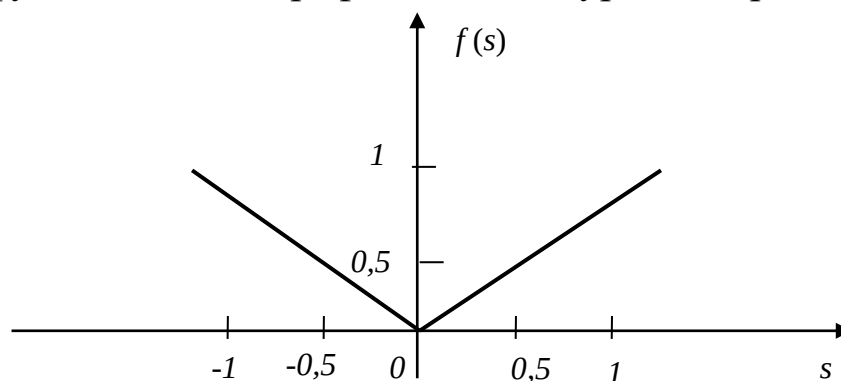
II.2.4. Модулдық белсендіру функциясы

Бұл тақырыпта Модулдық белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Модулдық белсендіру функциясы нейронның белсенділігіне байланысты үздіксіз сызықтық шығу сигналын тудырады және келесі өрнекпен көрсетіледі:

$$f(s) = |s|$$

Модулдық белсендіру функциясында $D(f) = (-\infty, +\infty)$ анықтау аймағы, ал $E(f) = (0, +\infty)$ мәндер аймағы болады. Модулдық белсендіру функциясының графигі II.2.4.1-суретте көрсетілген.



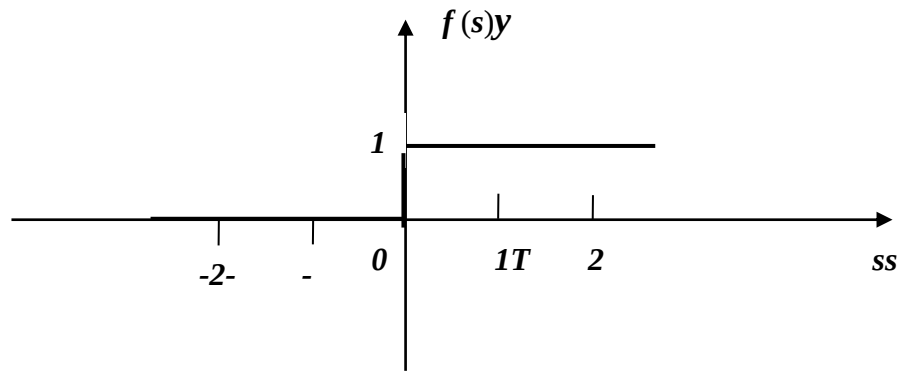
II.2.4.1-сурет. Модулдық белсендіру функциясының графигі

Модулдық белсендіру функциясымен нейронды оқыту үшін кателерді кері тарату әдісін қолдануға болады.

Осы функциядан туындыны алу үшін модулдық функция 0-ге тең болған жағдайда туындының болмайтындығын ескеру қажет, қалған нүктелерде модулды ашу керек және дифференциалдау қажет, яғни

$$f'(s) = \begin{cases} -s, & \text{егер } s < 0 \\ s, & \text{егер } s > 0 \end{cases}$$

Модулдық функция туындысының өрнегі межелік (таңбалық) белсендіру функциясының өрнегіне сәйкес келетіндігін байқауға болады. Сондықтан модулдық туынды Модулдық белсендіру функциясының графигі II.2.4.2-суреттегідей болады.



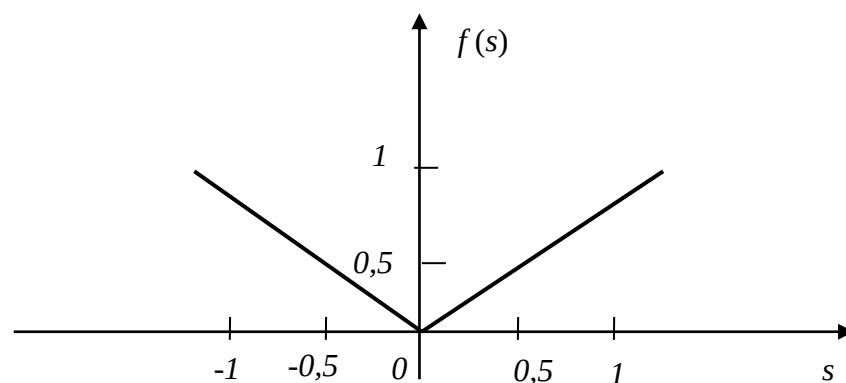
II.2.4.2-сурет. Туынды Модулдық белсендіру функциясының графигі

II.2.4-тапсырмалар.

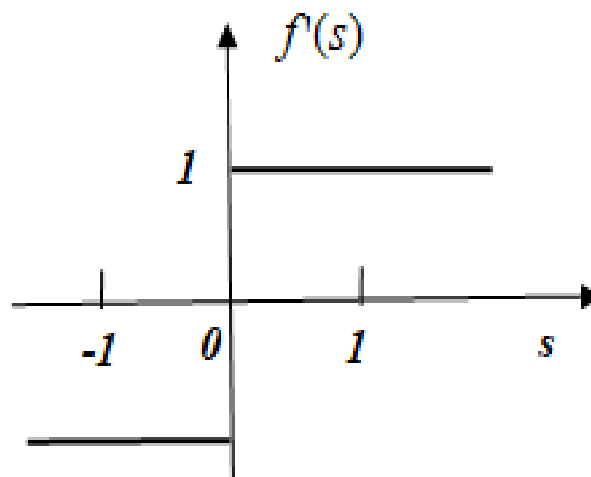
1. Модулдық белсендіру функциясы үшін өрнекті жазыңыз.
2. Модулдық белсендіру функциясының анықтау аймағын көрсетіңіз.
3. Модулдық белсендіру функциясының мәндер аймағын көрсетіңіз.

II.2.4-сұрақтар.

1. Төмендегі суретте қандай функцияның графигі көрсетілген?



2. Төмендегі суретте қандай функцияның графигі көрсетілген?



3. Модулдық белсендіру функциясының туындысын алу үшін нені ескерген жөн?
4. Модулдық белсендіру функциясының туындысы қандай формуламен есептеледі?

II.2.4-тестілер.

1. Модулдық белсендіру функциясының анықталу аймағын көрсетіңіз?

- A) $D(f) = (-\infty, +\infty)$;
- B) $D(f) = (-\infty, 0)$;
- C) $D(f) = (0, +\infty)$;
- D) $D(f) = (-1, +1)$;
- E) $D(f) = (-1, +\infty)$.

2. Модулдық белсендіру функциясының мәндер аймағын көрсетіңіз?

- A) $E(f) = (0, +\infty)$;
- B) $E(f) = (-\infty, 0)$;
- C) $E(f) = \{-1\} \cup \{0\}$;
- D) $E(f) = \{0\} \cup \{1\}$;
- E) $E(f) = (-1, +\infty)$.

3. Модулдық белсендіру функциясы нені өңдейді?

А) нейронның белсенділігіне байланысты үздіксіз сызықтық шығу сигналын;

В) нейронның белсенділігіне байланысты дискретті кіру сигналын;

С) нейронның белсенділігіне байланысты сызықтық кіру сигналын;

Д) нейронның белсенділігіне байланысты сызықтық сигналды;

Е) нейронның тек кіру сигналын.

4. Модулдық белсендіру функциялы нейронды оқыту үшін қандай қандай әдіс қолданылады?

А) қателерді кері тарату әдісі;

В) қателерді тура тарату әдісі;

С) қателерді тікелей тарату әдісі;

Д) қателерді таратпайтын әдіс;

Е) қателерді түзету әдісі;

5. Модулдық белсендіру функциясы қандай өрнекпен беріледі?

А) $f(s) = |s|$;

В) $f(s) = ks$;

С) $f(s) = \frac{1}{1 + e^{-\alpha s}}$;

Д) $f(s) = s^n$;

Е) $f(s) = s * |s|$.

II.2.5 Сигмоидалық белсендіру функциялары.

Бұл тақырыпта сигмоидалық белсендіру функциялары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Қазіргі кезде сигмоидалық белсендіру функциялары жиі қолданылатын белсендіру функцияларының бірі болады. Сигмоидалық типті функцияларды енгізу, нейрондардың межелік белсендіру функцияларымен берілген нейрондық желілерінің шектелуімен негізделді. Мұндай белсендіру функциясында желінің кез келген шығысы не нөлге, не бірге тең болады, бұл нейрондық желілерді сыныптауға жатпайтын есептерге қолдануды шектейді. Сигмоидалық функцияларды қолдану нейронның бинарлық (цифрлық) шығысынан нақты (аналогтық) шығысына ауысуына алып келді. Бұл типтес белсендіру функциялары, нейрондық желілердің ішкі қабатында болатын нейрондарға тән.

Сигмоидалық белсендіру функциялары мәні (a, b) интервалында жататын үздіксіз шығу сигналдарын өндіреді және сигмоида (S-түріндегі қисықты) қалыптастырады.

Сигмоидалық белсендіру функциялары *ығысқан сигмоидалық белсендіру функциялары, симметриялық сигмоидалық белсендіру функциялары және рационалды сигмоидалық белсендіру функциялары* болып бөлінеді.

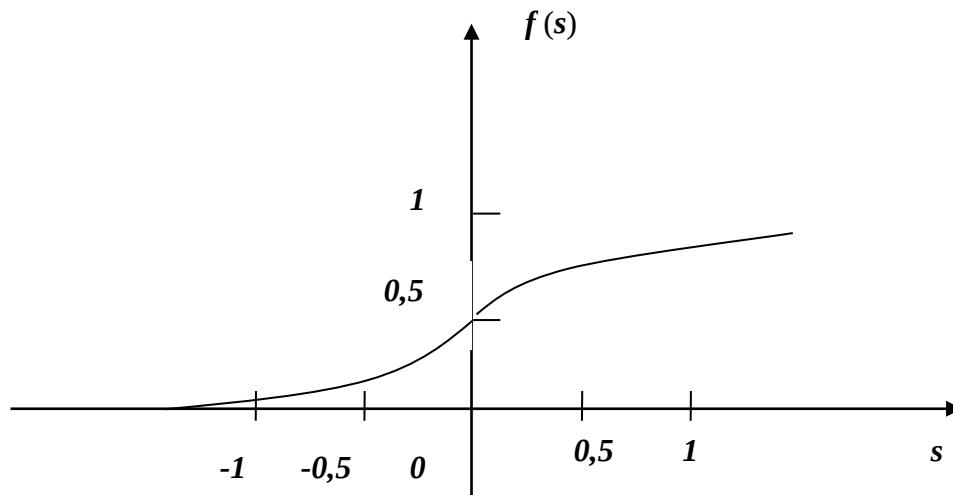
Ығысқан сигмоидалық белсендіру функциясы келесі өрнекпен беріледі:

$$f(s) = \frac{1}{1+e^{-\alpha s}}$$

мұнда α – функцияның құлдылауын анықтайтын параметр, ол шексіздікке ұмтылғанда функция межелік функцияға айналады, ал $\alpha=0$ болғанда сигмоидалық функция 0,5 мәнімен тұрақты функцияға айналады.

Ығысқан сигмоидалық функцияны кейде логистикалық белсендіру функциясы деп атайды.

Логистикалық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, 1)$ болады. Логистикалық белсендіру функциясының графигі II.2.5.1-суретте көрсетілген.



II.2.5.1-сурет. Логистикалық белсендіру функциясының графигі.

Логистикалық функция абциссаның барлық осінде дифференциалданады және өте қарапайым туындыға ие:

$$f'(s) = \alpha f(s)(1 - f(s))$$

Логистикалық функцияның туындысы оның мәнімен өрнектелуі мүмкін, бұл кері тарату алгоритмі бойынша нейрондық желіні оқытуда осы туындының қолданылуын жеңілдетеді. Осындай табыстама сипаттамасы бар нейрондар ерекшелігі мынада: олар күшті сигналдарды әлсіз сигналдарға қарағанда біршама аз күшейтеді, себебі күшті сигналдардың аймағы сипаттаманың еңіс аймақтарына сәйкес келеді. Бұл үлкен сигналдардың әсерінен болатын қанығуды болдырмау мүмкіндігін береді.

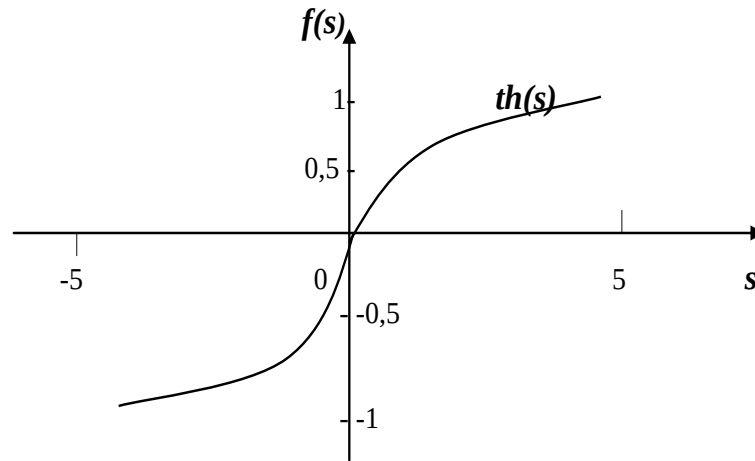
Симметриялық сигмоидалық белсендіру функциясы гиперболалық тангенс өрнегімен ұсынылады:

$$f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} = \frac{e^{2s} - 1}{e^{2s} + 1}.$$

Симметриялық сигмоидалық белсендіру функциясын кейде *гиперболалық функция (гиперболалық тангенс)* деп те атайды.

Гиперболалық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (-1, 1)$ болады.

Гиперболалық белсендіру функциясының графигі II.2.5.2-суретте көрсетілген.



II.2.5.2-сурет. Гиперболалық белсендіру функциясының графигі.

Гиперболалық белсендіру функциясы координат басына қатысты симметриялы, және $s = 0$ нүктесінде шығу сигналының мәні $y = f(s)$ нөлге тең.

Гиперболалық тангенстің туындысы келесі өрнекпен анықталады:

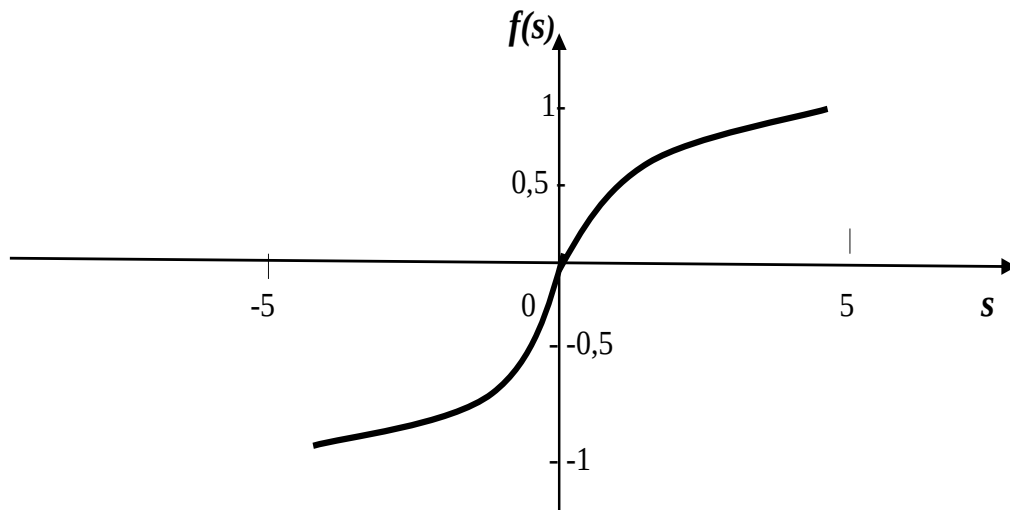
$$f'(s) = th(s) = \frac{1}{ch^2(s)}.$$

Гиперболалық функцияның логистикалық функциядан айырмашылығы, ол әр түрлі таңбалы мәндерді қабылдайды. Оның бұл қасиеті желінің бір тұтас қатары үшін қолданылады, ол көп қабатты персептрондарды баптау үшін өте қолайлы.

Рационалды сигмоидалық функция (қысқартылған сигмоида) сигмоидалық белсендіру функциясы сыныбына жатады және келесі өрнекпен беріледі:

$$f(s) = \frac{s}{k + |s|}, \quad k > 0,$$

Рационалды сигмоидалық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (-1, 1)$ болады, $k > 0$ – сигмоидалық белсендіру функциясының көлбею параметрі, бұл параметрді өзгерте отырып, құлдылаулығы әр түрлі функцияларды құруға болады. Рационалды сигмоидалық белсендіру функциясының графигі II.2.5.3-суретте бейнеленген.



II.2.5.3-сурет. Рационалды сигмоидалық белсендіру функциясының графигі.

Сигмоидалық белсендіру функциясы зерттелетін объектілердің іздеу кеңістігінде күрделі түрдегі, соның ішінде дөңес емес және байланыспаған аймақтарын ерекшелеу мүмкіндігін береді.

II.2.5-тапсырмалар.

1. Сигмоидалық белсендіру функцияларының түрлерін атап шығыңыз.
2. Ығысқан сигмоидалық функцияның формуласын жазыңыз.
3. Симметриялық сигмоидалық функцияның формуласын жазыңыз.
4. Рационалды сигмоидалық функцияның формуласын жазыңыз.
5. Логистикалық функцияның туындысының формуласын жазыңыз.

II.2.5-сұрақтар.

1. Қандай сигнал сигмоидалық белсендіру функциясын тудырады?
2. Гиперболалық белсендіру функциясы неге қатысты симметриялы?
3. Логистикалық функцияға қарағанда гиперболалық функция қандай мәндерді қабылдайды?
4. Қандай функциялар нейронның бинарлық (цифрлық) шығыстан нақты (аналогтық) шығысқа ауысуына мүмкіндік береді?
5. Белсендіру функцияларының қандай типі нейрондық желілердің ішкі қабаттарында болатын нейрондарға тән?
6. Қандай есептерді шешу үшін сигмоидалық белсендіру функциялары қолданылады?

II.2.5-тестілер.

1. Сигмоидалық белсендіру функциялары нені өндіреді?
 - А) мәні (a, b) интервалында жататын үздіксіз шығу сигналдарын;
 - В) нейронның белсенділігіне байланысты дискретті кіру сигналын;
 - С) нейронның белсенділігіне байланысты сызықтық кіру сигналын;
 - Д) нейронның белсенділігіне байланысты сызықтық сигналды;
 - Е) нейронның белсенділігіне байланысты үздіксіз сызықтық шығу сигналын.
2. Сигмоидалық белсендіру функциялары қандай түрлерге бөлінеді?
 - А) ығысқан, симметриялық, рационалды;
 - В) дискретті, үздіксіз;
 - С) антисимметриялық, симметриялық;
 - Д) иррационалды, рационалды;
 - Е) сызықтық, сызықтық емес.

3. Ығысқан сигмоидалық функцияны басқаша қалай атайды?

- A) логистикалық белсендіру функция;
- B) дискретті белсендіру функция;
- C) симметриялық белсендіру функция;
- D) рационалды белсендіру функция;
- E) сызықтық емес белсендіру функция.

4. Логистикалық белсендіру функциясының мәндер аймағы қандай?

- A) $E(f) = (0, 1)$;
- B) $E(f) = (0, 0)$;
- C) $E(f) = (-1, 1)$;
- D) $E(f) = (-\infty, +\infty)$;
- E) $E(f) = (0, +\infty)$.

5. Рационалды сигмоидалық функция қандай өрнекпен беріледі?

- A) $f(s) = \frac{s}{k + |s|}, k > 0$;
- B) $f'(s) = th(s) = \frac{1}{ch^2(s)}$;
- C) $f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} = \frac{e^{2s} - 1}{e^{2s} + 1}$;
- D) $f(s) = \frac{s}{|s|}$;
- E) $f(s) = \frac{s}{k}$.

II.2.6. Ара тәріздес белсендіру функциясы

Бұл тақырыпта ара тәріздес белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Ара тәріздес белсендіру функциясы сигмоидалды функцияның бөлікті-сызықтық нұсқасы болады және екі симметриялық шекті a мен b мәндерінің арасындағы мәндер есептеледі.

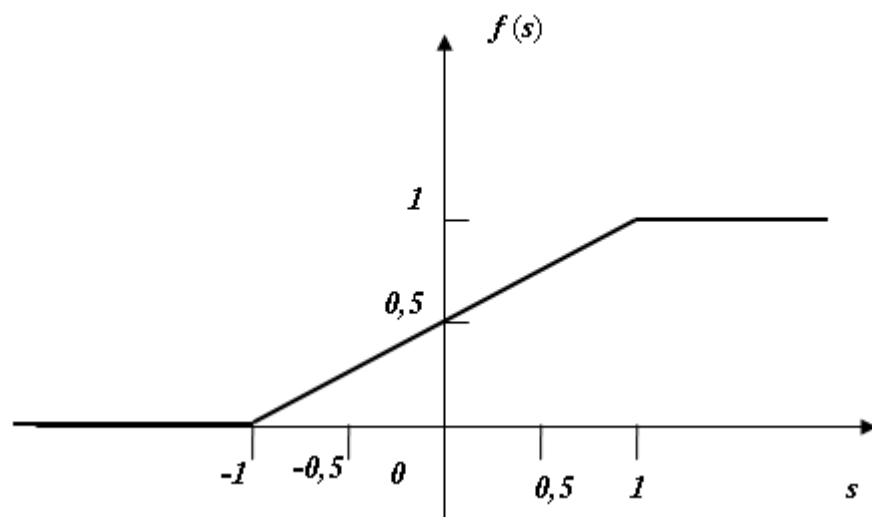
Ара тәріздес белсендіру функциясы ығыстырылған ара тәріздес белсендіру функциясына және симметриялық ара тәріздес белсендіру функциясына бөлінеді.

Ығыстырылған ара тәріздес белсендіру функциясы келесі өрнекпен беріледі:

$$f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } 0 < s < 1 \\ 0, & \text{егер } s \leq -1 \end{cases}$$

Ығыстырылған ара тәріздес белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, 1)$ бар.

Ығыстырылған ара тәріздес белсендіру функциясының графигі II.2.6.1-суретте көрсетілген.



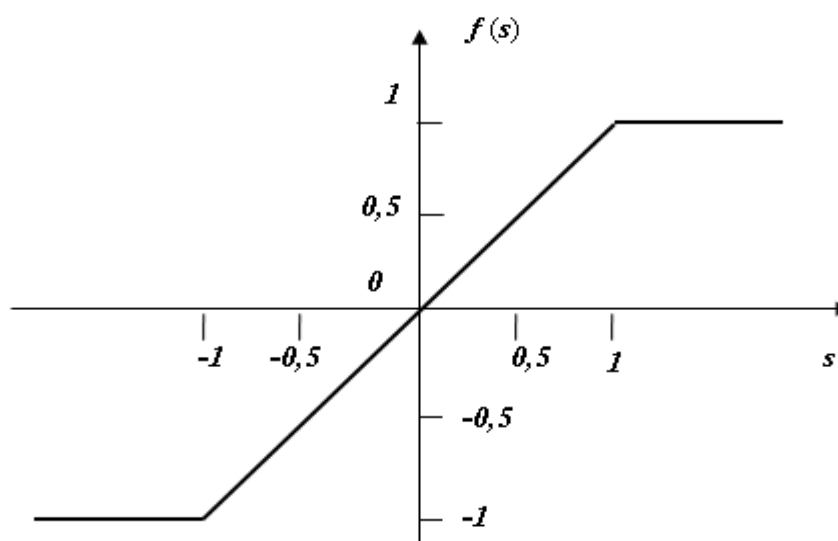
II.2.6.1-сурет. Ығыстырылған ара тәріздес белсендіру функциясының графигі

Симметриялы ара тәріздес белсендіру функциясы келесі өрнекпен беріледі:

$$f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } -1 < s < 1 \\ -1, & \text{егер } s \leq -1 \end{cases}$$

Симметриялы ара тәріздес белсендіру функциясында $D(f) = (-\infty, +\infty)$ анықтау аймағы және $E(f) = (-1, 1)$ мәндер аймағы.

Симметриялы ара тәріздес белсендіру функциясының графигі II.2.6.2-суретте көрсетілген.



II.2.6.2-сурет. Симметриялы ара тәріздес белсендіру функциясының графигі

Симметриялы ара тәріздес белсендіру функциясы сигмоидальды функцияның бөлікті-сызықтық нұсқасы болады.

II.2.6-тапсырмалар.

1. Ара тәріздес белсендіру функциясының түрлерін атаңыз.
2. Ығыстырылған ара тәріздес функцияның формуласын жазыңыздар.
3. Симметриялы ара тәріздес функцияның формуласын жазыңыздар.

II.2.6-сұрақтар.

1. Сигмоидалық функция мен ара тәріздес белсендіру функциясының арасында қандай байланыс бар?

2. Ығыстырылған сигмоидалық белсендіру функциясында мәндер аймағы болып не табылады?

3. Ығыстырылған және симметриялы ара тәріздес белсендіру функцияларының айырмашылығы неде?

II.2.6-тестілер.

1. Симметриялы ара тәріздес белсендіру функциясы қандай функцияның бөлікті-сызықтық нұсқасы болады?

- A) сигмоидалды;
- B) ығыстырылған;
- C) симметриялы;
- D) белсендіру;
- E) логистикалық.

2. Симметриялы ара тәріздес белсендіру функциясының формуласы қайсысы?

$$A) f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } -1 < s < 1 ; \\ -1, & \text{егер } s \leq -1 \end{cases}$$

$$B) f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} = \frac{e^{2s} - 1}{e^{2s} + 1};$$

$$C) f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } 0 < s < 1 ; \\ 0, & \text{егер } s \leq -1 \end{cases}$$

$$D) f(s) = \frac{s}{|s|};$$

$$E) f(s) = \frac{s}{k}.$$

3. Ығыстырылған ара тәріздес белсендіру функциясының формуласы қайсысы?

$$A) f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } 0 < s < 1 ; \\ 0, & \text{егер } s \leq -1 \end{cases}$$

$$B) f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} = \frac{e^{2s} - 1}{e^{2s} + 1};$$

$$C) f(s) = \begin{cases} 1, & \text{егер } s \geq 1 \\ s, & \text{егер } -1 < s < 1 ; \\ -1, & \text{егер } s \leq -1 \end{cases}$$

$$D) f(s) = \frac{s}{|s|};$$

$$E) f(s) = \frac{s}{k}.$$

4. Ара тәріздес белсендіру функциясы қандай түрлерге бөлінеді?

А) ығыстырылған ара тәріздес белсендіру функциясына және симметриялық ара тәріздес белсендіру функциясына;

В) сигмоидальды ара тәріздес белсендіру функциясына және симметриялық ара тәріздес белсендіру функциясына;

С) ығыстырылған ара тәріздес белсендіру функциясына және сигмоидальды ара тәріздес белсендіру функциясына;

Д) логистикалық ара тәріздес белсендіру функциясына және гиперболалық ара тәріздес белсендіру функциясына;

Е) сигмоидальды ара тәріздес белсендіру функциясына және логистикалық ара тәріздес белсендіру функциясына.

5. Ығыстырылған ара тәріздес белсендіру функциясының анықтау аймағы қайсысы?

A) $D(f) = (-\infty, +\infty);$

B) $D(f) = (0, 1);$

C) $D(f) = (0, 0);$

D) $D(f) = (-1, 1);$

E) $D(f) = (-\infty, 1).$

II.2.7. Синусоидалық белсендіру функциясы

Бұл тақырыпта синусоидалық белсендіру функциялары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

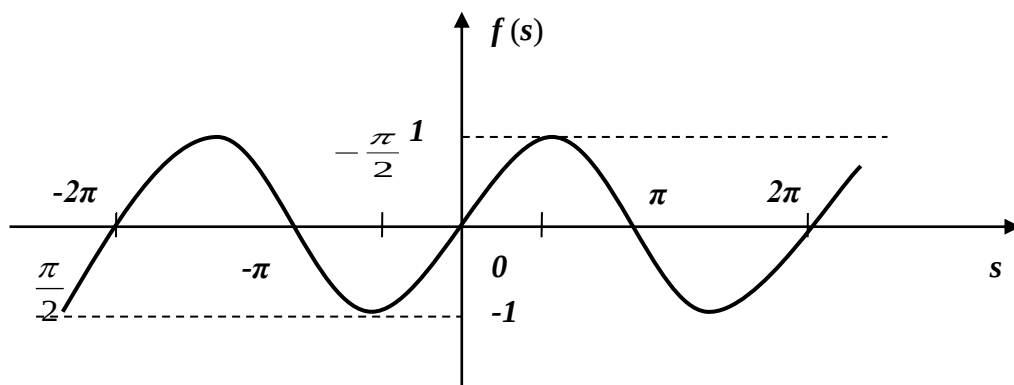
Синусоидалық белсендіру функциялары 2π периодымен периодты болады, мұнда $\pi = 3,14\dots$

Синусоидалық белсендіру функциясы келесі өрнекпен беріледі:

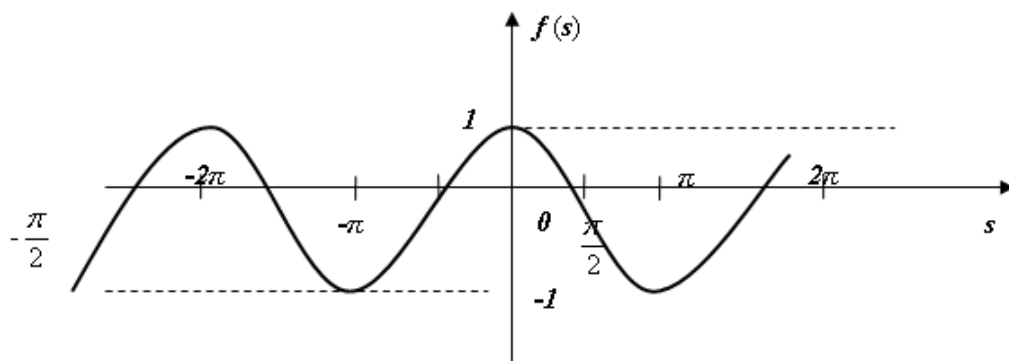
$$f(s) = \sin(s) \text{ немесе } f(s) = \cos(s)$$

Синусоидалық белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндерінің аймағы $E(f) = (-1, +1)$ болады.

Синусоидалық белсендіру функциясының графигі II.2.7.1 және II.2.7.2-суреттерде берілген.



II.2.7.1.-сурет. $\sin(s)$ белсендіру функциясының графигі.



II.2.7.2.-сурет. $\cos(s)$ белсендіру функциясының графигі.

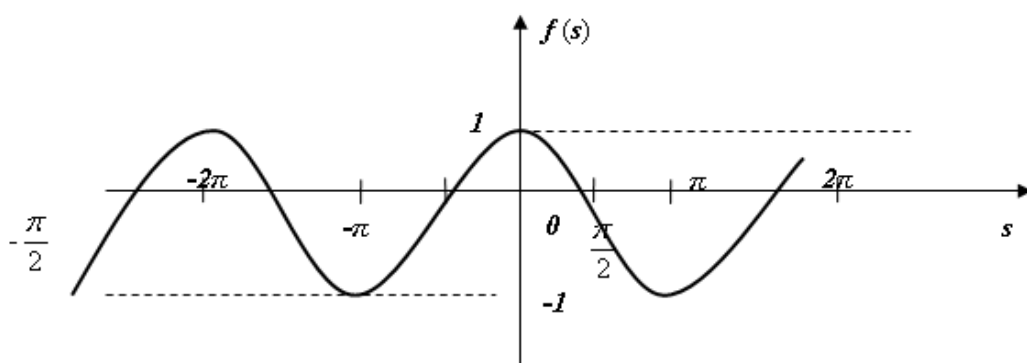
Синусоидалық белсендіру функциясы радиалды таралған деректерді тану үшін пайдалы, үнсіз келісім бойынша қолданылмайды.

II.2.7-тапсырмалар.

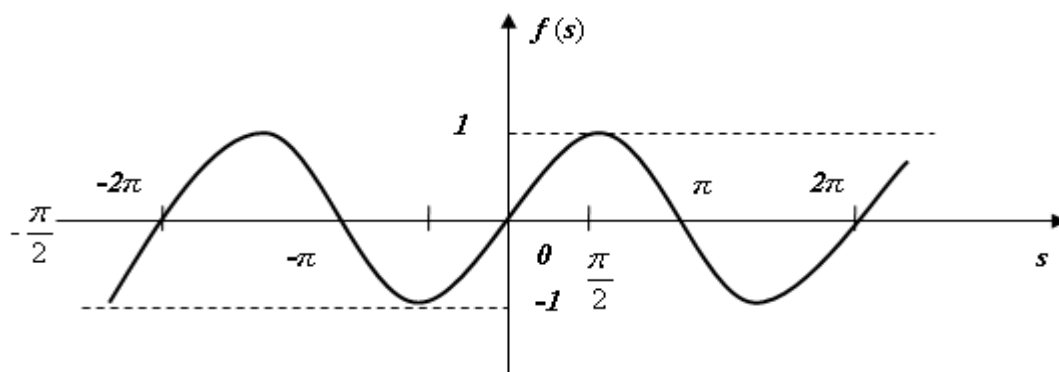
1. Синусоидалық белсендіру функциясының периодын көрсетіңіз.
2. Синусоидалық белсендіру функциясының түрлерін атаңыз.
3. Синусоидалық белсендіру функциясының өрнегін жазыңыз.
4. $\sin(x)$ белсендіру функциясының графигін салыңыз.
5. $\cos(x)$ белсендіру функциясының графигін салыңыз.

II.2.7-сұрақтар.

1. Синусоидалық белсендіру функциясының анықталу аймағы мен мәндер аймағы не?
2. Синусоидалық белсендіру функциясының мәндер аймағы не?
3. Синусоидалық белсендіру функциясы қандай есептер үшін пайдалы бола алады?
4. Төмендегі суретте қандай белсендіру функциясы берілген?



5. Төмендегі суретте қандай белсендіру функциясы берілген?



II.2.7-тестілер.

1. Синусоидалық белсендіру функциясы қандай өрнекпен беріледі?

- A) $f(s) = s \ln(s)$ немесе $f(s) = \cos(s)$;
- B) $f(s) = s \ln(s \ln(s))$ немесе $f(s) = \cos(s)$;
- C) $f = s \ln(s)$ немесе $f = \cos(s)$;
- D) $f(s) = \frac{s}{|s|}$;
- E) $f(s) = \frac{s}{k}$.

2. Синусоидалық белсендіру функциясының анықталу аймағы қайсысы?

- A) $D(f) = (-\infty, +\infty)$;
- B) $D(f) = (0, 1)$;
- C) $D(f) = (0, 0)$;
- D) $D(f) = (-1, 1)$;
- E) $D(f) = (-\infty, 1)$.

3. Синусоидалық белсендіру функциясының мәндерінің аймағы қайсысы?

- A) $E(f) = (-1, +1)$;
- B) $E(f) = (-1, +\infty)$;
- C) $E(f) = (-\infty, +1)$;
- D) $E(f) = (-\infty, +\infty)$;
- E) $E(f) = (-1, 0)$.

4. Синусоидалық белсендіру функциясы қандай деректерді тану үшін пайдалы?

- A) радиалды таралған;
- B) таралған;
- C) тұрақты;
- D) сандық;
- E) радиалды таралмаған.

5. Синусоидалық белсендіру функциясының периоды қайсысы?

- A) 2π ;
- B) π ;
- C) $f(s) = \frac{s}{|s|}$;
- D) 3π ;
- E) $f(s) = \frac{s}{k}$.

II.2.8. Дәрежелі белсендіру функциясы

Бұл тақырыпта дәрежелі белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Дәрежелі белсендіру функциясы келесі өрнекпен беріледі:

$$f(s) = s^n,$$

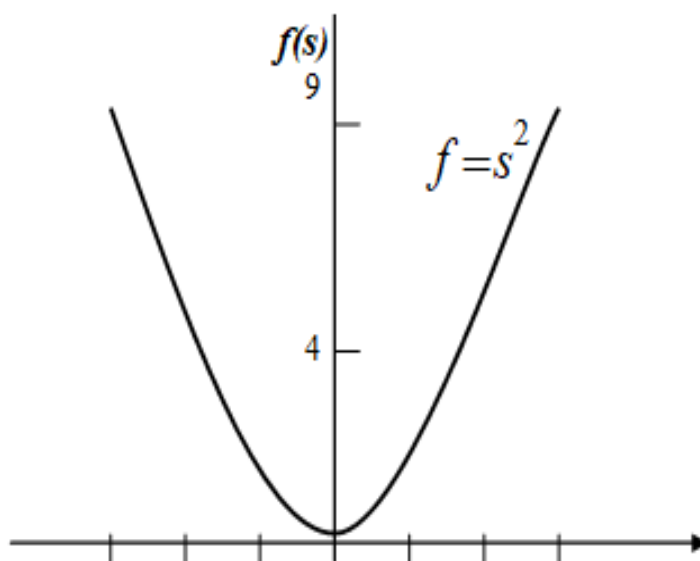
мұнда тұрақты шама $n = 2, 3, 4, \dots$ – дәреже көрсеткіші.

Дәрежелі белсендіру функциясында анықтау аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы дәреженің жұптық көрсеткішіне байланысты: егер көрсеткіш мәні жұп болса, онда мәндер аймағы $E(f) = (0, +\infty)$, кері жағдайда $E(f) = (-\infty, +\infty)$ болады.

Белсендіру функциясына қойылатын монотондықтың, үздіксіздіктің және дифференциалданулықтың талаптарына тек тақ көрсеткішті дәрежесі бар дәрежелі функциялар толықтай жауап береді.

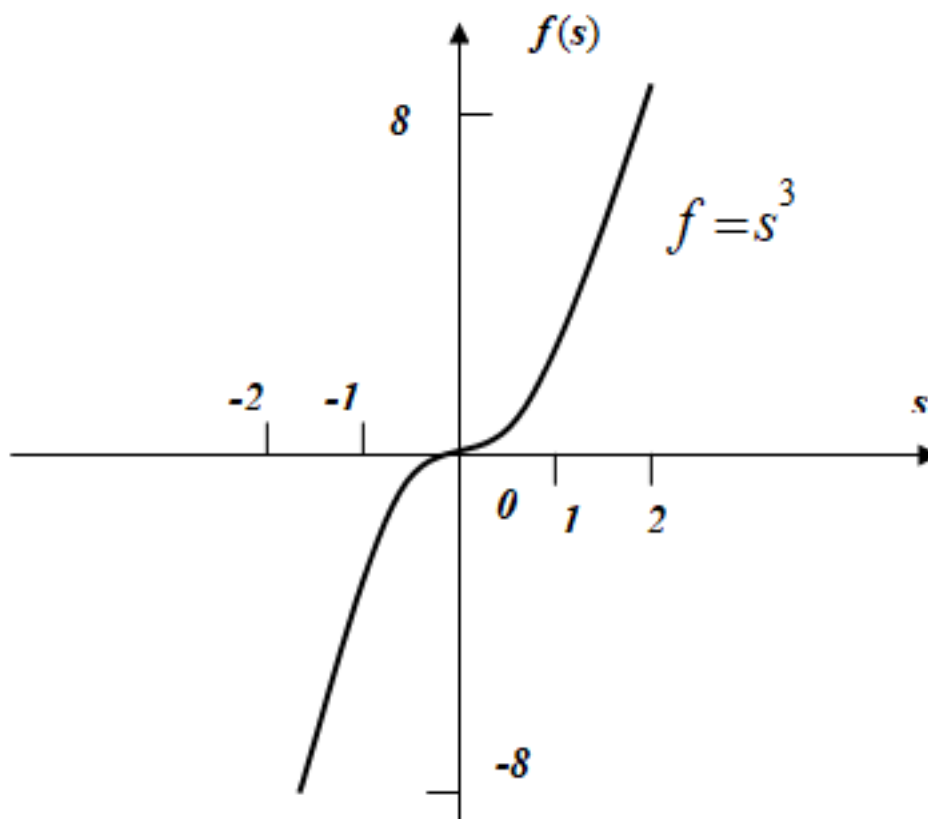
II.2.8. Мысалдар.

1. Дәрежелі функцияда, егер $n=2$ болса, онда квадратты белсендіру функциясы туралы айтылады. Квадратты белсендіру функциясының графигі II.2.8.1-суретте көрсетілген.



II.2.8.1-сурет. Квадратты белсендіру функциясының графигі

2. Дәрежелі функцияда $n=3$ болғанда, кубтық квадраттық белсендіру функциясымен жұмыс жасаймыз. Кубтық квадраттық белсендіру функциясының графигі II.2.8.2-суретте көрсетілген.



II.2.8.2-сурет. Кубтық квадраттық белсендіру функциясының графигі

II.2.8-тапсырмалар.

1. Жалпы түрдегі дәрежелі белсендіру функциясының өрнегін жазыңыз.
2. Квадраттық белсендіру функциясының анықтау аймағын көрсетіңіз.
3. Квадраттық белсендіру функциясының графигін салыңыз.
4. Кубтық белсендіру функциясының анықтау аймағын көрсетіңіз.
5. Квадраттық белсендіру функциясының графигінің кубтық белсендіру функциясының графигінен айырмашылығын айтыңыз.

II.2.8-сұрақтар.

1. Монотондық, үздіксіздік пен дифференциалданулық талаптары неге қатысты болады?
2. Дәрежелі функцияның белсендіру функциясында мәндер аймағы қалай анықталады, егер дәреже көрсеткішінің мәні жұп болса?
3. Дәрежелі функцияның белсендіру функциясында мәндер аймағы қалай анықталады, егер дәреже көрсеткішінің мәні тақ болса?
4. Квадраттық белсендіру функциясында мәндер аймағы болып не табылады?
5. Кубтық белсендіру функциясында мәндер аймағы қандай?

II.2.8-тестілер.

1. Дәрежелі белсендіру функциясы қандай өрнекпен беріледі?
A) $f(s) = s^n$;
B) $f(s) = s$;
C) $f(s) = \frac{s}{|s|}$;
D) $f(s) = \frac{s}{k}$;
E) $f(s) = \frac{s^m}{k}$.
2. Дәрежелі белсендіру функциясында мәндер аймағы қандай көрсеткішке байланысты?
A) дәреженің жұптық көрсеткішіне;
B) дәреженің тақтық көрсеткішіне;
C) дәреженің үздіксіз көрсеткішіне;
D) дәреженің тек қана 2-лік көрсеткішіне;
E) дәреженің кез-келген көрсеткішіне.

3. Егер дәрежелі белсендіру функциясында көрсеткіш мәні жұп болса, онда оның мәндер аймағы қайсысы?

- A) $E(f) = (0, +\infty)$;
- B) $E(f) = (-\infty, +\infty)$;
- C) $E(f) = (-\infty, 1)$;
- D) $E(f) = (-1, +1)$;
- E) $E(f) = (-\infty, +0)$.

4. Белсендіру функциясына қойылатын қандай талаптарына тек тақ көрсеткішті дәрежесі бар дәрежелі функциялар толықтай жауап береді?

- A) монотондық, үздіксіздік және дифференциалданулық;
- B) монотондық, үздіксіздік және сигматикалық;
- C) монотондық, синусоидалық және дифференциалданулық;
- D) сигмоидальды, үздіксіздік және дифференциалданулық;
- E) логистикалық, үздіксіздік және дифференциалданулық.

5. Кубтық белсендіру функциясы қандай формуламен беріледі?

- A) $f(s) = s^3$;
- B) $f(s) = s^n$;
- C) $f(s) = s^2$;
- D) $f(s) = \frac{s}{k}$;
- E) $f(s) = \frac{s^m}{k}$.

II.2.9. Квадраттық түбір белсендіру функциясы

Бұл тақырыпта квадраттық түбір белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Квадраттық түбір белсендіру функциясы келесі өрнекпен беріледі:

$$f(s) = \sqrt{s}$$

Квадраттық түбір белсендіру функциясының анықталу аймағы $D(f) = (0, +\infty)$, ал мәндерінің аймағы $E(f) = (0, +\infty)$ болады.

Квадраттық түбір белсендіру функциясы Кохоненнің нейрондық желісінде қолданылады. Квадраттық түбір Кохонен желісін белсендіруге, яғни ара қашықтардың квадраты ара қашықтықты ұсынатын шығыс мәндерге түрленеді. Кохонен желісі мұғалімсіз үйретуге негізделген және деректерді шолу арқылы талдауда қолданылады (деректерде кластерлерді тану, сыныптар жақындығын орнату, сыныптау есептерін шешу, жаңа құбылыстарды табу).

II.2.9-тапсырмалар.

1. Квадраттық түбір белсендіру функциясын ұсынатын өрнекті жазыңыз.
2. Квадраттық түбір белсендіру функциясының анықталу аймағын көрсетіңіз.
3. Квадраттық түбір белсендіру функциясының мәндер аймағын көрсетіңіз.

II.2.9-сұрақтар.

1. Қандай нейрондық желілерде квадраттық түбір белсендіру функциясы қолданылады?
2. Квадраттық түбір белсендіру функциясы Кохонен нейрондық желісімен не істейді?

3. Ара қашықтардың квадраты неліктен ара қашықтықты ұсынатын шығыс мәндерге түрленеді?

II.2.9-тестілер.

1. Квадраттық түбір белсендіру функциясы қандай өрнекпен беріледі?

A) $f(s) = \sqrt{s}$;

B) $f(s) = s^n$;

C) $f(s) = s^2$;

D) $f(s) = \frac{s}{k}$;

E) $f(s) = \frac{s^m}{k}$.

2. Квадраттық түбір белсендіру функциясының анықталу аймағы қайсысы?

A) $D(f) = (0, +\infty)$;

B) $D(f) = (0, 1)$;

C) $D(f) = (0, 0)$;

D) $D(f) = (-1, 1)$;

E) $D(f) = (-\infty, 1)$.

3. Квадраттық түбір белсендіру функциясының мәндерінің аймағы қайсысы?

A) $E(f) = (0, +\infty)$;

B) $E(f) = (-1, +\infty)$;

C) $E(f) = (-\infty, +1)$;

D) $E(f) = (-\infty, +\infty)$;

E) $E(f) = (-1, 0)$.

4. Квадраттық түбір белсендіру функциясы қайда қолданылады?

A) Кохоненнің нейрондық желісінде;

B) арнайы құрастырылған көп қабатты персептрондардың шығу қабатында;

С) интерпретациялауға болатын арнайы құрастырылған көп қабатты персептрондардың шығу қабатында;

Д) Радиалды ішкі синапстық потенциалды функцияның (PSP-функциялар) комбинациясында;

Е) Гауссианның стандартты ауытқуында.

5. Квадраттық түбір белсендіру функциясы Кохонен нейрондық желісімен не істейді?

А) ара қашықтықтар квадратын ара қашықтықты ұсынатын шығыс мәндерге түрлендіреді;

В) ара қашықтар квадратын ара қашықтықты ұсынатын шығыс мәндерге түрлендірмейді;

С) ара қашықтар квадратын ара қашықтықты ұсынатын шығыс және кіріс мәндерге түрлендірмейді;

Д) ара қашықтар квадратын ара қашықтықты ұсынатын кіріс мәндерге түрлендіреді;

Е) ара қашықтар квадратын ара қашықтықты ұсынатын шығыс және кіріс мәндерге түрлендіреді.

II.2.10. Экспоненциалды белсендіру функциясы

Бұл тақырыпта экспоненциалды белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Экспоненциалды белсендіру функциясы келесі түрде болады:

$$f(s) = e^s,$$

мұнда $e = 2,71$ – тұрақты шама.

Экспоненциалды белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, +\infty)$ болады.

Экспоненциалды белсендіру функциясының графигі төменгі II.2.10-суретте көрсетілген.



II.2.10-сурет. Экспоненциалды белсендіру функциясының графигі

Экспоненциалды белсендіру функциясы ықтималды нейрондық желілердегі нейрондарының бірінші жасырын қабатында қолданылады. Осындай желілердің артықшылығы, олардың архитектурасы объектілерді тек сыныптап қана қоймай, сонымен қоса қабылданатын шешімдердің әділдік ықтималдылығын анықтаудан байқалады.

Кейде шығыстары қабатты белсендірудің барлық қосындысы 1-ге тең болатындай нормаланған экспоненциалды функцияларды қолданады. Оларды сыныптау есептері үшін арнайы құрастырылған көп қабатты персептрондардың шығыс қабаттарында пайдаланады,

мұнда шығыстарды сыныпқа тән болу ықтималдығы ретінде интерпретациялайды.

Экспоненциалды белсендіру функциясы радиалды элементтер үшін өте жақсы келеді. Радиалды синапстық потенциалды функцияның (PSP-функциялар) комбинациясы және теріс экспоненциалды белсендіру функциясы салмақтар векторының ортасы болатын гаустық (қоңырау сияқты) функцияны моделдейтін элементті береді. Гауссианның стандартты ауытқуы $\sigma = \sqrt{1/d}$ формуласы бойынша анықталады, мұнда d арқылы оның межелік мәні ретінде сақталған элементтің «ауытқуы» белгіленген.

PSP-функция кіріс сигналдарға, оның салмағына және межелігіне қолданылатындығын және осы элементтің белсендіру деңгейін беретіндігін еске түсірейік. Сызықтық (салмақталған кірістер сомасы минус меже) және радиалды (салмақтар векторынан кірістер векторына дейінгі масштабталған арақашықтық квадраты) PSP-функциялар жиірек қолданылады.

II.2.10-тапсырмалар.

1. Экспоненциалды белсендіру функциясын көрсету үшін өрнекті жазыңыз.
2. Экспоненциалды белсендіру функциясының анықталу аймағын көрсетіңіз.
3. Экспоненциалды белсендіру функциясының графигін салыңыз.

II.2.10-сұрақтар.

1. Қандай нейрондық желілілерде экспоненциалды белсендіру функциясы қолданылады?
2. Экспоненциалды белсендіру функциясының мәндер аймағы болып не табылады?
3. Төмендегі суретте қандай белсендіру функциясының графигі салынған?



II.2.10-тестілер.

1. Экспоненциалды белсендіру функциясының өрнегі қайсысы?

A) $f(s) = e^s$;

B) $f(s) = \sqrt{s}$;

C) $f(s) = s^n$;

D) $f(s) = \frac{s}{|s|}$;

E) $f(s) = \frac{s}{k}$.

2. Экспоненциалды белсендіру функциясының формуласы $f(s) = e^s$, мұнда e неге тең?

A) =2,71; тұрақты шама;

B) =2,14; тұрақты шама;

C) =2,73; тұрақты шама;

D) кез-келген мәнге тең;

E) алдын ала анықталған тұрақты мән.

3. Экспоненциалды белсендіру функциясының анықталу аймағы қандай?

A) $D(f) = (-\infty, +\infty)$;

B) $D(f) = (0, 1)$;

C) $D(f) = (0, 0)$;

D) $D(f) = (-1, 1)$;

E) $D(f) = (-\infty, 1)$.

4. Экспоненциалды белсендіру функциясы қандай элементтер үшін өте жақсы келеді?

- A) радиалды;
- B) сақталған;
- C) белсенді;
- D) кез-келген;
- E) стандартты.

5. Гауссианның стандартты ауытқуы формуласы қандай формула бойынша анықталады?

- A) $\sigma = \sqrt{1/d}$;
- B) $f(s) = e^s$;
- C) $\sigma = \sqrt{1/s}$;
- D) $f(s) = \frac{s}{|s|}$;
- E) $\sigma = \sqrt{s}$.

II.2.11. Кіріс мәндерін азайтатын белсендіру функциясы

Бұл тақырыпта кіріс мәндерін азайтатын белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Кіріс мәндерін азайтатын белсендіру функциясы (*softmax*) экспоненциалды функция болады және нейрондарды белсендірудің барлық қосындысы бірге тең болатындай етіп шығу мәнін қалыптастырады.

Көп қабатты нейрондық желілерде, *softmax*-топтың әрбір *i*-ші нейронының белсендіру функциясы болады, ол келесі өрнекпен беріледі:

$$y_i = \frac{e^{s_i}}{\sum_j e^{s_j}}, \quad (\text{II.2.11.1})$$

мұнда $i \in N$ – нейронның индексі, $j \in N$ – қабаттың нөмірі.

Осы *softmax* белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, +1)$ болады.

(II.2.1.1)-өрнегінен әрбір нейронның шығысы *softmax*-топтағы барлық қалған нейрондардың қосқышынан тәуелді болатындығы, ал барлық топтың шығыс мәндерінің қосындысы бірге тең болатындығы байқалады.

Әрбір оқыту мысалдарында өзімізге қажетті ықтималдық үлестіруді моделдейтін желі шығысын алатын боламыз, ал екі ықтималдық үлестірулерді салыстыру үшін дұрыс өлшем қажет. Мұндай өлшем ретінде тоғыспалы энтропия қолданылатын болады:

$$C = -\sum_j d_j \log y_j, \quad (\text{II.2.11.2})$$

мұнда d_j – *j* үйрету мысалы үшін талап етілетін шығыстар, y_j – *j* үйрету мысалы үшін нақты шығыстар.

Бұл *softmax* белсендіру функциясы нейронның шығысын нақты сыныпқа тиістілік ықтималдығы ретінде түсіндіру керек болатын сыныптау есептерін шешу үшін, яғни нейрондық желімен ықтималдық үлестіруді моделдеу тәсілін қамтамасыз ету қажет

болғанда қолданылады. Ол үшін тура таратудың келесідегідей көп қабатты нейрондық желісі қолданылады:

- желі қандай да бір жасырын қабаттардан тұрады, барлық нейрондардың өзінің жеке белсендіру функциясы болуы мүмкін;
- соңғы қабатта сыныптар санына сәйкес келетін нейрондар саны болады; бұл барлық нейрондар softmax-тобы деп аталады.

II.2.11-тапсырмалар.

1. softmax белсендіру функциясы үшін формуланы жазыңыз.
2. softmax белсендіру функциясының анықталу және мәндер аймағын көрсетіңіз.
3. Тоғыспалы энтропия үшін формуланы жазыңыз.
4. Нейрондық желінің соңғы қабатында сыныптар санына сәйкес келетін нейрондар саны бар болатын топтың атын атаңыз.
5. «Нейрондық желіге ықтималдық үлестірудің нейрондық желісін моделдеу тәсілін қамтамасыз ету қажет болатын» шартты түсіндіріп беріңіз.
6. «Тоғыспалы энтропия» ұғымын түсіндіріп беріңіз.

II.2.11-сұрақтар.

1. softmax белсендіру функциясы экспоненциалды функция болады ма?
2. softmax белсендіру функциясы қандай мақсатпен жасанды нейронның мәнін тұрақтандырады?
3. softmax-топтың барлық басқа нейрондарының қосқышынан не тәуелді болады?
4. Барлық топтың шығыс мәндерінің қосындысы неге тең?
5. softmax белсендіру функциясы қандай мақсатта қолданылады?
6. «Нейронның шығысын нақты сыныпқа тиістілік ықтималдығы ретінде түсіндіру керек» шарты нені білдіреді?
7. Тура таратудың көп қабатты нейрондық желісі не үшін қолдналады?

II.2.11-тестілер.

1. Көп қабатты нейрондық желілерде, softmax-топтың әрбір І-ші нейронының белсендіру функциясы болады, ол қандай өрнекпен беріледі:

A) $y_i = \frac{e^{s_i}}{\sum_j e^{s_j}} ;$

B) $y_i = \frac{e}{\sum_j e^{s_j}} ;$

C) $y_i = \frac{e^{s_i}}{\sum_j e} ;$

D) $y_i = \frac{e^{s_i}}{\sum_j s} ;$

E) $y_i = \frac{e^i}{\sum e^{s_j}} .$

2. softmax белсендіру функциясының анықталу аймағы?

A) $D(f) = (-\infty, +\infty);$

B) $D(f) = (0, 1);$

C) $D(f) = (0, 0);$

D) $D(f) = (-1, 1);$

E) $D(f) = (-\infty, 1).$

3. softmax белсендіру функциясының мәндер аймағы?

A) $E(f) = (0, +1);$

B) $E(f) = (-1, +\infty);$

C) $E(f) = (-\infty, +1);$

D) $E(f) = (-\infty, +\infty);$

E) $E(f) = (-1, 0).$

4. Екі ықтималдық үлестіруді салыстыру үшін қолданылатын тоғыспалы энтропия өрнегі?

A) $C = -\sum_j d_j \log y_j ;$

B) $C = -\sum_j \log y_j ;$

C) $C = -\sum_j d \log y_j ;$

D) $C = -\sum_j d_j ;$

E) $C = \sum_j d_j \log y_j .$

5. $C = -\sum_j d_j \log y_j$, тоғыспалы энтропиядағы d_j нені сипаттайды?

A) үйрету мысалы үшін талап етілетін шығыстар;

B) үйрету мысалы үшін нақты шығыстар;

C) үйрету мысалы үшін тұрақты шығыстар;

D) үйрету мысалы үшін шығыстар;

E) үйрету мысалы үшін кез-келген шығыстар.

II.2.12. Құлдилау айнымалысы бар белсендіру функциясы

Бұл тақырыпта құлдилау айнымалысы бар белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Құлдилауы айнымалысы бар дәрежелі белсендіру функциясы келесі түрдегі өрнекпен анықталады:

$$y = k \left(\sum_{i=1}^n x_i w_i + w_0 \right)^t,$$

Бұл моделде k коэффициенті t дәрежелі параболаның құлдилауын өзгертеді. Ол жеке нейронның функционалдық мүмкіншіліктерін барынша кеңейтеді. Осындай нейронның тікелей жүзеге асырылуы w_I және k екі параметрдің жетілдіруін талап етеді. Нейронның бапталу уақытын, соңында нақты есептің шешілу уақытын да азайту үшін осы моделдің өрнегінің түрін өзгерте отырып, k жетілдіруінен бас тартуға болады. Ол үшін k коэффициентін қосынды белгісінің астына енгізу қажет, яғни келесі түрде жазу керек:

$$y = k \left(\sum_{i=1}^n x_i w_i + w_0 \right)^t = \left(\sum_{i=1}^n x_i k^t w_i + k^t w_0 \right)^t = \left(\sum_{i=1}^n x_i u_i + u_0 \right)^t,$$

мұнда $u_j = k^t w_j$, $j=0, 1, 2, \dots, t$ коэффициенттері, ол нейронның синапстық коэффициенттерінің жалпы жетілдіруі кезінде құлдилауды түрлендіруге мүмкіндік береді.

Сонымен қоса, сигмоидалды, экспоненциалды және синусоидалды белсендіру функциялары үшін k сигмоидтың құлдилау коэффициентін қолдану пайдалы. Сонда оларды анықтау өрнегі келесі түрде жазылады:

$$f(s) = \frac{1}{1+e^{-ks}} \quad - \quad \text{ығыстырылған сигмоидалды белсендіру}$$

функция үшін;

$$f(s) = \frac{e^{ks} - e^{-ks}}{e^{ks} + e^{-ks}} \quad - \quad \text{гиперболалық белсендіру функциясы үшін;}$$

$f(s) = \frac{s}{k + |s|}$ – рационалды сигмоидалды белсендіру функция

үшін;

$f(s) = e^{-ks}$ –экспоненциалды белсендіру функция үшін;

$f(s) = \sin(ks)$ –синусоидалы белсендіру функция үшін.

Тәжірибеде белсендіру функцияларының жоғарыда аталғандардан басқаларын қолданады. Әрбір нақты жағдайда өңдеуші өз таңдауын дәлелдеу қажет:

1) Ықтималды нейрондық желілерде бірлік қосынды деректерді қайта масштабтау үшін барлық мәндердің қосындысы 1-ге тең болуы қажет. Бұл белсендіру функция келесі өрнекпен сипатталады:

$$f(s) = \frac{s}{\sum_i s_i}$$

Осы белсендіру функцияның мәндер аймағы $E(f) = (0, +1)$.

2) квадраттық түбір $\sqrt{s}$ арақашықтықтардың квадратын шығу мәндеріне өзгертеді. Арақашықтықтар келесі түрде болады:

$$\sigma = \sqrt{\frac{1}{d}}$$

Енді жоғарыда қарастырылған белсендіру функциялары бар нейрондардың кемшілігі мен артықшылықтарын қарастыруға болады.

Сатылық (межелік) белсендіру функциясы бар нейрондардың негізгі кемшілігіне шешілетін есепке баптау және нейрондық желіні оқыту кезінде жеткілікті икемділіктің болмауы жатады. Себебі, егер

$\sum_{i=1}^n x_i w_i$ мәні межеге аздап жетпесе де, онда шығу сигналы қалыптаспайды және нейрон «жұмыс істемейді». Бұл аталған нейронның шығу сигналының қарқындылығын жоғалтады деген

сөз. Сондықтан, нейрондардың келесі қабатында салмақталған кірісте деңгейдің жоғары емес мәні құрылады.

Осындай кемшіліктен көп жағдайда сызықтық (бөлікті-сызықты) белсендіру функциясы айырылған, оны іске асыру жоғары емес есептеу күрделілігін қамтамасыз етеді.

Сигмоидалды функция сатылық және сызықтық функциялардың артықшылықтарын қамтиды. Ол сатылық функцияға ұқсас сызықты емес және бұл зерттелінетін объектілердің іздеу кеңістігінде күрделі пішінді аймақты ерекшелеуге мүмкіндік береді, соның ішінде дөңессіз және байланыссыз. Басқа жағынан, сатылы функцияға қарағанда, ол кіріс сигналының бір мәнінен басқасына үздіксіз өтуге мүмкіндік береді. Алайда, белсендіру функцияларының кез-келгенін жуықтау ретінде қарастыру керек. Олардың архитектурасының күрделілігі мен параметрлерін нақты есептерді шешуге баптау қиындығын ескере отырып, анағұрлым икемді еркін сызықтық емес функцияларға көшу қажет.

II.2.12-тапсырмалар.

1. Құлдилау айнымалысы бар дәрежелі белсендіру функция үшін өрнек жазыңыз.

2. Қосынды белгісінің астына k коэффициентін енгізу үшін өрнектің түрлендірілуін жазыңыз.

3. Сигмоидалды, экспоненциалды және синусоидалды белсендіру функциялар үшін анықталатын өрнектерді жазыңыз.

4. Құлдилау айнымалысы бар дәрежелі белсендіру функция үшін өрнек жазыңыз.

5. Ықтималды нейрондық желілерде бірлік қосындының функциясы үшін өрнекті жазыңыз.

II.2.12-сұрақтар.

1. Құлдилау айнымалысы белсендіру функциясының жазбасында k коэффициенті қандай қызмет атқарады?

2. Сатылы (межелік) белсендіру функциясы бар нейрондар қандай кемшіліктерге ие?

3. Сигмоидалды функция сатылы және сызықтық функциялардың артықшылықтарын қамти ма?

II.2.12-тестілер.

1. Құлдилау айнымалысы бар дәрежелі белсендіру функциялар қандай түрдегі өрнекпен анықталады?

A) $y = k \left(\sum_{i=1}^n x_i w_i + w_0 \right)^t ;$

B) $y = k \left(\sum_{i=1}^n x_i w_i \right) ;$

C) $f(s) = \frac{s}{\sum_i s_i} ;$

D) $y = k \left(\sum_{i=1}^n x_i w_i \right)^t ;$

E) $\sigma = \sqrt{\frac{1}{d}} .$

2. Рационалды сигмоидалды белсендіру функциялар үшін анықталатын өрнек қалай жазылады?

A) $f(s) = \frac{s}{k + |s|} ;$

B) $f(s) = e^{-ks} ;$

C) $f(s) = \sin(ks) ;$

D) $f(s) = \frac{e^{ks} - e^{-ks}}{e^{ks} + e^{-ks}} ;$

E) $f(s) = \frac{1}{1 + e^{-ks}} .$

3. Гиперболалық белсендіру функциялар үшін анықталатын өрнек қалай жазылады?

$$A) f(s) = \frac{e^{ks} - e^{-ks}}{e^{ks} + e^{-ks}};$$

$$B) f(s) = e^{-ks};$$

$$C) f(s) = \frac{s}{k + |s|};$$

$$D) f(s) = \sin(ks);$$

$$E) f(s) = \frac{1}{1 + e^{-ks}}.$$

4. Квадраттық түбір $\sqrt{s}$ арақашықтық квадратын шығыс мәндеріне өзгертеді. Арақашықтық қандай түрде болады?

$$A) \sigma = \sqrt{\frac{1}{d}};$$

$$B) \sigma = \sqrt{d};$$

$$C) \sigma = \sqrt{\frac{s}{d}};$$

$$D) \sigma = \sqrt{\frac{1}{d}};$$

$$E) \sigma = \sqrt{\frac{2}{d}}.$$

5. Синусоидалды белсендіру функциялар үшін анықталатын өрнек қалай жазылады?

$$A) f(s) = \sin(ks);$$

$$B) f(s) = e^{-ks};$$

$$C) f(s) = \frac{s}{k + |s|};$$

$$D) f(s) = \frac{e^{ks} - e^{-ks}}{e^{ks} + e^{-ks}};$$

$$E) f(s) = \frac{1}{1+e^{-ks}}.$$

II.2.13. Радиалды-базистік белсендіру функциясы

Бұл тақырыпта радиалды-базистік белсендіру функциясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Радиалды-базистік белсендіру функциясы аргумент ретінде кіріс векторлары және қандай да бір алдын ала берілген белсендіру функциясының ортасы арасындағы ара қашықтықты алады. Бұл функцияның мәні неғұрлым жоғары болса, соғұрлым ортаға кіру векторына жақын болады. Радиалды-базистік функция моделдерінде векторлар арасындағы қашықтықты өлшеудің әр түрлі тәсілдері, сондай-ақ жасырын қабаттың нейрондарын белсендіру функциялары қолданылуы мүмкін.

Радиалды базистік функциялардың көптеген сыныптары бар, олардың ішінде келесілері аса көп қызығушылық тудырады:

1. Мультикватраттық функция:

$$f(s) = (s^2 + c^2)^{\frac{1}{2}}, \quad c > 0.$$

2. Кері мультикватраттық функция:

$$f(s) = \frac{1}{(s^2 + c^2)^{\frac{1}{2}}}, \quad c > 0.$$

3. Гаусс функциясы:

$$f(s) = e^{-s^2/2\sigma^2}, \quad \sigma > 0.$$

мұнда s – нейронның кіру сигналдарының векторы, σ – функция терезесінің ені, $f(s)$ – кемітуші функция (көбінесе, қандай да бір кесіндіден тыс жерде нөлге тең). Қызығы, 2 және 3 функциялары локализацияланған болады, яғни $s \rightarrow 0$ болғанда, ал 1 функция керісінше s өскенде ол да шексіз өседі, ол оған RBF желісінде қолдану үшін жарамды функция болуына кедергі емес.

Радиалды-базистік функцияның анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, +1)$ болады.

Үш артықшылығымен сипатталады:

1. Жалғыз жасырын қабат;
2. Жасырын қабаттардың нейрондары ғана сызықты емес белсендіру функциядан тұрады;
3. Кіріс және шығыс қабаттардың барлық синапстық байланыстар қосындысы бірге тең.

Радиалды-базистік функция ретінде көбінесе Гаусстың жетілдірілген функциясын қолданады:

$$y = f(s) = e^{-(S-R)^2/2\sigma^2},$$

мұнда $S = \|X-C\|$ – C ортасы мен X кіріс сингалдар векторының арасындағы қашықтық. σ скалярлық параметрі центрден векторды жою кезіндегі функцияның ортаю жылдамдығын анықтайды және *терезенің ені* деп аталады, R параметрі абцисса осі бойынша белсендіру функциясының ығысуын анықтайды.

Векторлар арасындағы қашықтық ретінде әр түрлі метрикалар қолданылуы мүмкін, көбінесе евклидтік қашықтық қолданылады:

$$S = \sqrt{\sum_{j=1}^n (x_j - c_j)^2},$$

мұнда x_j – нейрон кірісіне берілетін вектордың j -ші компоненті, ал c_j – табыстама функциясының орталығының орналасуын анықтайтын вектордың j -ші компоненті.

Радиалдық базистік белсендіру функциясы, оның кірісі 0-ге тең болғанда ең үлкен мән 1-ге ие болады. x_{ij} және c_{ij} арасындағы қашықтық азайған кезде, шығыс арта түседі. Ендеше, радиалдық базистік нейрон датчик ретінде жұмыс жасайды, ол x_{ij} кірісі c_{ij} радиалдық функциясы орталығының векторына ұқсас болғанда әрдайым 1 мәнін береді. Бұл кеңістікті шеңберлерге немесе (жалпы жағдайда) гиперсфераларға бөлу мүмкіндігін береді. Мұндай

нейрондары бар желілер ықтималдылықты немесе регрессиялық деп аталады.

Нақты желілерде мұндай нейрондарды белсендіру функциялары қандай да бір кездейсоқ шаманың ықтималдығын үлестіруді бейнелей алады, немесе шамалар арасындағы эвристикалық тәуелділікті белгілей алады.

Радиалды-базистік функцияларды қолданатын нейрондар, радиалды симметриялы деп аталады. Олар сыныптау есептерін шешу үшін қолданылады және оқытушы векторлардың саны көп болғанда аса тиімді болады.

Радиалдық базистік функциялардың нейрондық желілері көп қолданысқа ие, соның ішінде жуықтау функциялары, уақыттық қатарларды болжау, сыныптау және басқару жүйелерінде.

II.2.13-тапсырмалар.

1. Мультиквадраттық белсендіру функциясының формуласын жазыңыз.

2. Кері мультиквадраттық белсендіру функциясының формуласын жазыңыз.

3. Радиалды-базистік белсендіру функциясының формуласын жазыңыз.

4. Евклидтік қашықтық формуласын жазыңыз.

5. Гаусстың белсендіру функциясының формуласын жазыңыз.

6. Гаусстың жетілдірілген белсендіру функциясының формуласын жазыңыз.

7. Нақты желілерде радиалды-базистік белсендіру функциясы орындайтын әрекеттерді сипаттаңыз.

II.2.13-сұрақтар.

1. Радиалды-базистік белсендіру функциясы аргумент ретінде нені қабылдайды?

2. Кіру векторының ортаға жақындығына байланысты радиалды-базистік белсендіру функциясының мәні қалай өзгереді?

3. Радиалдық базистік функциясында қандай сыныптар болады?

4. Евклидтік қашықтықты қолдану қандай мүмкіндіктер береді?

5. Радиалды-базисті нейрондық желілер қандай ерекшеліктерімен сипатталады?

6. Қандай нейрондық желілер ықтималдылықты немесе регрессиялық деп аталады?

7. Радиалды-базистік функцияларды қолданатын нейрондар қандай есептерді шешу үшін қолданылады?

II.2.13-тестілер.

1. Радиалды базистік функциялардың сыныбына жататын мультиквадраттық функцияның өрнегі қалай жазылады?

A) $f(s) = (s^2 + c^2)^{\frac{1}{2}}, \quad c > 0;$

B) $f(s) = \frac{1}{(s^2 + c^2)^{\frac{1}{2}}}, \quad c > 0;$

C) $f(s) = e^{-s^2/2\sigma^2}, \quad \sigma > 0;$

D) $y = f(s) = e^{-(S-R)^2/2\sigma^2};$

E) $f(s) = (s^2 + c^2)^2, \quad c > 0.$

2. Радиалды базистік функциялардың сыныбына жататын кері мультиквадраттық функцияның өрнегі қалай анықталады?

A) $f(s) = \frac{1}{(s^2 + c^2)^{\frac{1}{2}}}, \quad c > 0;$

B) $f(s) = (s^2 + c^2)^{\frac{1}{2}}, \quad c > 0;$

C) $f(s) = e^{-s^2/2\sigma^2}, \quad \sigma > 0;$

D) $y = f(s) = e^{-(S-R)^2/2\sigma^2};$

E) $f(s) = (s^2 + c^2)^2, \quad c > 0.$

3. Радиалды базистік функциялардың сыныбына жататын Гаусс функцияның өрнегін қалай анықталады?

A) $f(s) = e^{-s^2/2\sigma^2}, \sigma > 0;$

B) $f(s) = (s^2 + c^2)^{\frac{1}{2}}, c > 0;$

C) $f(s) = \frac{1}{(s^2 + c^2)^{\frac{1}{2}}}, c > 0;$

D) $y = f(s) = e^{-(s-R)^2/2\sigma^2};$

E) $f(s) = (s^2 + c^2)^2, c > 0.$

4. Радиалды-базистік функцияның артықшылықтары неде?

A) Жалғыз жасырын қабат, жасырын қабаттардың нейрондары ғана сызықты емес белсендіру функциядан тұрады, кіріс және шығыс қабаттардың барлық синапстық байланыстар қосындысы бірге тең;

B) Жалғыз жасырын қабат, жасырын қабаттардың нейрондары ғана сызықты емес белсендіру функциядан тұрады;

C) Жасырын қабаттардың нейрондары ғана сызықты емес белсендіру функциядан тұрады, кіріс және шығыс қабаттардың барлық синапстық байланыстар қосындысы бірге тең;

D) Жасырын қабаттың нейрондарын белсендіру функциялары қолданылуы мүмкін;

E) Жасырын қабаттардың нейрондары ғана сызықты емес белсендіру функциядан тұрады, жасырын қабаттың нейрондарын белсендіру функциялары қолданылуы мүмкін.

5. Радиалды-базистік функция ретінде қолданылатын Гаусстың жетілдірілген функциясының өрнегі қалай жазылады?

A) $y = f(s) = e^{-(s-R)^2/2\sigma^2}$

B) $f(s) = (s^2 + c^2)^{\frac{1}{2}}, c > 0;$

C) $f(s) = \frac{1}{(s^2 + c^2)^{\frac{1}{2}}}, c > 0$

$$D) f(s) = e^{-s^2/2\sigma^2}, \sigma > 0;$$

$$E) f(s) = (s^2 + c^2)^2, c > 0$$

II.2.14. Бәсекелесуші белсендіру функциялары

Бұл тақырыпта бәсекелесуші белсендіру функциялары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Бәсекелесуші белсендіру функцияның кірісі болып есептелген арақашықтықты b ығыстыру векторымен қосу нәтижесі болатын n^1 векторы табылады. Бәсекелесуші белсендіру функциясы n^1 векторы элементтерінің мәндерін талдайды және кірісінде максималды мәнге ие болатын бір нейрон-жеңімпаздан басқа барлық нейрондар үшін 0-ге тең нейрондардың шығысын құрады. Осылайша, a^1 қабатының шығу векторында нейрон-жеңімпазға сәйкес 1-ге тең жалғыз элемент, ал қалғандары 0-ге тең болады. Егер барлық ығыстырулар нөлдік болса, n^1 -дің максималды мәні 0-ден аса алмайды. n^1 -дің 0-дік мәні, қашан s кіру векторы нейрондардың біреуінің салмақ векторына тең болғанда ғана мүмкін болады. Егер ығыстырулар 0-ден бөлек болса, онда n^1 векторының элементтері үшін оң мәндер болуы мүмкін.

Осындай белсендіру функциясы (белсендіру сипаттамасы) келесі түрде жазылуы мүмкін:

$$f(s) = \begin{cases} 1, & \text{егер } i = i^*, i^* = \arg(\max_{n_i}^1) \\ 0, & \text{кері жағдайда} \end{cases}$$

Бұл белсендіру сипаттама жекелеген нейронға емес, қабатқа орнатылғандығын ескерейік. Сондықтан осындай белсендіру сипаттамасы *бәсекелесуші* атауына ие болған. Белсенді нейронның

нөмірі I^* кіру векторы аса жақын болып келетін топты (кластерді) анықтайды.

Бәсекелесуші белсендіру функциясының анықталу аймағы $D(f) = (-\infty, +\infty)$, ал мәндер аймағы $E(f) = (0, +1)$ болады.

Бәсекелесуші белсендіру функциялар ықтималды және өздігінен ұйымдастырылған нейрондық желілерді қалыптастыру үшін қолданылады.

II.2.14-тапсырмалар.

1. Бәсекелесуші белсендіру функцияларды анықтайтын формуланы жазыңыз.

2. Бәсекелесуші белсендіру функцияларының функционалды жұмысын сипаттаңыз.

3. Бәсекелесуші белсендіру функцияларының анықталу аймағын анықтаңыз.

4. Бәсекелесуші белсендіру функцияларының мәндер аймағын анықтаңыз.

5. I^* нейрон нөмірінің тағайындалуын жазыңыз.

6. n^1 нөлдік мәнге ие болғандағы оқиғаны сипаттаңыз.

7. Бәсекелесуші белсендіру функцияның анықталу аймағын көрсетіңіз.

II.2.14-сұрақтар.

1. Қандай вектор бәсекелесуші белсендіру функциясының кірісі болады?

2. n^1 векторының элементтерінің мәндерін талдаудан соң бәсекелесуші белсендіру функциясы нені құрады?

3. Қандай есептерді шешу үшін бәсекелесуші белсендіру функциясы қолданылады?

4. Барлық ығыстырулар нөл болса не болады?

5. n^1 векторының элементтері үшін мәндер қай уақытта оң болады?

6. s кіру векторы салмақ векторларының біріне тең болса не болады?

7. Қандай есепті шешу үшін бәсекелесуші белсендіру функциясы қолданылады?

II.2.14-тестілер.

1. Бәсекелесуші белсендіру функциясына кіріс болып табылатын вектор қалай жазылады?

- A) n^1 векторы;
- B) a^1 қабатының векторы;
- C) a қабатының векторы;
- D) b векторы;
- E) n векторы.

2. a^1 шығу қабатының векторы нейрон-жеңімпазға сәйкес келетін қандай элементтен тұрады?

- A) 1-ге тең жалғыз;
- B) тұрақты;
- C) кез-келген;
- D) бірнеше;
- E) тек оң.

3. Бәсекелесуші белсендіру функциясының мәндер аймағын анықтаңыз?

- A) $E(f) = (0, +1)$;
- B) $E(f) = (-1, +\infty)$;
- C) $E(f) = (-\infty, +1)$;
- D) $E(f) = (-\infty, +\infty)$;
- E) $E(f) = (-1, 0)$.

4. I^* нейрон нөмірінің тағайындалуы қалай анықталады?

- A) $i^* = \arg(\max n_i^1)$;
- B) $i^* = \arg(\max n)$;
- C) $i^* = \arg(\min n_i^1)$;
- D) $i^* = \arg(\min n)$;
- E) $i^* = \arg(\max n_i)$.

5. Егер барлық ығыстырулар нөл болса не болады?

- A) n^1 -дің максималды мәні 0-ден аса алмайды;
- B) n^1 -дің максималды мәні 0-ден асады;
- C) n^1 -дің минималды мәні 0-ден аса алмайды;

- D) n^1 -дің максималды мәні 1-ден асады;
- E) n^1 -дің максималды мәні 1-ден аса алмайды.

II.3. Нейрондардың моделдері

II.3.1. Персептрондар

Бұл тақырыпта персептрондар жайында мағлұматтар, персептрондардың мүмкіншіліктері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Бір нейрондық персептронның құрылымдық сұлбасы мен жұмыс істеу алгоритмі Мак-Каллок пен Питтс нейрондарымен ұқсас келеді.

Персептрон үш типті элементтерден тұрады: *S-элементі*, *A-элементі* және бір *R-элементі*.

S-элемент – сенсорлар немесе рецепторлар қабаты. Әр рецептор қозған немесе тыныштық күйде болады, және тек қозған күйде ғана бірлік сигнал келесі қабатқа ассоциативті элементтерге, содан соң әрі қарай жауап беруші элементтерге береді. Формалды түрде *A-элементі* *R-элементтері* сияқты шекті қосқыштардан тұрады, яғни олар жалғыз нейрондар болады. *A-элементтері* ассоциативті деп аталады, себебі әрбір мұндай элементке *S-элементтерінің* жиыны (ассоциациясы) сәйкес келеді. *A-элементі*, өзінің кірісіндегі *S-элементтерінен* сигналдары қандай да бір ω_0 шамасынан асып кеткенде, белсендіріледі. Сонымен, егер *S-элементтеріне* сәйкес келетін топ сенсорлық өрісте «Д» әрпі тәріздес орналасқан болса, онда *A-элементі* белсендіріледі.

Қоздырылған *A-элементтерінен* шыққан сигналдар, өз кезегінде, *R* қосқышқа беріледі, мұнда *i*-ші ассоциативті элементтен сигнал ω_i коэффициентімен беріледі. Бұл коэффициент *A–R* байланысының салмағы деп аталады.

Сондай-ақ, *A-элементтері* секілді *R-элементі* де салмақтарға көбейтілген кіру сигналдары қосындысының мәндерін есептейді. Егер сызықтық пішін w_0 межесінен асса, онда *R-элемент* және онымен қоса персептронда «1»-ді береді, әйтпесе шығыста «-1»

шығады. R -элементімен жүзеге асырылатын функцияны математикалық түрде келесідегідей жазуға болады:

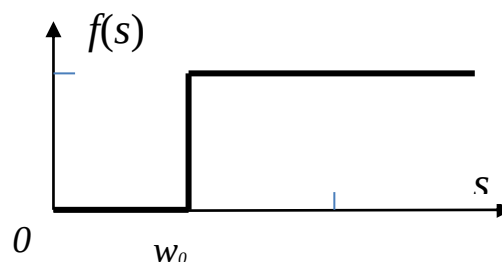
$$y = f(s) = f\left(\sum_{i=1}^n w_i x_i - w_0\right),$$

мұнда x_i – i -ші шығыс сигналы, w_i – i -ші синапстық байланыстардың салмағы, w_0 – ығысу (меже), s – салмақталған қосынды, y – шығу сигналы, f – межелік белсендіру функциясы, n – кірістер саны, i -ші синапстық байланыс қозушы (тежеуші) болып саналады, егер w_i синапс салмағының мәні оң (теріс) болса

Персептронды оқып үйрену синапстық A – R байланысының w_i салмағының өзгеруінен тұрады. S – A байланысының салмағы ($\{-1; 0; +1\}$ мәндерін қабылдай алатын) A -элементінің межелік мәндері ең басында кездейсоқ таңдап алынады да, содан соң соңына дейін өзгертілмейді. Персептронды оқып үйренген соң тану немесе жалпылау режимінде жұмыс істеуге дайын. Бұл режимде персептронға оған таныс емес объектілер ұсынылады және персептрон олардың қандай сыныпқа жататынын анықтауы қажет. Персептронның жұмысы келесіден тұрады: объектіні ұсынғанда қозған A -элементтері сигналды сәйкес w_i коэффициенттерінің қосындысына тең R -элементіне береді. Егер бұл қосынды оң болса, бұл объект бірінші сыныпқа жатады деген, ал егер ол теріс болса – екінші сыныпқа жатады деген шешім қабылданады. Ендеше, персептронның белсендіру функциясы келесі түрге ие:

1) S және A элементтерімен орындалатын белсендіру функциясы

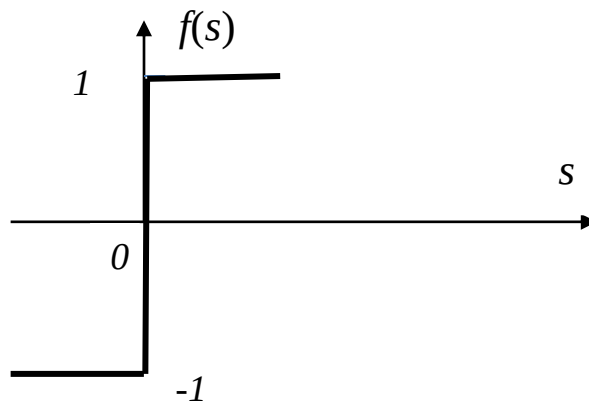
$$y = f(s) = \begin{cases} 0, & \text{егер } s < w_0 \\ 1, & \text{егер } s \geq w_0 \end{cases}$$



II.3.1.2-сурет. S және A элементтерін белсендіру функциясының графигі

2) Қарапайым R-элементімен орындалатын белсендіру функциясы

$$y = f(s) = \begin{cases} -1, & \text{егер } s < 0 \\ 1, & \text{егер } s \geq 0 \end{cases}$$



II.3.1.2-сурет. Қарапайым KR-элементін белсендіру функциясының графигі.

Енді персептронның авторы Ф. Розенблаттың сигналдар негізінде көрсеткен сипаттамасын келтіруге болады. Ол үшін ең алдымен межелі белсендіру функциясы бар жасанды нейронның дербес жағдайы болып табылатын персептронның құрамдас элементін анықтаймыз:

- қарапайым S-элемент (сенсорлы) энергияның қандай да бір түрлерінің (мысалы, жарық, дыбыс, қысым, жылу және т.с.с) әсерінен сигнал тудыратын сезгіш элемент болады. Егер кіру сигналы қандай да бір w_0 межелік шамадан асып кетсе, элемент шығысында +1-ді, кері жағдайда 0-ді аламыз.

- қарапайым A-элементі (ассоциативті) деп, оның кіру сигналдарының алгебралық қосындысы қандай да бір w_0 межелік шамадан асып кетсе, шығу сигналы +1-ді беретін (элемент белсендірілген), кері жағдайда шығысы нөлге тең болатын логикалық шешуші элементті айтады.

- қарапайым R -элементі (әрекет етуші, яғни жауап қайтарушы) деп аталатын және егер оның кіру сигналдарының сомасы қатаң оң болса, онда шығыс сигналы $+1$, ал егер оның кіру сигналдарының сомасы теріс болса, онда шығыс сигналы -1 беретін элементті айтады. Егер кіру сигналдарының сомасы нөлге тең болса, онда шығысы не нөлге тең, не анықталмаған деп саналады.

Егер кез-келген элементтің шығысында біз 1 -ді алатын болсақ, онда элемент *белсендірілген* немесе *қозған* күйде деп есептеледі.

Барлық қарастырылған элементтер *қарапайым* деп аталады, себебі олар *сатылы (секіртпелі)* функцияларды жүзеге асырады. Розенблаттың пайымдауынша қиын мәселерді шешу кезінде функцияның басқа түрлері, мысалы, сызықтық функциялар қолданылуы мүмкін.

Нәтижесінде Розенблатт келесі анықтамаларды енгізді:

- *Персептрон* S -, A -, R - элементтерінен тұратын W өзара әрекеттесудің матрицасымен (элементтері w_{ij} — синапстық байланыстардың салмағы болатын) желі белсенділігінің бұрынғы күйлерінің тізбегімен анықталатын желіден тұрады.

- *Тізбекті байланыстармен персептрон*, ол барлық байланыстары жақын орналасқан S -элементінен логикалық d ара қашықтығымен берілген элементтерден бастап, жақын орналасқан S -элементінен логикалық $d+1$ ара қашықтығымен берілген элементтерде аяқталатын жүйе.

- *Қарапайым персептрон* деп келесі бес шартты қанағаттандыратын кез-келген жүйені айтады:

- 1) жүйеде тек бір R -элементі болады (әрине, ол барлық A -элементтерімен байланысқан);

- 2) жүйе тек S -элементтерінен A -элементтеріне және A -элементтерінен R -элементтеріне жүретін тізбекті байланыстары бар персептроннан тұрады.

- 3) S -элементтерден A -элементтерге (S - A байланысы) дейінгі барлық байланыстың салмағы өзгермейді;

4) әр байланысты беру уақыты не нөлге тең, не бекітілген τ тұрақтысына тең;

5) S -, A -, R -элементтерінің барлық белсендіру функциялары келесі түрге ие болады:

$$u_i(t) = f(a_i(t))$$

мұнда $a_i(t) = u_i$ элементтерінің кірісіне бір уақытта келіп түсетін барлық сигналдардың алгебралық қосындысы.

- *Қарапайым персептрон* деп барлық элементтері қарапайым болып келетін қарапайым персептрондарды атайды. Бұл жағдайда оны белсендендіруші функциясы келесі түрге ие болады:

$$c_{ij}(t) = u_i(t - \tau) \cdot w_{ij}(t) .$$

Розенблаттың ұсынған және кейіннен нейрондық желілер теориясының аясында дамыған концепцияларды қосымша көрсетіледі:

1) *Айқыш байланыстары бар персептрондар* S -элементтен бірдей логикалық қашықтықта болатын бір типті элементтер (S , A немесе R элементтер) арасында байланыстар бар жүйе, мұнымен қоса барлық басқа байланыстар – тізбекті типтес болатын.

2) *Кері байланысты персептрон* – логикалық тұрғыдан аса қашық элементтен қашықтығы аз элементке дейін кем дегенде бір байланыс бар жүйе. Бұл желілерді рекуррентті деп атайды.

3) *S - A байланысты айнымалымен персептрондар* S -элементтен A -элементке дейінгі бекітілген байланысқа шектеуі алынған жүйе. S - A байланысты оптимизациялау жолымен персептронның сипаттамаларын біршама жақсартуға болатындығы дәлелденген.

Марвин Минскийге персептронның қасиеттерін талдау үшін персептрон теориясын предикаттар тілінде қайта баяндауға тура келді. Бұл әдістің мәнісі мынада:

- S -элементтерінен сигналдар жиынына X айнымалысы қойылды;

- әрбір A -элементіне жеке предикат деп аталған $\phi(X)$ предикаты сәйкес қойылды;

- әрбір R -элементіне жеке предикаттардан тәуелді болатын ψ предикаты сәйкес келді;

- персептрон деп ψ типті барлық предикаттарды есептей алатын құрылғы аталды.

Кіріс сигналдардың мәні екілік сандар болады, ал синапстар салмақтарының және ығысудың мәндері нақты сандар болады. Шығу сигналының мәні белсендіру функциясының түрімен анықталады және нақты немесе екілік сан бола алады.

Егер w_i синапс салмағының мәні оң (теріс) болса, онда i -ші синапстық байланыс қоздыраушы (тежеуші) болады.

Персептрондарды 5 тобы мына қызықты қасиеттерге ие болады:

1. *Диаметрі бойынша шектелген персептрондар*—жеке предикттармен танылатын әр X пішіні, диаметрі бойынша қандай да бір бекітілген шамадан аса алмайды.

2. *Шектелген ретті персептрондар* —әр жеке предикат X -тегі нүктелердің шектелген санынан тәуелді болады;

3. *Гамбы персептрондары* — әр жеке предикат сызықтық межелік функция, яғни мини-персептрон болуы керек;

4. *Кездейсоқ персептрондар* —жеке предикаттар кездейсоқ таңдап алынған бульдік функциядан тұратын, реті шектелген персептрондар;

5. *Шектелген персептрондар* —жеке предикаттар жиыны шексіз, ал a_i коэффициенттерінің мүмкін мәндер жиыны шектеулі.

Персептрондарды оқытудағы классикалық әдіс—қателерді дұрыстау әдісі. Персептрондардың ағымдағы реакциясы дұрыс болып қалғанға дейін байланыс салмағы өзгермейтін, мұғаліммен оқытылатын әдіс. Бұрыс реакция пайда болғанда салмақ бірлікке өзгереді, ал таңбасы (+/-) қателік таңбасын қарама-қарсы таңбаға өзгереді.

Мысалы, біз бірінші сыныптағы объектілерді ұсынған кезде персептронның шығысы оң (+1) болатындай, ал екінші сыныптағы объектілерді ұсынғанда теріс (-1) болатындай объектілердің екі

сыныпқа бөлетін персептронды оқытқымыз келді дейік. Ол үшін келесі алгоритмді орындаймыз:

1. Кездейсоқ түрде A -элементтері үшін межені таңдап аламыз және $S-A$ байланысын орнатамыз (олар ары қарай өзгермейді).

2. Бастапқы коэффициенттер w_I нөлге тең деп аламыз;

3. Оқыту іріктеуін ұсынамыз: олар тиісті болатын сыныбы көрсетілген объектілер (мысалы, дөңгелек немесе шаршылар).

- Персептронға бірінші сыныптағы объектіні көрсетеміз. Мұнда қандай да бір A -элементтері қозады. Қозған элементтерге сәйкес келетін w_I коэффициенттерін 1-ге ұлғайтамыз.

- Персептронға екінші сыныптағы объектіні көрсетеміз. Бұл көрсетілімде қозатын A -элементтерінің w_I коэффициенттерін 1-ге кемітеміз.

4. Үшінші қадамның екі бөлігін де барлық оқыту іріктеуі үшін орындаймыз. Оқыту нәтижесінде w_I байланыстар салмағының мәні қалыптасады.

Ф. Розенблат (Блоктың, Джозефтың, Кестеннің және онымен бірге жұмыс істеген тағы басқа зерттеушілердің қатысуымен) сипаттаған және дәлелдеген *персептронның жинақтылық теоремасы*, мұндай алгоритм бойынша оқытылатын қарапайым персептрон, салмақтық коэффициенттердің бастапқы күйінен және ынтаның пайда болу тізбегінен тәуелсіз түрде ақырғы уақыт аралығында шешімге жету мүмкіндігіне алып келетіндігін көрсетеді. Персептрондарды оқыту сұрақтары бойынша толық және нақты материалдар III.3.1-бөлімде қарастырылады.

III.3.1-тапсырмалар.

1. Персептрондардың құрылымдық элементтерін атап шығыңыз және олардың қасиеттері мен мақсаттарын сипаттаңыз.

2. Персептронның құрамдық элементтері арқылы кіру сигналының өту үдерісін сипаттаңыз.

3. Бірнейронды персептронның формуласын жазыңыз.

4. Персептронның белсендіру функциясының формуласын жазыңыз.

5. Предикаттар тілінде персептронға сипаттама беріңіз.

II.3.1-сұрақтар.

1. Персептронның таратушы қабаты не үшін арналған?

2. A -элементтері қай кезде белсенді болады?

3. Салмаққа көбейтілген кіру сигналдарының мәндерінің қосындысын қандай элемент есептейді?

4. Егер сызықтық форма w_0 межесінен асса, персептрон қандай мән береді?

5. Кіру сигналдарының, синапстар салмақтарының және ығысудың мәндері қандай типке жатады?

II.3.1-тестілер.

1. Персептронның қандай элементтері бар?

A) S -элементі, A -элементі және бір R -элементі;

B) S -элементі, A -элементі және екі R -элементтері;

C) S -элементі, бір A -элементі және R -элементі;

D) бір S -элементі, A -элементі және R -элементі;

E) S -элементі, A -элементі және R -элементі.

2. Қандай белсендіру функциясы S , A элементтерімен орындала?

A) $y = f(s) = \begin{cases} 0, & \text{егер } s < w \\ 1, & \text{егер } s \geq w \end{cases}$

B) $y = f(s) = \begin{cases} -1, & \text{егер } s < 0 \\ 1, & \text{егер } s \geq 0 \end{cases}$

C) $y = f(s) = \begin{cases} 0, & \text{егер } s > w_0 \\ 1, & \text{егер } s \leq w_0 \end{cases}$

D) $y = f(s) = \begin{cases} 0, & \text{егер } s < w_0 \\ 1, & \text{егер } s = w_0 \end{cases}$

E) $y = f(s) = \begin{cases} 1, & \text{егер } s < w_0 \\ 0, & \text{егер } s \geq w_0 \end{cases}$

4. Қандай белсендіру функциясы R -элементімен орындалады?

A) $y = f(s) = \begin{cases} -1, & \text{егер } s < 0 \\ 1, & \text{егер } s \geq 0 \end{cases}$

B) $y = f(s) = \begin{cases} 0, & \text{егер } s < w_0 \\ 1, & \text{егер } s \geq w_0 \end{cases}$

C) $y = f(s) = \begin{cases} 1, & \text{егер } s < w_0 \\ 0, & \text{егер } s \geq w_0 \end{cases}$

D) $y = f(s) = \begin{cases} 0, & \text{егер } s < w_0 \\ 1, & \text{егер } s = w_0 \end{cases}$

E) $y = f(s) = \begin{cases} 1, & \text{егер } s < w_0 \\ 0, & \text{егер } s \geq w_0 \end{cases}$

4. Айқыш байланысты персептрондардың анықтамасы қандай?

A) S -элементтен бірдей логикалық қашықтықта болатын, мұнымен қоса барлық басқа байланыстар – тізбекті типтес болатын, бір типті элементтер арасында байланыстар бар жүйе;

B) логикалық тұрғыдан аса қашық элементтен, қашықтығы аз элементке дейін кем дегенде бір байланыс бар жүйе. Замануи терминологияға сәйкес бұл желілерде рекуррентті деп аталады;

C) S -элементтен A -элементке дейінгі нақтыланған байланысқа шектеу қойылмаған жүйе;

D) жеке предикттармен танылатын әр X пішіні, диаметрі бойынша қандай да бір бекітілген шамадан аса алмайды;

E) әр жеке предикат X -тегі нүктелердің шектелген санынан тәуелді болады.

5. Кері байланысты персептронның анықтамасы қандай?

A) логикалық тұрғыдан аса қашық элементтен қашықтығы аз элементке дейін кем дегенде бір байланыс бар жүйе және замануи терминологияға сәйкес бұл желілерде рекуррентті деп аталады;

B) S -элементтен бірдей логикалық қашықтықта болатын, мұнымен қоса барлық басқа байланыстар – тізбекті типтес болатын, бір типті элементтер (S , A немесе R элементтер) арасында байланыстар бар жүйе;

С) S-элементтен A-элементке дейінгі нақтыланған байланысқа шектеу қойылмаған жүйе. S-A байланысты оптимизациялау жолымен персептронның сипаттамаларын біршама жақсартуға болатындығы дәлелденген;

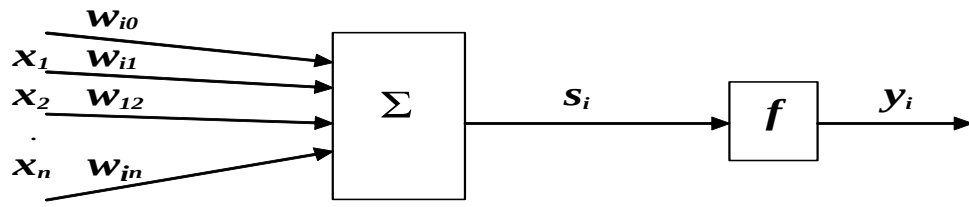
D) жеке предикттармен танылатын әр X пішіні, диаметрі бойынша қандай да бір бекітілген шамадан аса алмайды;

E) әр жеке предикат X -тегі нүктелердің шектелген санынан тәуелді болады.

II.3.2. Сигмоидалық нейрондар

Бұл тақырыпта сигмоидалық нейрондардың мүмкіндіктері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

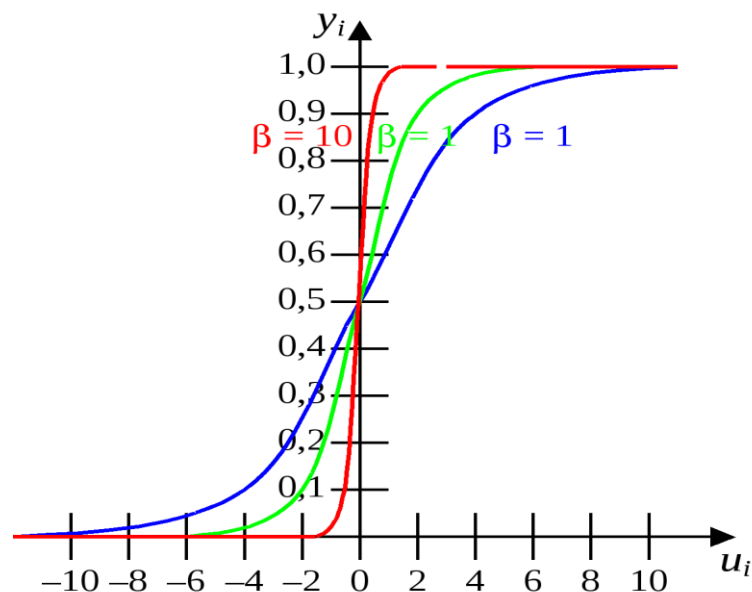
Сигмоидалық нейрон Мак-Каллок-Питтс нейронның моделіндей құрылымға ие, алайда оның сигмоидалды униполярлы немесе биполярлы функция түрінде өрнектеле алуы мүмкін дифференциалды белсендіру функциясы бар. Сигмоидалық нейронның құрылымы II.3.2.1-суретте көрсетілген.



II.3.2.1-сурет. Сигмоидалық нейронның құрылымы
Униполярлы белсендіру функциясы келесідегідей болады:

$$y_i = f(s_i) = \frac{1}{1 + e^{-\beta s_i}} \quad (\text{II.3.2.1})$$

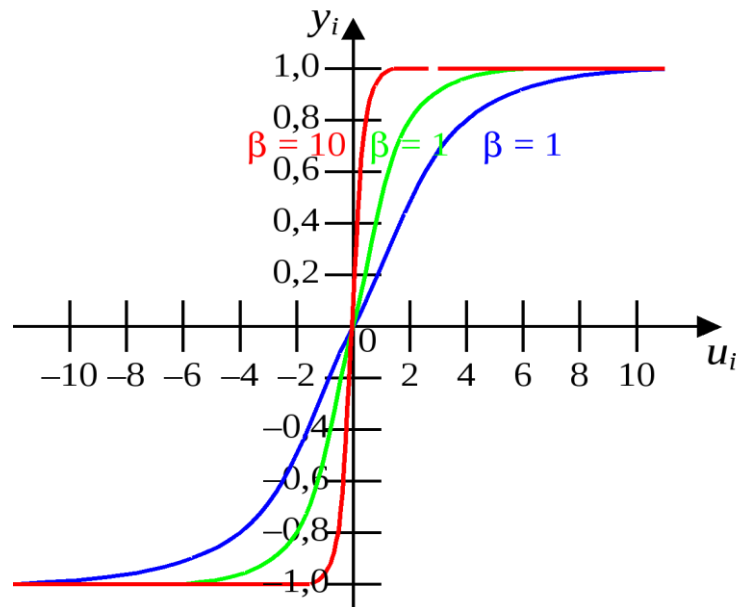
Униполярлы функцияның графигі II.3.2.2-суретте көрсетілген.



II.3.2.2-сурет. Униполярлы функцияның графигі
Биполярлы белсендіру функциясы келесідегідей болады:

$$y_i = f(s_i) = \tanh(\beta s_i) = \frac{e^{\beta s_i} - e^{-\beta s_i}}{e^{\beta s_i} + e^{-\beta s_i}} \quad (\text{II.3.2.2})$$

Биполярлы функцияның графигі II.3.2.3-суретте көрсетілген.



II.3.2.3- сурет. Биполярлы функцияның графигі

β параметрінің мәні тұтынушымен таңдалады және белсендіру функциясының формасына әсер етеді. $\beta \rightarrow \infty$ болғанда, сигмоидалды функция персентронның біркелкі белсендіру функциясының сатылы типті функциясына айналады. Тәжірибеде әдетте жеңілдету үшін $\beta = 1$ мәні қолданылады.

Сигмоидалды функцияның маңызды қасиеті оның дифференциалдандылығы болады. Униполярлы функция үшін туынды келесі формула бойынша анықталады:

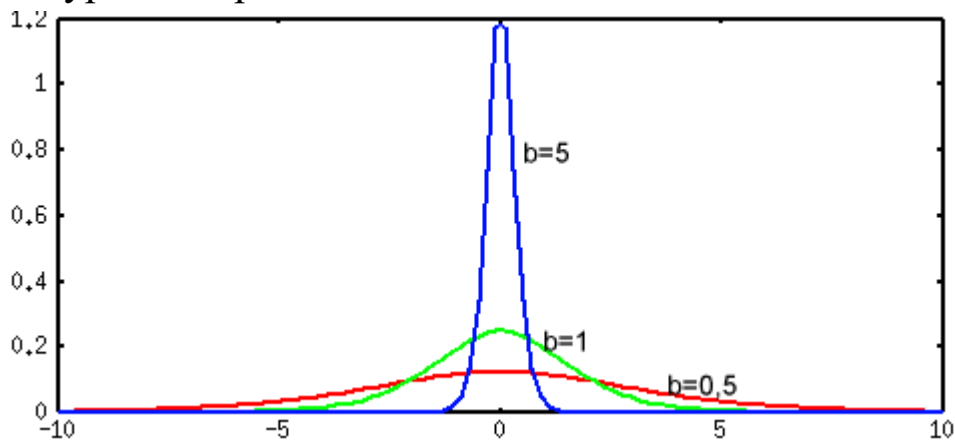
$$\frac{dy_i}{ds_i} = \beta \cdot f(s_i) \cdot (1 - f(s_i)) = \beta \left(\frac{1}{1 + e^{-\beta s_i}} \right) \cdot \left(1 - \frac{1}{1 + e^{-\beta s_i}} \right) = \frac{\beta e^{-\beta s_i}}{(1 + e^{-\beta s_i})^2},$$

ал биполярлы функция үшін:

$$\frac{dy_i}{ds_i} = \beta \cdot (1 - f^2(s_i)) = \beta \cdot \left(1 - \left(\frac{e^{\beta s_i} - e^{-\beta s_i}}{e^{\beta s_i} + e^{-\beta s_i}} \right)^2 \right) = \frac{4\beta}{e^{2\beta s_i} + e^{-2\beta s_i} + 2}$$

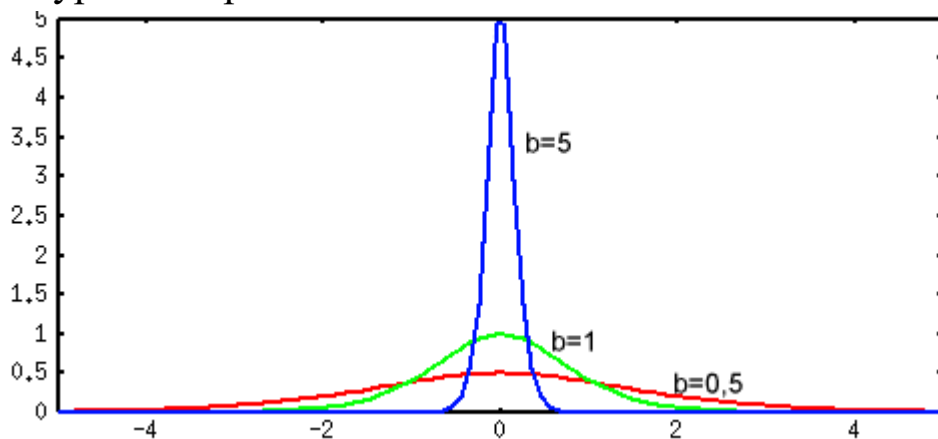
Сигмоидалық функциялардың туындыларының графиктері қоңырау тәріздес түрге ие.

Униполярлы сигмоидалық функция туындысының графигі
 II.3.2.4- суретте көрсетілген.



II.3.2.4-сурет. Униполярлы сигмоидалық функция туындысының графигі

Биполярды сигмоидалды функция туындысының графигі
 II.3.2.5-суретте көрсетілген.



II.3.2.5- сурет. Биполярды сигмоидалды функция туындысының графигі

Сигмоидалды нейронды оқыту үшін оқытудың $E(w_i)$ мақсатты функциясын азайтуға мүмкіндік беретін градиентті әдістерді қолдану арқылы «мұғаліммен» парадигмасы (стратегиясы) қолданылады. Мұнда персептронға қарағанда, мақсатты функцияның минимумын іздеу үшін салмақтық коэффициенттердің w_{ij} мақсатты өзгерісі теріс градиент бағытында жүретін бірінші кезектегі іздеу оңтайлылығы әдісі қолданылады:

$$E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r \left(\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k \right)^2 \rightarrow \min_w \quad (\text{II.3.2.3})$$

мұнда $y_i^k = f(s_i^k) = f\left(\sum_{i=0}^m w_{ij}^k x_i^k\right)$ сигмоидалды функция болады; x – кіру векторы $x = [x_0, x_1, \dots, x_n]^T$ поляризация бар кезде $x_0 = 1$ мәнімен және ол жоқ кезде $x_0 = 0$ мәнімен, ал d_i – I -нші нейронның шығысындағы оған сәйкес күтілетін мән.

Үздіксіз белсендіру функциясын қолдану оқыту барысында градиентті әдістерді пайдалануға мүмкіндік береді. $w = [w_0, w_1, \dots, w_n]^T$ салмақ векторын анықтау мақсатты функцияның теріс градиенті бағытында жүретін ең жылдам түсу әдісін іске асыру жеңіл болып келеді. Егер бұл функция (II.3.2.3) өрнекпен анықталса, градиенттің j -нші құрамдасы келесі түрде:

$$\nabla_j E(w_i) = \frac{dE(w_i)}{dw_{ij}} = \sum_{k=1}^r (y_i^k - d_i^k) \frac{dy_i^k}{dw_{ij}} = \sum_{k=1}^r (y_i^k - d_i^k) \frac{df(s_i^k)}{ds_i^k} x_j^k,$$

мұнда ең жылдам түсу әдісі бойынша w_{ij} салмақты анықтау мақсатты функцияның теріс градиенті бағытында жүргізіледі.

$$\delta_i^k = (y_i^k - d_i^k) \frac{df(s_i^k)}{ds_i^k} \text{ өрнектейміз, одан шығады: } \frac{dE(w_i)}{dw_{ij}} = \sum_{k=1}^r \delta_i^k \cdot x_j^k.$$

Сигмоидалық нейронды дискретті әдіспен – әрбір эталонды жұп $\langle x^k, d_i^k \rangle$ үшін кіріс салмақтарын анықтау циклінің сериясымен оқыту да мүмкін (персептрон ережесі сияқты). Әрбір циклден соң салмақтарды түзету келесі функция бойынша орындалады:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k,$$

мұнда η – мәні $(0,1)$ интервалдан эмпирикалық тұрғыда таңдалынатын оқыту коэффициенті.

Салмақтардың өзгеруін келесі айырымдық тендеудің шешімімен де орындауға болады:

$$\frac{dw_{ij}(t)}{dt} = -\mu \delta_i x_j,$$

мұнда $\mu \in (0, 1)$ – оқыту коэффициенті.

1-ші реттегі іздеу оңтайландырудың барлық әдістері – глобалды экстремумға жетуге кепілдік бермейтін локалды іздеудің әдістері екенін ескерген жөн. Бұл кемшіліктерден өту үшін моментпен оқытудың әдісі ұсынылды, онда салмақтарды түзету келесідегідей орындалады:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \cdot \delta_i^k(t) \cdot x_j^k + \alpha \cdot (w_{ij}(t) - w_{ij}(t)). \quad (\text{II.3.2.4})$$

Формулада соңғы қосылғыш момент деп аталады және алдыңғы циклдегі салмақтардың нақты өзгерісін сипаттайды (α мәні $(0, 1)$ диапазонында таңдалады). Локалды минимум нүктесінде жақындағанда (мұнда $\delta_i^k(t) \cdot x_j^k$ градиентті құрамдас нөлге ұмтылғанда) момент құрамдасы локалды минимум аймағынан перспективалық аймаққа іздеуге шығарады.

Нейронды оқытудың градиентті әдісін пайдалану локалды экстремумға ғана жетуге кепілдік беретінін ескеру қажет, ол полимодалды $E(w)$ үшін глобалдық минимумнан жеткілікті алыс болуы мүмкін. Оның аймағынан ең жылдам түсу қарапайым алгоритмін пайдаланғанда шығу мүмкін емес, сондықтан практикада шашумен немесе моментпен оқытуды пайдаланады.

Мұнда салмақты айқындау үдерісі функцияның градиенті туралы ақпаратпен ғана емес, салмақ өзгерісінің нақты трендімен де анықталады:

$$\Delta w_{ij}(t+1) = -\eta \delta_i x_j + \alpha \Delta w_{ij}(t) \quad (\text{II.3.2.4})$$

мұнда $\alpha \in (0, 1)$ – момент коэффициенті, немесе оңтайландырудың стохастикалық әдісін пайдалану.

Моменттің салмақты таңдауға ықпалы α мәні өскен сайын арта береді, сондықтан α -ны момент оқу барысында үстемдік етпейтін етіп таңдау қажет, себебі ол оқыту алгоритмінің орнықсыздығына (шашырауына) алып келеді.

II.3.2-тапсырмалар.

1. Униполярлы белсендіру функциясының формуласын жазыңыз.
2. Биполярлы белсендіру функциясының формуласын жазыңыз.
3. Униполярлы белсендіру функциясының туындысын есептеудің формуласын жазыңыз.
4. Биполярлы белсендіру функциясының туындысын есептеудің формуласын жазыңыз.
5. Градиенттің j -нші құрамдасына арналған формуланы жазыңыз.
6. Сигмоидалды нейронды дискретті әдіспен оқытудың формуласын жазыңыз.
7. Сигмоидалды нейронды үздіксіз әдіспен оқытудың формуласын жазыңыз.

II.3.2-сұрақтар.

1. Униполярлы белсендіру функциясының графигі қандай түрде?
2. Туынды униполярды белсендіру функциясының графигі қандай?
3. Биполярлы белсендіру функциясының графигі қандай түрде?
4. Туынды биполярды белсендіру функциясының графигі қандай?
5. Сигмоидалды нейронды оқыту үшін қандай парадигма қолданылады?
6. Нейронды оқытудың градиентті әдісі нені кепілдендіреді?
7. Салмақты таңдау неге байланысты артады?

II.3.2-сұрақтары.

1. Униполярлы белсендіру функциясы қандай өрнекпен көрсетіледі?

$$A) y_i = f(s_i) = \frac{1}{1 + e^{-\beta s_i}};$$

$$B) y_i = f(s_i) = \frac{1}{1 + e};$$

$$C) y_i = f(s_i) = \frac{1}{1 + e^{s_i}};$$

$$D) y_i = f(s_i) = \tanh(\beta s_i) = \frac{e^{\beta s_i} - e^{-\beta s_i}}{e^{\beta s_i} + e^{-\beta s_i}};$$

$$E) y_i^k = f(s_i^k) = f\left(\sum_{i=0}^m w_{ij}^k x_i^k\right).$$

2. Биполярлы белсендіру функциясы қандай өрнекпен көрсетіледі?

$$A) y_i = f(s_i) = \tanh(\beta s_i) = \frac{e^{\beta s_i} - e^{-\beta s_i}}{e^{\beta s_i} + e^{-\beta s_i}};$$

$$B) y_i = f(s_i) = \frac{1}{1 + e};$$

$$C) y_i = f(s_i) = \frac{1}{1 + e^{s_i}};$$

$$D) y_i = f(s_i) = \frac{1}{1 + e^{-\beta s_i}};$$

$$E) y_i^k = f(s_i^k) = f\left(\sum_{i=0}^m w_{ij}^k x_i^k\right).$$

3. Сигмоидалды нейронды дискретті әдіспен оқытудың формуласы қандай болады?

$$A) w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k;$$

$$B) w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k(t);$$

- C) $w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k x_j^k$;
- D) $w_{ij}(t) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k$;
- E) $w_{ij}(t-1) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k$.

4. Сигмоидалды нейронды үздіксіз әдіспен оқытудың формуласы қандай болады?

- A) $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k)^2 \rightarrow \min_w$;
- B) $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k)^2 \rightarrow \max_w$;
- C) $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k)^2$;
- D) $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k)^2 \rightarrow \min_w$;
- E) $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k)^2 \rightarrow \min_w$.

5. Сигмоидалық нейрон қандай формуламен өрнектеле алады?

- A) сигмоидалды униполярлы немесе биполярлы функция түрінде;
- B) тек сигмоидалды униполярлы функция түрінде;
- C) тек биполярлы функция түрінде;
- D) үздіксіз немесе биполярлы функция түрінде;
- E) үздіксіз немесе дискретті функция түрінде.

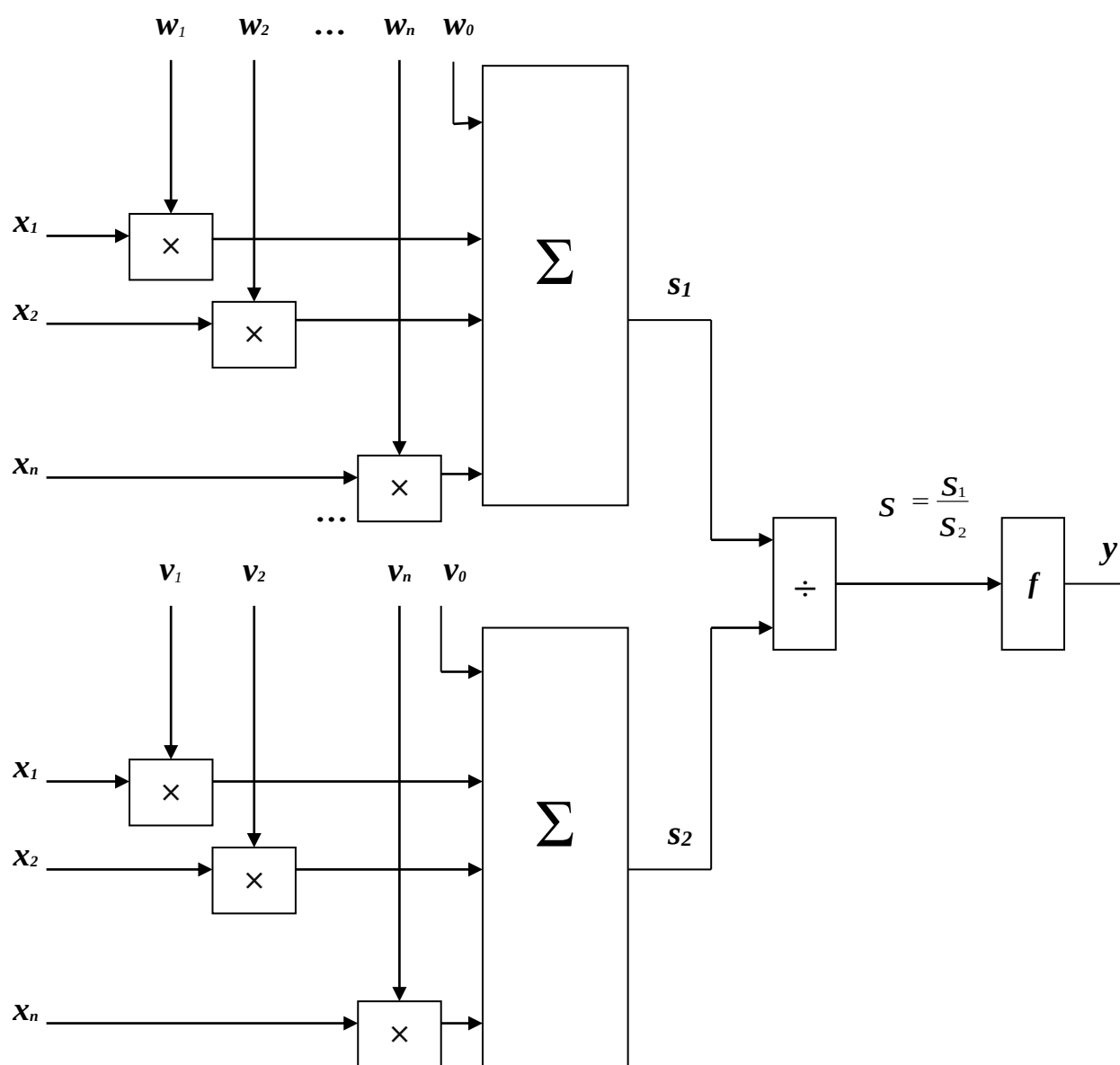
II.3.3. Паде-нейрондар

Бұл тақырыпта паде-нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Паде-нейрондар екі қосқыштан, салмақталған қосындылардың бөліндісін есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады. Паде-нейрон өзінің атауын рационалды функцияның аппроксимациясының паде-әдісінен алған.

Паде-нейронның құрылымы II.3.3-суретте көрсетілген.



II.3.3-сурет. Паде-нейрондардың құрылымдық сұлбасы

Практикады мұндай әдіс көп қолданылмайды, себебі ол айтарлықтай қосымша мүмкіндіктер берместен, нейрожүйелік модел құру үдерісін күрделендіреді. Нейрондағы белсендіру деңгейі мына формуламен анықталады:

$$s = \frac{\sum_{i=0}^n w_i x_i}{\sum_{i=0}^n v_i x_i}$$

Осы жағдайда квадраттық қатені минималдау келесі өрнекпен көрсетілген.

$$E(w) = \sum_{i=1}^n \left(a \left(\frac{w_i x_i}{v_i x_i} \right) - y_i \right)^2 \rightarrow \min_w$$

Паде-нейронның математикалық моделі келесі түрде беріледі:

$$y = f(s) = f\left(\frac{s_1}{s_2}\right) = f\left(\frac{w_0 + \sum_{i=1}^n w_i x_i}{v_0 + \sum_{i=1}^n v_i x_i}\right),$$

мұнда w_i – бірінші қосқыштың салмақтары, w_0 – бірінші қосқыштың бастапқы жағдайы, v_i – екінші қосқыштың салмақтары, v_0 – екінші қосқыштың бастапқы жағдайы, яғни бапталатын салмақтың саны $2n$ болады.

Паде-нейронда бапталатын салмақтың саны классикалық нейронға қарағанда екі есе көп, яғни $2n$ болады.

Паде-нейрон сызықтық функция жеткіліксіз болған жағдайда, «адалайн» типті нейрондардың жиынтығын жалпылау ретінде қолданылуы мүмкін.

II.3.3-тапсырмалар.

1. Паде-нейронды белсендірілу деңгейін есептеу формуласын жазыңыз.

2. Паде-нейронның квадраттық қатесін минималдау формуласын жазыңыз.

3. Паде-нейронның математикалық моделін жазыңыз.

II.3.3-сұрақтар.

1. Паде-нейрон немен ерекшеленеді?
2. Паде-нейронның қанша бапталатын салмағы болады?
3. Паде-нейрон не үшін және қандай жағдайда қолданылады?

II.3.3-тестілер

1. Паде-нейрондарының құрамына неден тұрады?

A) екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады;

B) салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады;

C) екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен тұрады;

D) бір қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады;

E) екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бірнеше блогынан тұрады.

2. Паде-нейронды белсендірілу деңгейін қалай есептейді?

$$A) s = \frac{\sum_{i=0}^n w_i x_i}{\sum_{i=0}^n v_i x_i} ;$$

$$B) s = \frac{\sum_{i=0}^n w_i}{\sum_{i=0}^n v_i} ;$$

$$C) s = \frac{\sum_{i=0}^n x_i}{\sum_{i=0}^n v_i x_i} ;$$

$$D) s = \frac{\sum_{i=0}^n x_i}{\sum_{i=0}^n v_i x_i} ;$$

$$E) s = \frac{1}{\sum_{i=0}^n v_i x_i} .$$

3. Паде-нейронда бапталатын салмақтың саны классикалық нейронға қарағанда неше есеге көп болады?

- A) 2 есе;
- B) 3 есе;
- C) 4 есе;
- D) 10 есе;
- E) 5 есе.

4. Паде-нейрон не үшін және қандай жағдайда қолданылады?

A) Паде-нейрон сызықтық функция жеткіліксіз болған жағдайда, «адалайн» типті нейрондардың жиынтығын жалпылау ретінде қолданылуы мүмкін;

B) Паде-нейрон сызықтық функция жеткіліксіз болған жағдайда, жалпылау ретінде қолданылуы мүмкін;

C) Паде-нейрон нейрондардың жиынтығын жалпылау ретінде қолданылуы мүмкін;

D) Паде-нейрон белсендіру функция жеткіліксіз болған жағдайда, «адалайн» типті нейрондардың жиынтығын жалпылау ретінде қолданылуы мүмкін;

E) Паде-нейрон үздіксіз функция жеткіліксіз болған жағдайда, «адалайн» типті нейрондардың жиынтығын жалпылау ретінде қолданылуы мүмкін.

5. Паде-нейронның қанша бапталатын салмағы болады?

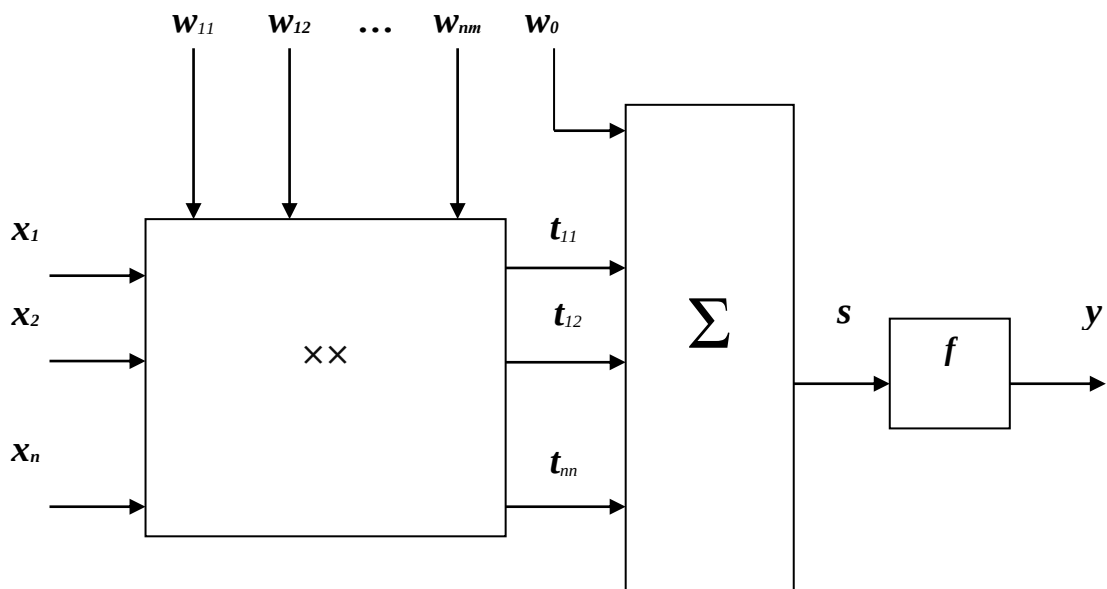
- A) $2n$;
- B) $3n$;
- C) n ;
- D) $5n$;
- E) $10n$.

II.3.4. Квадраттық қосқышы бар нейрондар

Бұл тақырыпта квадраттық қосқышы бар нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Квадраттық қосқышы бар нейрондар жасанды нейронның стандартты емес моделіне жатады. Нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді.

Квадраттық қосқышы бар нейронның құрылымы II.3.4-суретте көрсетілген.



II.3.4-сурет. Квадраттық қосқышы бар нейронның құрылымы

Квадраттық қосқышы бар нейронда ең алдымен кіріс сигналдардың жұп комбинациясының квадраттық формалары синапстық байланыстардың салмағына көбейтіледі, яғни $(i = 1, 2, \dots, n; j = 1, 2, \dots, m)$, сосын салмақталған квадраттық формалар z_{ij} және бастапқы жағдай сигналының мәні w_0 қосылады, ол келесі өрнекпен көрсетіледі:

$$s = w_0 + \sum_{i=1}^n v_i x_i + \sum_{i=1}^n \sum_{j=1}^m w_{ij} x_i x_j .$$

Квадраттық қосқышы бар нейронда шығыс блогына кіру сигналының салмақталған квадраттық формасы беріледі және бұл салмақталған сома f функциясына сәйкес түрленеді және шығу сигналының $y = f(s)$ мәні анықталады.

Квадраттық қосқышы бар нейронның математикалық моделін келесі түрде көрсетуге болады:

$$y = f(s) = f\left(w_0 + \sum_{i=1}^n v_i x_i + \sum_{i=1}^n \sum_{j=1}^m w_{ij} x_i x_j\right)$$

Квадраттық қосқышы бар нейронда қалыптасатын салмақтың саны $n(n+1)/2$ -ге тең.

Көпөлшемді қалыпты тарату үшін квадраттық қосқышы бар нейрон ең үздік сипаттауыш болады. Квадраттық таратушы жазықтық ең аз қателік ықтималдылығын береді:

– егер $f(s) > 0$, объект бірінші сыныпқа жатады;

– егер $f(s) \leq 0$, объект екінші сыныпқа жатады ($Q(x)$

коэффициентін дұрыс таңдаған жағдайда).

Квадраттық қателік былай анықталады:

$$E(w_{ij}, v_i, w_0) = \frac{e^2}{2} = \frac{(d - f(s))^2}{2}$$

Квадраттық қосқыштың коэффициенттері квадраттық қателердің анықтамаларына байланысты келесі формулалармен анықталады:

$$w_{ij}(t+1) = w_{ij}(t) + 2\alpha \cdot e \cdot x_i \cdot x_j,$$

$$v_i(t+1) = v_i(t) + \alpha \cdot e \cdot x_i,$$

$$w_0(t+1) = w_0(t) + \alpha \cdot e,$$

мұнда $\alpha \in (0, 1)$ – оқыту коэффициенті.

Квадраттық қосқышы бар нейрон көпөлшемді нормалды таратулар үшін жасанды нейрондар арасындағы ең үздік сипаттауыш болады. Мұндай сипаттауыштың кемшілігі – қалыптанатын параметрлердің көп болуы.

Жасанды нейронның басқа үлгілерін де құрастыруға болады. Алайда кез келген болжамда өзара әрекеттесетін нейрондар желісінің моделі мидың шынайы мүмкіндіктерімен салыстыруға болатын қасиеттерге ие және бай болып келеді.

Қарапайым элементтерден тұратын квадраттық қосқышы бар желілерді құру барысында байланысты жүргізу мен туындыны есептеуді ұйымдастыру бойынша тұтынушыға көп жұмыс жүктеледі.

II.3.4-тапсырмалар.

1. Квадраттық қосқышы бар нейронның құрамын көрсетіп, құрылымын салыңыз.
2. Квадраттық қосқышы бар нейронның математикалық моделін жазыңыз.
3. Квадраттық қосқышы бар нейрон үшін квадраттық қатені есептеу формуласын жазыңыз.
4. Квадраттық қосқышы бар нейрон үшін қосқыш коэффициентін айқындау әдісін сипаттаңыз.
5. Квадраттық қосқышы бар нейронның артықшылықтары мен кемшіліктерін көрсетіңіз.

II.3.4-сұрақтар.

1. Квадраттық қосқышы бар нейронда қалыптанатын салмақтардың саны қанша?
2. Квадраттық қосқышы бар нейрон ең үздік сипаттауыш болады ма?
3. Квадраттық қосқышы бар нейрон негізінде сипаттауыш коэффициенті қандай формула бойынша анықталады?
4. Көпөлшемді қалыпты тарату үшін ең үздік сипаттауыш болып не табылады?

5. Квадраттық қосқышы бар нейрон – сипаттауыштың кемшілігі неде?

II.3.4-тестілер.

1. Квадраттық қосқышы бар нейрондар қандай моделге жатады?

- A) Жасанды нейронның стандартты емес моделіне;
- B) Жасанды нейронның стандартты моделіне;
- C) Жасанды нейронның кері таратылған моделіне;
- D) Жасанды нейронның үздіксіз моделіне;
- E) Жасанды нейронның кез-келген моделіне.

2. Квадраттық қосқышы бар нейронда қалыптанатын салмақтардың саны қанша?

- A) $n(n+1)/2$;
- B) $n(n+1)$;
- C) $n/2$;
- D) $2n$;
- E) $n(n+2)/2$.

3. Квадраттық қосқышы бар нейрон негізінде сипаттауыш коэффициентінің формуласы қайсысы?

- A) $w_{ij}(t+1) = w_{ij}(t) + 2\alpha \cdot e \cdot x_i \cdot x_j$;
- B) $w_{ij}(t+1) = w_{ij}(t) + 2\alpha \cdot e$;
- C) $w_{ij}(t+1) = 2\alpha \cdot e \cdot x_i \cdot x_j$;
- D) $w_{ij}(t+1) = w_{ij}(t) + 2\alpha \cdot e \cdot x_i \cdot x_j + 1$;
- E) $w_{ij}(t+1) = w_{ij}(t) - 2\alpha \cdot e \cdot x_i \cdot x_j$.

4. Квадраттық қосқышы бар нейрон сипаттауыштың кемшілігі қандай?

- A) қалыптанатын параметрлердің көп болуы;
- B) ең жылдам түсім әдісінің қолданылуы;
- C) бір ғана қалыптанатын параметрінің болуы;
- D) қалыптанатын параметрлердің жоқ болуы;
- E) Жасанды нейронның стандартты емес моделге жатуы.

5. Квадраттық қосқышы бар нейрон үшін квадраттық қатені есептеу формуласы қандай?

A) $E(w_{ij}, v_i, w_0) = \frac{e^2}{2} = \frac{(d - f(s))^2}{2}$;

B) $E(w_{ij}, v_i, w_0) = \frac{e^2}{2} = \frac{(d)^2}{2}$;

C) $w_{ij}(t+1) = w_{ij}(t) + 2\alpha \cdot e \cdot x_i \cdot x_j$;

D) $E(w_{ij}, v_i, w_0) = \frac{e^2}{2} = \frac{(d + f(s))^2}{2}$;

E) $v_i(t+1) = v_i(t) + \alpha \cdot e \cdot x_i$.

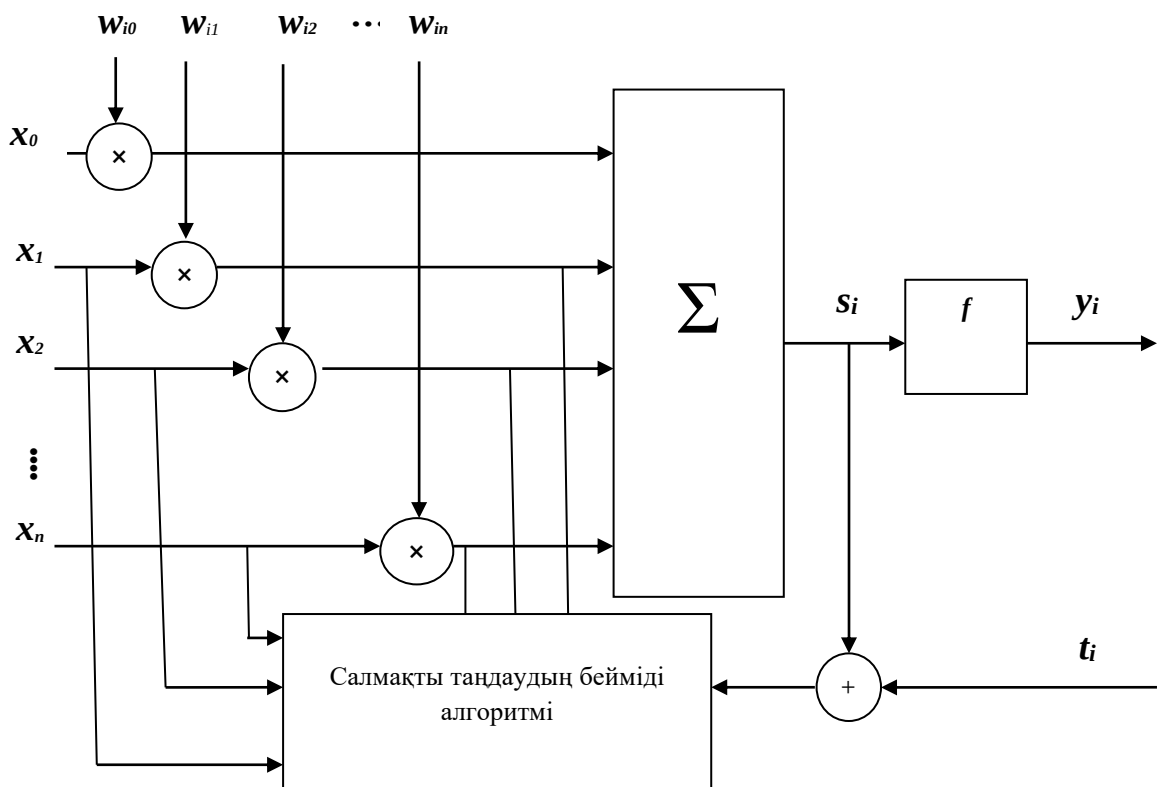
II.3.5. Бейімделген сызықтық нейрондар

Бұл тақырыпта бейімделген сызықтық нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Бейімделген сызықтық нейрондарды—ADaptive Linear Neuron (немесе «адалайн»—«adalline» типті нейрон) Б.Уидроу (B.Widrow) ұсынған болатын. Олар жасанды нейронның стандартты емес моделіне жатады. Сигналдардың салмақталған қосындысы әдісі бойынша «адалайн» типті нейрондар Мак-Каллок-Питтс моделіне ұқсас, алайда белсендіру функциясы ретінде сигнатуралық функцияны пайдаланады, яғни бейімделген сызықтық нейронда белсендіру функциясы «сигнум»—«sgn» типті функция:

$$y_i = f(s_i) = \begin{cases} 1, & \text{егер } s_j > 0 \\ -1, & \text{егер } s_j \leq 0 \end{cases}$$

Бейімделген сызықтық нейронның құрылымы II.3.5-суретте көрсетілген.



II.3.5-сурет. Бейімделген сызықтық нейрон құрылымы

Бұл нейронда салмақтың бейімді таңдауы квадраттық қатені азайту үдерісінде орындалады:

$$E(w) = \frac{1}{2} (d_i - \sum_{j=1}^n w_{ij} x_j)^2 \rightarrow \min_w, (12)$$

мұнда d_i – нейронның күтілетін шығу сигналы.

Уидроу салмақтарды таңдау бейімделген алгоритмінде мақсаттық функция үздіксіз болады, сондықтан оны азайту үшін ең жылдам түсім әдісі қолданылады. Сонымен қатар, салмақ келесі формула бойынша нақтыланады:

$$w_{ij}(t+1) = w_{ij}(t) + \eta e_i x_i$$

мұнда η – оқыту қарқындылығы, $e_i = d_i - \sum_{j=1}^n w_{ij} x_j$.

II.3.5-тапсырмалар.

1. Бейімделген сызықтық нейронның белсендіру функциясының формуласын жазыңыз.
2. Бейімделген сызықтық нейронның құрылымын салыңыз.
3. Бейімделген сызықтық нейрон үшін салмақты таңдау үшін квадраттық қатені азайту үдерісінің формуласын жазыңыз.
4. Бейімделген сызықтық нейронға арналған мақсатты функцияның формуласын жазыңыз.
5. Бейімделген сызықтық нейронға арналған салмақты айқындау формуласын жазыңыз.

II.3.5-сұрақтар.

1. Бейімделген сызықтық нейронды кім ұсынды?
2. Сигналдардың салмақталған қосындысы әдісі бойынша «адалайн» типті нейрон қандай моделге ұқсас?
3. Бейімделген сызықтық нейрон қандай белсендіру функциясын пайдаланады?
4. Бейімделген сызықтық нейронда белсендіру функциясы дифференциалдана ма?

5. «Салмақты таңдаудың бейімделген алгоритмі» құрылымдық элементі қандай рөл атқарады?

II.3.5.-тестілер

1. Бейімделген сызықтық нейрондарды кім ұсынды?

- A) Б.Уидроу;
- B) Ф. Розенблатт;
- C) У. Маккалох;
- D) У. Питтс;
- E) Е. Хинтон.

2. Сигналдардың салмақталған қосындысы әдісі бойынша «адалайн» типті нейрон қандай моделге ұқсас?

- A) Мак-Каллок-Питтс моделіне;
- B) Френсиса Розенблатт моделіне;
- C) Хопфилд моделіне;
- D) Б.Уидроу моделіне;
- E) У. Маккалох моделіне.

3. Бейімделген сызықтық нейронда белсендіру функциясы қандай болады?

A) $y_i = f(s_i) = \begin{cases} 1, & \text{егер } s_j > 0 \\ -1, & \text{егер } s_j \leq 0 \end{cases};$

B) $y_i = f(s_i) = \begin{cases} 1, & \text{егер } s_j > 0 \\ -1, & \text{егер } s_j < 0 \end{cases};$

C) $y_i = f(s_i) = \begin{cases} 1, & \text{егер } s_j > 0 \\ -1, & \text{егер } s_j = 0 \end{cases};$

D) $y_i = f(s_i) = \begin{cases} -1, & \text{егер } s_j > 0 \\ 1, & \text{егер } s_j \leq 0 \end{cases};$

E) $y_i = f(s_i) = \begin{cases} 1, & \text{егер } s_j = 0 \\ -1, & \text{егер } s_j \leq 0 \end{cases}.$

4. Бейімделген сызықтық нейронда белсендіру функциясы қандай типті функция болады?

- A) сигнум;
- B) квадраттық;
- C) рационалды;
- D) векторлы;
- E) сызықтық.

5. Бейімделген сызықтық нейронға арналған салмақты айқындау формуласында η мәні нені білдіреді?

- A) оқыту қарқындылығы;
- B) оқыту мерзім;
- C) уақыт;
- D) тұрақты мән;
- E) қателік мәні.

II.3.6. Сигма-Пи нейрондар

Бұл тақырыпта Сигма-Пи нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Сигма-Пи нейрондар жасанды нейронның стандартты емес моделіне жатады және n деңгейінің (n -нейрондар кірісінің саны) s көпмүшелігімен белсендіру функциясын көрсету жағдайында сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады және былай өрнектеледі:

$$s = \sum_{k=1}^m w_k \prod_{i \in I_k} x_i + \theta,$$

мұнда I_k – алғашқы n бүтін санының мүмкін болатын 2^n комбинациясының біреуінен тұратын индекстер жиынтығы, $m = 2^n$.

Сигма-Пи нейрондар белсендіру функциясы ретінде бастапқы-таңбалық функциясын қолданады:

$$y = \text{sgn}(s) = \begin{cases} 1, & \text{егер } s \geq 0 \\ 0, & \text{егер } s < 0 \end{cases},$$

мұнда $y \in \{0,1\}$, $X \in \{0,1\}^n$, $\theta \in Z$, $W \in Z$ (Z – скалярлық көбейту), $W \cdot X$ – W және X векторларының скалярлық көбейтіндісі.

Сигма-Пи нейрондары диагностика, болжау және бейнелерді танып-білу тапсырмаларын шешу үшін пайдаланылады.

II.3.6-тапсырмалар.

1. Сигма-Пи нейрондарының формуласын жазыңыз.
2. Сигма-Пи нейрондарының белсендіру функциясын жазыңыз.

II.3.6-сұрақтар.

1. Сигма-Пи нейрондары ненің кеңеюі болады?
2. Индекстер жиындығы қанша комбинациядан тұрады?

3. Сигма-Пи нейрондары қандай тапсырмаларды шешуге қолданылады?

II.3.6- тестілер.

1. Сигма-Пи нейрондар қандай нейрондардың жиынтығынан тұрады?

А) Сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы;

В) Сызықтық және сигнумды белсендірілу функциялары бар нейрондардың жиынтығы;

С) Рационал және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы;

Д) Сызықтық және рационал белсендірілу функциялары бар нейрондардың жиынтығы;

Е) Векторлық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы.

2. Сигма-Пи нейрондарды сипаттайтын өрнек қайсысы?

А)
$$s = \sum_{k=1}^m w_k \prod_{i \in I_k} x_i + \theta;$$

В)
$$s = \sum_{k=1}^m w_k \prod_{i \in I_k} x_i;$$

С)
$$s = \sum_{k=1}^m w_k + \theta;$$

Д)
$$s = \sum_{k=1}^m w_k \prod_{i \in I_k} x_i + 1;$$

Е)
$$s = \sum_{k=1}^m w_k \prod_{i \in I_k} x_i - \theta.$$

3. Сигма-Пи нейрондары қандай жағдайда пайдаланылады?

- A) диагностика, болжау және бейнелерді танып-білу тапсырмаларын шешу үшін;
- B) тек бейнелерді танып-білу тапсырмаларын шешу үшін;
- C) диагностикалық тапсырмаларды шешу үшін;
- D) салмақтарды анықтауға арналған тапсырмаларды шешу үшін;
- E) диагностикалық және дыбыстық сигналдарды анықтауға арналған тапсырмаларды шешу үшін.

4. Сигма-Пи нейрондар функциясында индекстер жиынтығы қанша комбинациядан тұрады?

- A) 2^n ;
- B) 3^n ;
- C) 4^n ;
- D) 5^n ;
- E) $1/2^n$.

5. Сигма-Пи нейрондар белсендіру функциясы қалай жазылады?

- A) $y = \text{sgn}(s) = \begin{cases} 1, & \text{егер } s \geq 0; \\ 0, & \text{егер } s < 0; \end{cases}$
- B) $y = \text{sgn}(s) = \begin{cases} 0, & \text{егер } s \geq 0; \\ 1, & \text{егер } s < 0; \end{cases}$
- C) $y = \text{sgn}(s) = \begin{cases} 1, & \text{егер } s \geq 0; \\ 0, & \text{егер } s > 0; \end{cases}$
- D) $y = \text{sgn}(s) = \begin{cases} 1, & \text{егер } s \geq 0; \\ 0, & \text{егер } s = 0; \end{cases}$
- E) $y = \text{sgn}(s) = \begin{cases} 1, & \text{егер } s = 0 \\ 0, & \text{егер } s < 0. \end{cases}$

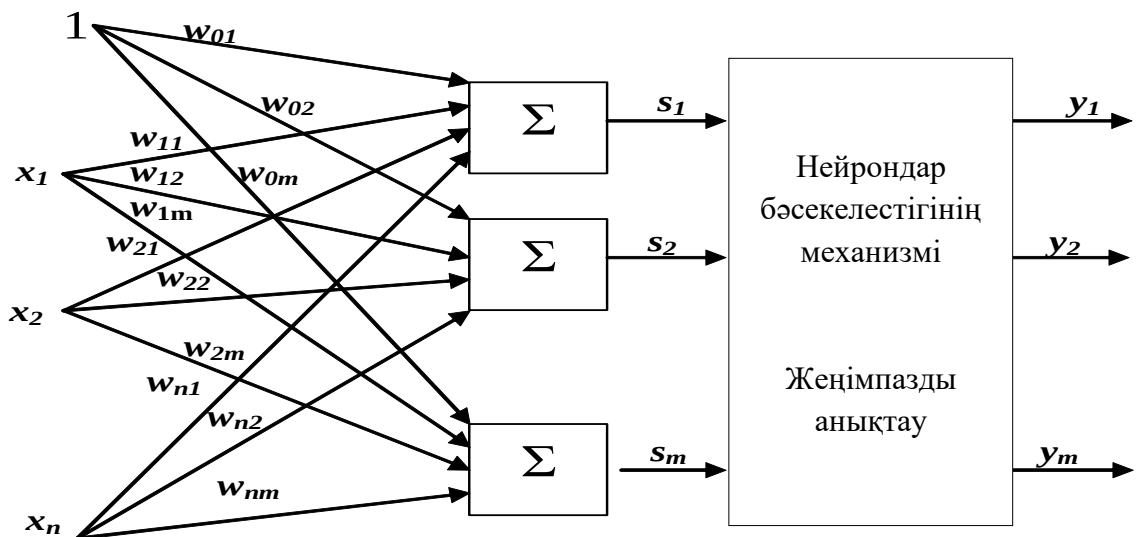
II.3.7. WTA типті нейрондар

Бұл тақырыпта WTA типті нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

WTA типті нейрондар (WInner Takes All – Жеңімпаз бәрін алады) жасанды нейрондардың стандартты емес моделіне жатады. Бұл типтің нейрондары өзара бәсекеге түсетін және сәйкес салмақтарымен w_{ij} кіру сигналдарының сомасын есептейтін стандартты қосқыштың кіріс модуліне ие топтарда қолданылады.

«Жеңімпаз бәрін алады» типті нейрондар топтарының құрылымы II.3.7-суретте көрсетілген.



II.3.7- сурет. WTA типті нейрондардың байланысу механизмі

Топтағы әрбір бәсекелес нейрон бір кіру сигналына ие. Әрбір нейрон өзінің қосқышының шығу сигналын қарапайым түрде есептейді. Қосқыштың j -ші шығу сигналы мына формула бойынша анықталады:

$$s_j = f\left(\sum_{i=0}^n w_{ij} x_i\right), \quad j=1, 2, \dots, m$$

Барлық s_j салыстыру нәтижелері бойынша s_j ең көп мәнге тең нейрон- жеңімпаз анықталады. Шығу y_j сигналы нейрон- жеңімпаздың мәні 1, қалған барлық нейрондардың (жеңілгендердің) шығу сигналдары 0 болады және олардың салмақты анықтау үдерістері бұғатталады.

WTA типті нейрондар мұғалімсіз Гроссбергтің Instar оқытуына ұқсас. Барлық нейрондар салмақтарының бастапқы мәндері салыстырмалы 1-ге кезекті қалыптандыруы бар кездейсоқ әдіспен таңдалады. Әрбір білім алушы X^k векторын ұсынғанда, жеңілдетілген ереже бойынша өзінің салмақтарын анықтауға рұқсат берілетін нейрон-жеңімпаз анықталады:

$$w_{ij}(t+1) = w_{ij}(t) + \eta(x_j^k - w_{ij}(t))$$

Барлық жеңілген нейрондар өздерінің синапстық байланыстарының салмақтарын өзгеріссіз қалдырады. Білім алудың әрбір циклінде кіріс салмағының w_i ағымдағы векторы X^k кіру векторына ең жақын болатын нейрон жеңіске жетеді. Сонымен қатар, w_i вектор X^k вектор жағына қарай түзеледі. Сондықтан оқыту үдерісінде бір-біріне жақын кіріс векторларының әрбір тобына (кластеріне) жеке нейрондар қызмет етеді.

II.3.7-тапсырмалар.

1. «Жеңімпаз бәрін алады» типті нейрондардың байланыс сұлбасын салыңыз.
2. j -ші қосқыштың шығу сигналының есептеу формуласын жазыңыз.
3. «Жеңімпаз бәрін алады» типті нейрондар үшін салмақ өзгерісінің формуласын жазыңыз.

II.3.7-сұрақтар.

1. Қандай себептерге байланысты «Жеңімпаз бәрін алады» типті нейрондары топ болып пайдаланылады?
2. «Жеңімпаз бәрін алады» типті нейрон үшін оқытудың қандай парадигмасы қолданылады?

3. Жеңілген нейрон синапстық салмақтарымен не істейді?

II.3.7-тестілер.

1. WTA типті нейрондар қандай нейрондар типіне жатады?

A) «Жеңімпаз бәрін алады» типті нейрондар;

B) «адалайн» типті нейрондар;

C) «сигнум»—«sIgnum» типті нейрондар;

D) квадраттық қосқышы бар нейрондар;

E) Паде типті нейрондар.

2. Топтағы әрбір бәсекелес нейрон неше кіру сигналына ие?

A) 1;

B) 2;

C) 3;

D) n ;

E) $2n$.

3. WTA типті нейрондарды үйрету үшін неше мұғалім қажет?

A) мұғалім қажет емес;

B) бір мұғалім жеткілікті;

C) екі мұғалім қажет;

D) көп мұғалім қажет;

E) үш мұғалім қажет.

4. WTA типті нейрондар салмақтарының бастапқы мәндері қандай әдіспен таңдалады?

A) кездейсоқ;

B) Паде;

C) Сигма пи;

D) салмақталған қосындысы;

E) ең жылдам құлау.

5. WTA типті нейрондар қайда қолданылады?

A) өзара бәсекелес және w_{lj} салмақтарымен кіру сигналдарының қосындысын есептейтін қосқыштың кіріс модуліне ие топтарда;

B) диагностика, болжау және бейнелерді танып-білу тапсырмаларын шешу үшін;

C) салмақтарды анықтауға арналған тапсырмаларды шешу үшін;

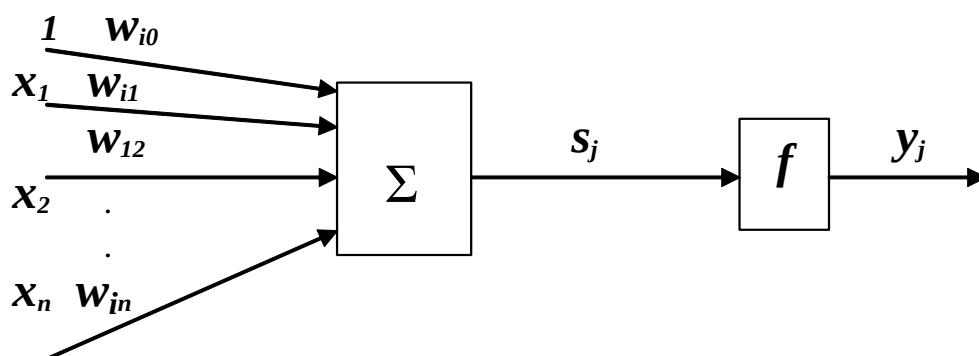
- D) тек бейнелерді танып-білу тапсырмаларын шешу үшін;
 E) диагностикалық тапсырмаларды шешу үшін.

II.3.8. Хебб нейрондары

Бұл тақырыпта Хебб нейрондары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Д.Хебб өзінің зерттеулерінде құрылымы II.3.8-суретте көрсетілгендей нейронның стандартты моделін қарастырған.



II.3.8.1-сурет. Нейронның стандартты моделінің құрылымы

Ол егер екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс күшейетіндігін байқаған. Егер y_j шығу сигналдары бар j -ші нейрон y_j шығу сигналдары бар i -ші нейронмен, w_{ij} салмақты байланысса, онда бұл клеткалардың байланыс күштеріне y_i және y_j шығу сигналдарының мәндері әсер етеді. Бұл үдерісті моделдеу үшін Д.Хебб w_{ij} нейронның салмағы оның кіру және шығу сигналдарының көбейтіндісіне пропорционал өзгертін формалды ереже ұсынды:

$$\Delta w_{ij} = \eta y_j y_i ,$$

мұнда η – мәні (0, 1) аралығында жатқан оқыту коэффициенті.

Нейронды Хебб ережесі бойынша оқыту мұғалім көмегімен де, оның көмегінсіз де жүреді. Екінші жағдайда, Хебб ережесінде нейронның шығу сигналының y_i нақты мәні қолданылады. Мұғаліммен бірге оқыту барысында y_i шығу сигналының мәнінің

орнына осы нейроннан күтілетін d_i реакция қолданылады. Бұл жағдайда Хебб ережесі келесі түрде жазылады:

$$\Delta w_{ij} = \eta y_j d_i$$

Хебб ережесін қолдану нәтижесінде w_{ij} салмақтар үлкен мәнге ие болуы мүмкін. Себебі оқытудың әрбір циклінде салмақтың ағымдағы мәнімен Δw_{ij} оның үстелуі қосылады:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}$$

Хебб ережесі бойынша оқыту үдерісін тұрақтандыру әдістерінің бірі γ ұмыту коэффициентіне азайған w_{ij} -нің соңғы мәнін есептеу болады. Сонымен бірге, Хебб ережесі келесі түрде болады:

$$w_{ij}(t+1) = w_{ij}(t)(1-\gamma) + \Delta w_{ij}$$

γ ұмыту коэффициентінің мәні $(0,1)$ интервалынан таңдалады және η оқыту коэффициентінен бірнеше пайыз құрайды. γ -дің үлкен мәнін қолдану нейронның бұрын үйренгенінің көп бөлігін ұмытуына алып келеді. Ұмыту коэффициентінің ұсынылатын мәні – $\gamma < 0,1\eta$, осы кезде нейрон оқыту үдерісінде жинақталған ақпараттың көп бөлігін сақтайды және белгілі-бір деңгейде салмақтардың мәнін тұрақтардыру мүмкіндігін алады.

Хебб ережесі жеке нейрондардың әртүрлі белсендірілу функциялары бар әртүрлі типті нейрондық желілер үшін қолданылады.

II.3.8-тапсырмалар.

1. Нейронның стандартты моделінің құрылымын салыңыз.
2. Екі белсендіру нейрон үшін ұсынылған салмақтардың өзгеріс формуласын жазыңыз.
3. Мұғаліммен бірге үйрену барысындағы Хебб ережесінің формуласын жазыңыз.

II.3.8-сұрақтар.

1. Екі белсендіру нейрондар арасында не болады?
2. Оқудың әрбір циклінде не орын алады?
3. Қандай нейрондар үшін Хебб ережесі қолданылады?

II.3.8-тестілер.

1. Д.Хебб өзінің зерттеулерінде нейронның стандартты моделін қарастыра отырып, екі нейрон арасында нені байқаған?

А) екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс күшейетіндігін;

В) екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс азаятындығын;

С) екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс 2 есеге күшейетіндігін;

Д) екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс болмайтындығын;

Е) екі нейронда уақыттың бір сәтінде белсендіру болса, екі нейрондар арасындағы байланыс 2 есеге азаятындығын.

2. Нейронды Хебб ережесі бойынша оқыту үшін мұғалім қажеттілігі қандай?

А) мұғаліммен де, мұғалімсіз де ұйымдастырылады;

В) тек бір мұғаліммен ұйымдастырылады;

С) екі мұғаліммен ұйымдастырылады;

Д) төрт мұғаліммен ұйымдастырылады;

Е) үш мұғаліммен ұйымдастырылады.

3. Хебб ережесін қолдану нәтижесінде w_{ij} салмақтар мәні қандай болуы мүмкін?

А) үлкен мәнге ие болуы мүмкін;

В) кіші мәнге ие болуы мүмкін;

С) тек бір мәнге ие болуы мүмкін;

Д) үлкен немесе кіші мәнге ие болуы мүмкін;

Е) үлкен немесе тең мәнге ие болуы мүмкін.

4. Хебб ережесіне сай келетін өрнек қайсысы?

A) $w_{ij}(t+1)=w_{ij}(t)(1-\gamma)+\Delta w_{ij}$;

B) $w_{ij}(t+1)=w_{ij}(t)(1-\gamma)$;

C) $w_{ij}(t+1)=w_{ij}(t)+\Delta w_{ij}$;

D) $w_{ij}(t+1)=(1-\gamma)+\Delta w_{ij}$;

E) $w_{ij}(t+1)=w_{ij}(t)(1-\gamma)-\Delta w_{ij}$.

5. Хебб ережесіне сай келетін өрнектегі Ұмыту коэффициентінің ұсынылатын мәні қандай?

A) $\gamma < 0,1\eta$;

B) $\gamma > 0,1\eta$;

C) $\gamma < \eta$;

D) $\gamma < 0,5\eta$;

E) $\gamma > 0,5\eta$.

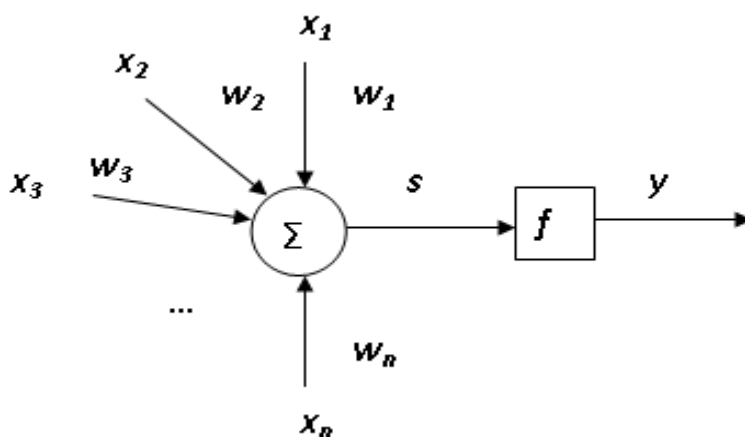
II.3.9. Гроссберг жұлдыздары

Бұл тақырыпта Гроссберг жұлдыздары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Гроссберг жұлдыздары биологиялық мидың жеке бөліктерің үлгісі болады және жасанды нейронның стандартты емес моделіне жатады. Олар белгілерді танып-білу тапсырмаларын шешу үшін нейрожелілердің компоненттері ретінде қолданылады.

Кіру жұлдызы—Input Star (Instar) формасындағы нейронның синапстық байланыстардың санына сәйкес келетін n кірістер мен y бір шығысқа ие. Кіру жұлдызының құрылымы II.3.9.1-суретте көрсетілген.



II.3.9.1-сурет. Кіру жұлдызы

Кіру жұлдызы кіріске белігілі-бір вектор түскенде, әр жолы шығысқа сигнал беруге үйренеді, яғни ол өздігінен кіріс жағдайының детекторы болады және тек өзінің кіру векторына жауап береді. Нейронның шығысында кіру векторының салмақтық вектормен түйісуін көрсететін оның кірістерінің салмақталған қосындысы құралады және белсендіру функциясына сәйкес шығу сигналы құрастырылады:

$$y = f(s) = f(W \cdot X) = f\left(\sum_{i=1}^n w_i x_i\right),$$

мұнда $X=(x_1, x_2, \dots, x_n)$ – кіру векторы; $W=(w_1, w_2, \dots, w_n)$ – салмақ векторы, y – шығу сигналы.

Instar-да белсендіру функциясының сызықтық формасы жиі қолданылады, және сол кезде $y=s$. Instar-ды үйрету– w_{ij} салмақтар Гроссберг ережесі бойынша іріктеледі:

$$w_{ij}(t+1)=w_{ij}(t)+\eta y_i(x_j-w_{ij}(t)) ,$$

мұнда $\eta \in (0,1)$ – оқыту коэффициенті.

x векторы түрінде көрсетілетін кіру деректері көп жағдайда $|x|=1$ болатын қалыптанған формада өрнектеледі. x векторының компоненттерін қалыптандыру келесі формула бойынша орындалады:

$$x_j^* = \frac{x_j}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}}$$

Гроссберг әдісі бойынша оқытудың нәтижелері белгілі бір деңгейде η оқыту коэффициентіне тәуелді. $\eta = 1$ таңдағанда, w_{ij} салмақтары бірінші интеграциядан соң-ақ x_j мәндеріне тең болады. x кезекті кіру векторын енгізу салмақтардың жаңа векторға бейімделуіне және бұрынғы мәндерге абсолютті «ұмытуға» алып келеді. $\eta < 1$ таңдауы w_{ij} синапстық байланыстар салмағын оқыту нәтижесінде x оқушы векторлардың орташаланған мәндерін қабылдайды.

Кіру жұлдызы бірлік вектордың елеусіз өзгерістеріне жауап беру мүмкіндігіне ие. Бұған осы сыныпты кез-келген векторға жауап беру мақсатында оқушы векторлардың өлшемін орташаландыру үшін оқыту үдерісінде салмақтарды баптау арқылы қол жеткізіледі.

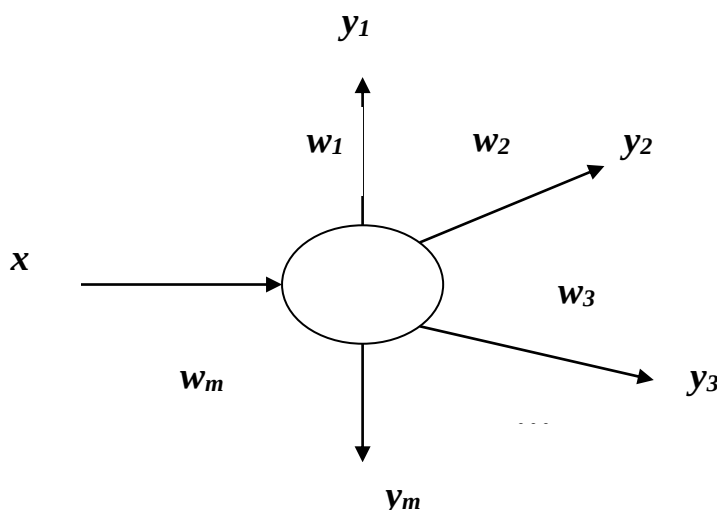
Шығу жұлдызы—Output Star (Outstar) формасындағы нейрон бір кіріске және салмақтармен бірге m шығысқа ие.

Шығу жұлдызы кіру жұлдызына қарағанда қарама-қайшы функция атқарады – кіріске сигнал түскенде салмағы бар белгілі-бір шығу векторы беріледі:

$$\sum_{i=1}^n y_i w_i = Y \cdot W = x ,$$

мұнда x – кіру сигналы; $Y=(y_1, y_2, \dots, y_m)$ – шығу векторы; $W=(w_1, w_2, \dots, w_m)$ – басында $(0,1)$ аралығынан кездейсоқ мәннің векторы иемделетін салмақтық вектор.

Шығу векторының құрылымы II.3.9.2-суретте көрсетілген.



II.3.9.2-сурет. Шығу жұлдызы

Гроссберг ережесі бойынша Outstar оқыту келесі өрнекпен көрсетілген:

$$w_{ij}(t+1) = w_{ij}(t) + \eta(y_i - w_{ij}(t)) ,$$

мұнда $w_i(t)$ және $w_i(t+1)$ – оқудың t -шы және $(t+1)$ -шы қадамындағы I -ші жұлдыздың сәйкесінше салмақтары; η – басында шамамен 1-ге тең, кейіннен оқыту үдерісінде нөлге дейін азаятын мөлшерлейтін оқыту коэффициенті.

Гроссберг жұлдызы формасындағы нейрондардың ерекшелігі жадтың іріктемелігі болады. Шығу жұлдызы формасындағы әрбір нейрон өзі соған қатысты «өзінің» бейнесін есте сақтап, қалғандарын елемейді. Әрбір шығу жұлдызына кейбір нақты командалық функциялар сәйкес келеді. Жад белгісі белгілі бір нейронмен байланысады, ол желіде нейрондар арасындағы өзара әрекет нәтижесінде туындамайды.

Кіру және шығу жұлдыздары өзара кез-келген күрделіліктегі НЖ-де байланыса алады және жеке-жеке көптеген нейрондық желілерде пайдаланыла алады. Оның ішінде, Гроссберг кіру және шығу жұлдыздарының конфигурациялары иерархиялық нейрондық желілерді құру үшін пайдаланылады.

II.3.9-тапсырмалар.

1. Алдын ала кірістер мен шығыстардың санын анықтап алып, кіру жұлдызының құрылымын салыңыз.
2. Белсендіру функциясына сәйкес шығу сигналының кіру жұлдызының құрылу формуласын жазыңыз.
3. Кіру жұлдыздарына арналған Гроссберг ережесі бойынша синапстық байланыстардың салмақ таңдау формуласын жазыңыз.
4. Кіру жұлдызы үшін кіру векторының компоненттерін қалыптандыру формуласын жазыңыз.
5. Алдын ала кірістер мен шығыстардың санын анықтап алып, шығу жұлдызының құрылымын салыңыз.
6. Белсендіру функциясына сәйкес шығу сигналының шығыс жұлдызының құрылу формуласын жазыңыз.
7. Гроссберг ережесі бойынша шығу жұлдызының оқыту формуласын жазыңыз.
8. Шығу жұлдызы қандай функция атқарады?

II.3.9-сұрақтар.

1. Биологиялық ми үшін Гроссберг ұсынған жұлдыздар не болады?
2. Белгілерді танып-білу тапсырмаларын шешу үшін Гроссберг жұлдыздары қалай қолданылады?
3. Қандай кіру жұлдызы кіріс жағдайының детекторы болады және тек өзінің кіру векторына жауап береді?
4. Кіру жұлдызын Гроссберг әдісі бойынша оқыту нәтижелері қандай параметрлерге тәуелді?
5. Салмақтар жаңа векторға бейімделуі мен бұрынғы білімдерді толық «ұмытуы» қашан орын алады?

6. Кіру жұлдызы қандай мүмкіндікке ие?
7. Кіру және шығу жұлдыздары өзара байланыса алады ма және жеке дара қолданылады ма?
8. Әрбір шығу жұлдызына не сәйкес келеді?

II.39.-тестілер.

1. Гроссберг жұлдыздарының қандай ерекшеліктері бар?
 - A) биологиялық мидың жеке бөліктердің үлгісі болады және жасанды нейронның стандартты емес моделіне жатады;
 - B) «Жеңімпаз бәрін алады» типті нейрондар, жасанды нейрондардың стандартты емес моделіне жатады;
 - C) Жасанды нейронның стандартты емес моделіне жатады және сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады;
 - D) Жасанды нейронның стандартты емес моделіне жатады және нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді;
 - E) Екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады.
2. Гроссберг жұлдыздары қандай жағдайда қолданылады?
 - A) Белгілерді танып-білу тапсырмаларын шешу үшін нейрожелілердің компоненттері ретінде;
 - B) Өзара бәсекеге түсетін және сәйкес салмақтарымен w_{lj} кіру сигналдарының қосындысын есептейтін стандартты қосқыштың кіріс модуліне ие топтарда;
 - C) Диагностика, болжау және бейнелерді танып-білу тапсырмаларын шешу үшін;
 - D) Салмақтарды анықтауға арналған тапсырмаларды шешу үшін;
 - E) Тек бейнелерді танып-білу тапсырмаларын шешу үшін.
3. Гроссберг ережесі бойынша Outstar (шығу жұлдызы) қандай өрнекпен беріледі?

- A) $w_{ij}(t+1)=w_{ij}(t)+\eta(y_i-w_{ij}(t))$;
- B) $w_{ij}(t+1)=w_{ij}(t)+\eta$;
- C) $w_{ij}(t+1)=w_{ij}(t)+\eta(y_i+w_{ij}(t))$;
- D) $w_{ij}(t-1)=w_{ij}(t)+\eta(y_i-w_{ij}(t))$;
- E) $w_{ij}(t+1)=\eta(y_i-w_{ij}(t))$.

4. Кіру жұлдыздары қандай мүмкіндікке ие?

- A) Бірлік вектордың елеусіз өзгерістеріне жауап беру мүмкіндігіне ие;
- B) Бір кіріске және салмақтармен бірге m шығысқа ие;
- C) Әрбір нейрон өзі соған қатысты «өзінің» бейнесін есте сақтап, қалғандарын елемей мүмкіндігіне ие;
- D) Өзара кез-келген күрделіліктегі НЖ-де байланыса алады және жеке-жеке көптеген нейрондық желілер де пайдаланыла алады;
- E) Оған сигнал түскенде салмағы бар белгілі-бір шығу векторы беріле алады.

5. Шығу жұлдызы қандай формуламен жазылады?

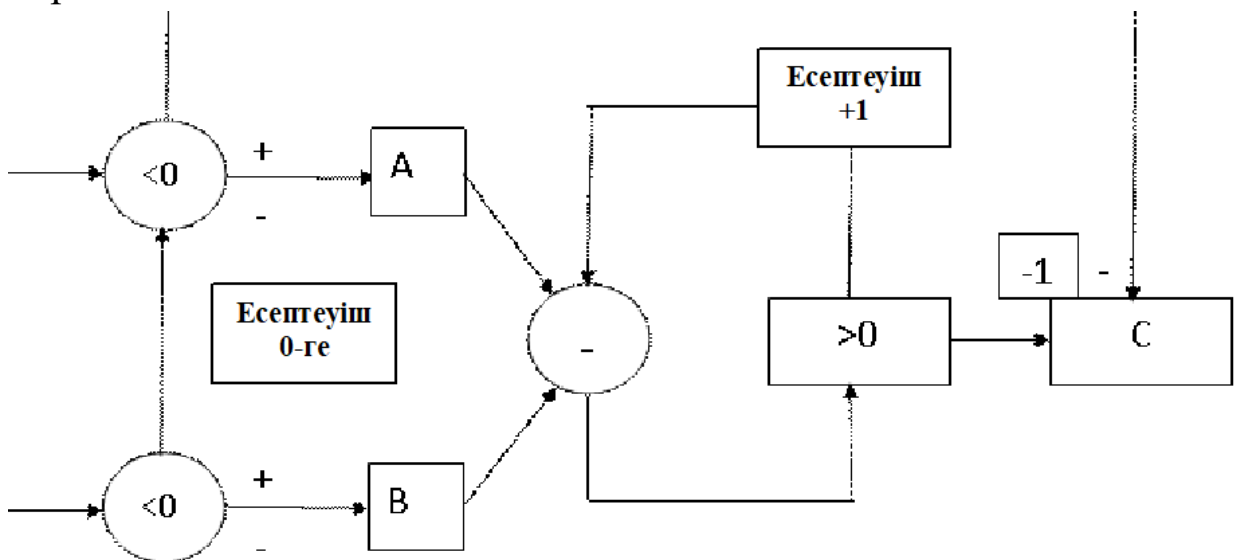
- A) $\sum_{i=1}^n y_i w_i = Y \cdot W = X$
- B) $y = f(s) = f(W \cdot X) = f(\sum_{i=1}^n w_i x_i)$
- C) $w_{ij}(t+1)=w_{ij}(t)+\eta(y_i+w_{ij}(t))$;
- D) $w_{ij}(t-1)=w_{ij}(t)+\eta(y_i-w_{ij}(t))$;
- E) $w_{ij}(t+1)=\eta(y_i-w_{ij}(t))$.

II.3.10. Есептеуіші сәйкесті нейрондар

Бұл тақырыпта есептеуіші сәйкесті нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Есептеуіші сәйкесті нейрондар жасанды нейрондардың стандартты емес моделіне жатады. Мұндай нейрон, кірісте n -өлшемді X векторын қабылдап, шығыста x_i мәніне a_i мәні сәйкестік санына тең бүтін сан, яғни $x_i = a_i$ болатын I санын береді.

Есептеуіші сәйкесті нейрондардың сұлбасы II.3.10.1-суретте көрсетілген.



II.3.10.1-сурет. Есептеуіші сәйкесті нейрондардың жалпы сұлбасы

Есептеуіші сәйкесті нейрондар жұмысының алгоритмі үш кезеңге бөлінеді:

Бірінші кезеңде X және A векторлары кіріске беріледі. Нәтижесінде $(X-A)$ вектор алынатын, n элементке бөлу арқылы (X) кіру векторынан (A) салмақтық коэффициенттер векторын алу операциясы жүреді.

Екінші кезеңде, сәйкестік санын есептеу үшін $(X-A)$ векторында сәйкестіктер үшін 0-ге және сәйкестіктер емес үшін 1-ге өзгерту қажет. Ол үшін теріс мәнді жою үшін алдын ала жасыру арқылы

белсендірілудің арифметикалық функциясы қолданылады. $(X-A)$ векторының белсендірілу нәтижесінде $x_i \neq a_i$ позициясында бірлікке, $x_i = a_i$ позициясында нөлге тең M векторы алынады.

Үшінші кезеңде M векторы X кіріске түседі және оның барлық элементтерін қосады. Нәтижесінде X және A векторлары элементтерінің сәйкессіздік саны шығады, ол соңғы кезеңде x_i мәнінің a_i мәніне сәйкестік санын алу үшін элементтердің жалпы санынан алынады.

II.3.10-тапсырмалар.

1. Кірісте n -өлшемді X векторын алғандағы есептеуіші сәйкесті нейронның жұмысын суреттеңіз.
2. Есептеуіші сәйкесті нейронның құрамын көрсетіп, жалпы сұлбасын сызыңыз.
3. Есептеуіші сәйкесті нейрон жұмысының бірінші кезеңдегі әрекетін сипаттаңыз.
4. Есептеуіші сәйкесті нейрон жұмысының екінші кезеңдегі әрекетін сипаттаңыз.
5. Есептеуіші сәйкесті нейрон жұмысының үшінші кезеңдегі әрекетін сипаттаңыз.

II.3.10-сұрақтар.

1. Есептеуіші сәйкесті нейрондар жасанды нейронның қандай моделіне жатады?
2. Есептеуіші сәйкесті нейрон қандай жағдайда шығыста бүтін сан береді?
3. Есептеуіші сәйкесті нейрон жұмысының алгоритмі қанша кезеңнен тұрады?
4. $(X-A)$ векторының белсендірілуі нәтижесінде есептеуіші сәйкесті нейронда не орын алады?
5. Есептеуіші сәйкесті нейронда x_i -ның a_i -мен сәйкестік санын алу үшін не істелінеді?

II.3.10-тестілер.

1. Есептеуіші сәйкесті нейронның жұмысы қалай сипатталады?

А) кіріске n -өлшемді X векторын ала отырып, шығыста x_i мәніне a_i мәні сәйкестік санына тең бүтін сан, яғни $x_i = a_i$ болатын I санын береді;

В) «Жеңімпаз бәрін алады» типті нейрондар, жасанды нейрондардың стандартты емес моделіне жатады;

С) Жасанды нейронның стандартты емес моделіне жатады және сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады;

Д) Жасанды нейронның стандартты емес моделіне жатады және нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді;

Е) екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады.

2. Есептеуіші сәйкесті нейрондар жұмысының алгоритмі неше кезеңге бөлінеді?

А) 3;

В) 2;

С) 4;

Д) 5;

Е) 6.

3. Есептеуіші сәйкесті нейрон жұмысының бірінші кезеңінде не істеледі?

А) X және A векторлары кіріске беріледі;

В) Сәйкестік санын есептеу үшін $(X-A)$ векторында сәйкестіктер үшін 0-ге және сәйкестіктер емес үшін 1-ге өзгерту қажет;

С) M векторы X кіріске түседі және оның барлық элементтерінің жиынтығы жүреді;

D) A векторы X кіріске түседі және оның барлық элементтерін қосады;

E) X және A векторлары шығысқа беріледі.

4. Есептеуіші сәйкесті нейрондар жұмысының екінші кезеңінде қандай әрекеттер жасалады?

A) Сәйкестік санын есептеу үшін $(X-A)$ векторында сәйкестіктер үшін 0-ге және сәйкестіктер емес үшін 1-ге өзгерту қажет;

B) X және A векторлары кіріске беріледі;

C) M векторы X кіріске түседі және оның барлық элементтерінің жиынтығы жүреді;

D) A векторы X кіріске түседі және оның барлық элементтерінің жиынтығы жүреді;

E) X және A векторлары шығысқа беріледі.

5. Есептеуіші сәйкесті нейрондар жұмысының үшінші кезеңінде қандай әрекеттер жасалады?

A) M векторы X кіріске түседі және оның барлық элементтерінің жиынтығы жүреді;

B) X және A векторлары кіріске беріледі;

C) Сәйкестік санын есептеу үшін $(X-A)$ векторында сәйкестіктер үшін 0-ге және сәйкестіктер емес үшін 1-ге өзгерту қажет;

D) A векторы X кіріске түседі және оның барлық элементтерінің жиынтығы жүреді;

E) X және A векторлары шығысқа беріледі.

II.3.11. Кубтық нейрондар

Бұл тақырыпта кубтық нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Кубтық нейрондар жасанды нейрондардың стандартты емес моделіне жатады. Нейрондардың кубтық моделінде X екілік кіріс сигналдардың векторы құрамы 0-ге немесе 1-ге тең жад ұяшықтарының адресі ретінде қарастырылады. X векторының n өлшемділігі үшін 2^n мүмкін адрестер бар болады.

Жад ұяшықтарын n -өлшемді *гиперкубтың* төбесі ретінде қарастыруға болады. Жад ұяшықтары бір-біріне тәуелсіз мәндер алады. Жад ұяшықтарын полярланған ± 1 екілік мәндерді сақтаушы ретінде қарастырған жөн. Сонда кубтық модульдің жұмысы келесі түрде суреттеледі: x екілік кірісі жад адресі ретінде қолданылады, полярланған екілік шама оқылады және h функциясымен полярланбаған формаға конвертттеледі, яғни, $y=h(v_x)$, мұнда v_x – x адресі бойынша мән.

Мұндай модульдерді біз, белсендірілу мәндер адрестері жиынының геометриялық көрінісін гиперкуб шыңдарының жиыны ретінде көрсетуі үшін, кубтық деп атаймыз.

Екі кірісті кубтық модульді қарастырайық. Белсендірілудің 4 мәні бар $\{v_{00}, v_{01}, v_{10}, v_{11}\}$. Белсендірілу өрнегі келесі түрде болады:

$$s=v_{00}(1-x_1)(1-x_2)+v_{01}(1-x_1)x_2+v_{10}x_1(1-x_2)+v_{11}(1+x_1)(1+x_2)$$

$$x=(x_1, x_2) \text{—кіру векторы.}$$

Мұндай жазу сомадағы көбейтінділердің біреуі нөлдік болу қажеттілігінен туындаған. x_1 және x_2 полярланған кіріс үшін белсендірілу болады

$$s=[v_{00}(1-x_1)(1-x_2)+v_{01}(1-x_1)x_2+v_{10}x_1(1-x_2)+v_{11}(1+x_1)(1+x_2)]/4$$

n кірісі бар модуль жағдайында аламыз:

$$s = \left[\sum_{i=0}^m v_{ij} \prod_{j=1}^n (1 - x_j) \right] / 2^n$$

Кубтық нейрондар олардың жад ұяшықтарының мазмұнын өзгерту арқылы оқытылады. ‘+’ арқылы ұяшық ішіндегісін +1-ге орнату – *инкремент* амалын, ал ‘-’ арқылы ұяшық ішіндегісін -1-ге орнату – *декремент* амалын белгілейміз.

Мейлі, бастапқы жағдайда кубтық нейронның барлық ұяшықтары нөлге орнатылған болсын. Оқыту таңдамасымен адрестелетін ұяшықтарды орталық ұяшықтар немесе орталықтар ретінде белгілейміз. Хемминг қашықтығы тұрғысынан орталықтарға жақын ұяшықтарды, орталықтың өзіне немесе соған жақын мәндерге баптаймыз, яғни орталық айналасындағы ұяшық мәндерінің кластерленуі орын алуы қажет. Бұл шарт кубтық нейрондардың желісі үшін орындалуы керек.

Оқыту алгоритмі ұяшық мәнін ең жақын орталық мәнімен анықтайтын, ал орталықтан бірдей қашықтықтағы ұяшықтар мәндерін нөлге тең күйде қалдыратын *Вороной бөлінуі* дегенді құрастырады.

Кубтық нейрондар жартылай сызықтық нейрондарға қарағанда үлкен функционалдылыққа жол береді, сондықтан, модульдердің саны аз болғанда сондай есептерді де шешуге мүмкіндік береді.

II.3.11-тапсырмалар.

1. Екі кірісті кубтық модульдың белсендірілуіне арналған өрнекті жазыңыз.
2. Полярланған n кірістерге арналған өрнекті жазыңыз.
3. Кубтық және жартылай сызықтық нейрондарды салыстырыңыз.

II.3.11-сұрақтар.

1. Нейрондардың кубтық моделіне кіретін екілік сигналдардың векторы қалай қарастырылады?
2. Оқудың әрбір циклында не орын алады?
3. Оқыту алгоритмі нені құрастырады?

II.3.11-тестілер.

1. Нейрондардың кубтық моделінің кірісі қалай сипатталады?

А) Нейрондардың X екілік кіріс сигналдардың векторы құрамы 0-ге немесе 1-ге тең жад ұяшықтарының адресі ретінде қарастырылады;

В) Жасанды нейрондардың стандартты емес моделіне жатады және кіріске n -өлшемді X векторын ала отырып, шығыста x_i мәніне a_i мәні сәйкестік санына тең бүтін сан, яғни $x_i = a_i$ болатын I санын береді;

С) Жасанды нейронның стандартты емес моделіне жатады және сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады;

Д) Жасанды нейронның стандартты емес моделіне жатады және нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді;

Е) Екі қосқыштан, салмақталған қосындыдан бөліндіні есептейтін бір элементтен және функционалды түрлендіргіштің бір блогынан тұрады.

2. Бастапқы күйде кубтық нейронның барлық ұяшықтары неде орналасқан?

А) 0-де;

В) 1-де;

С) соңғы мәнде;

Д) бастапқы мәнде;

Е) кездейсоқ мәнде.

3. Кубтық нейрондардан тұратын желі үшін қандай шарт орындалу керек?

А) орталық айналасындағы ұяшық мәндерінің кластеризациясы орын алуы қажет;

В) Модульдардың саны аз болғандағы тапсырмаларды шешу қажет;

С) Хемминг қашықтығы тұрғысынан орталықтарға алыс ұяшықтарда сол немесе соған жақын мәндерге бапталуы қажет;

Д) Жад ұяшықтарын n –өлшемді *гиперкубтың* төбесі ретінде қарастыру қажет;

Е) Екі кірісті кубтық модулды қарастыру қажет.

4. Екі кірісті кубтық модулды қарастырғанда белсендірілудің неше мәні бар?

А) $\{v_{00}, v_{01}, v_{10}, v_{11}\}$, яғни 4;

В) $\{v_{00}, v_{01}, v_{10}\}$, яғни 3;

С) $\{v_{00}, v_{01}\}$, яғни 2;

Д) v_x , яғни тек бір ғана;

Е) $\{v_{00}, v_{01}, v_{10}, v_{11}, v_{12}\}$, яғни 5;

5. Белсендірілу өрнегінің кіру векторы қайсысы?

А) $x=(x_1, x_2)$;

В) $x=1$;

С) $x=0$;

Д) $x=n$;

Е) $x=-1$.

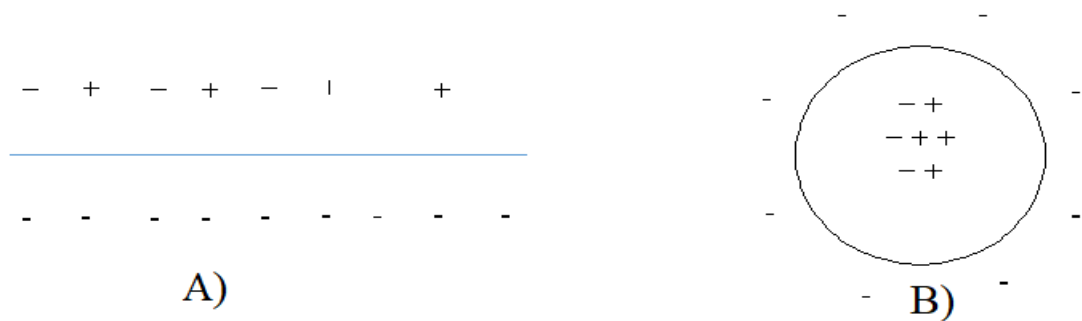
II.3.12 Радиалды нейрондар

Бұл тақырыпта радиалды нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Радиалды нейрондар стандартты түрге жатпайтын жасанды нейрон моделіне жатады. Олар орталық нүктенің айналасында шар тәріздес кеңістік бөлігін жасайтын гиперсфера пішіндес болады. Дәл осы көз қараспен қарағанда ол сигмоидалды нейронның табиғи толықтауышы болады, себебі деректердің шеңберлі симметриялы болғандары әртүрлі сыныптарға бөлгенде нейрондардың сандарын әлдеқайда азайтуға мүмкіндік береді.

II.3.12.1-суретте деректер кеңістігін сигмоидалық және радиалдық нейрондармен бөлу тәсілдері көрсетілген.



II.3.12.1-сурет. Деректер кеңістігін бөлу әдісі

A) сигмоидалды нейронмен;

B) радиалды нейронмен.

Радиалды нейрондар таңдалған орталықтың айналасында радиалды өзгертін және тек осы орталық маңайында нөлдік мән қабылдайтын функцияны жүзеге асырады. Осындай функциялар мына түрде $f(X,C)=\|X-C\|$ анықталып *радиалды базистік функциялар* деп аталады, мұнда $X \in R^n$, C – ортасы.

Радиалды базистік функцияларының мысалдыры:

1. Мультиквадраттық функция: $f(X,C)=\left(\|X-C\|^2+a^2\right)^{\frac{1}{2}}$, где $a>0$.

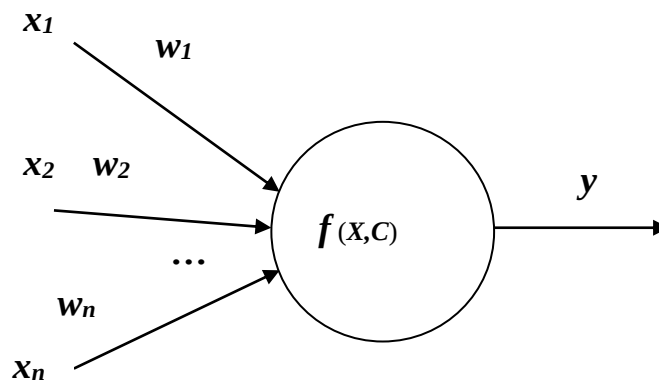
2. Жұқа қабыршық сплайны: $f(X,C)=\left(\|X-C\|^2 \ln(\|X-C\|)\right)^{\frac{1}{2}}$.

3. Кері мультиквадраттық функция: $f(X,C) = \frac{1}{(\|X-C\|^2 + a^2)^{\frac{1}{2}}}, a > 0$

4. Гаусс функциясы: $f(X,C) = \exp\left(-\frac{\|X-C\|^2}{2\sigma^2}\right), \sigma > 0$ –

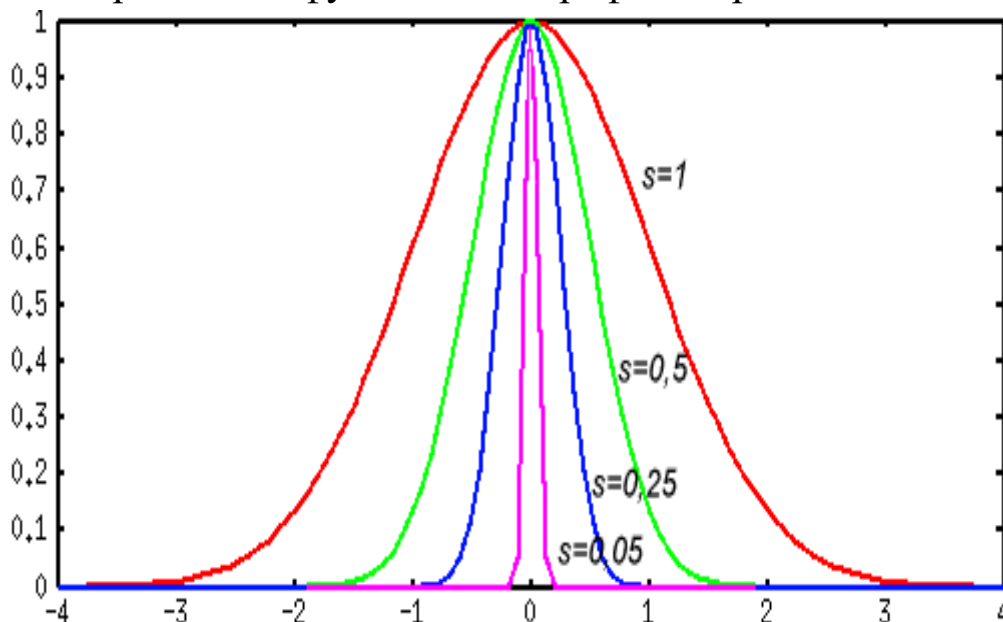
функцияның "енін" анықтайтын параметр.

II.3.12.2-суретте радиалды нейрон құрылымы көрсетілген:



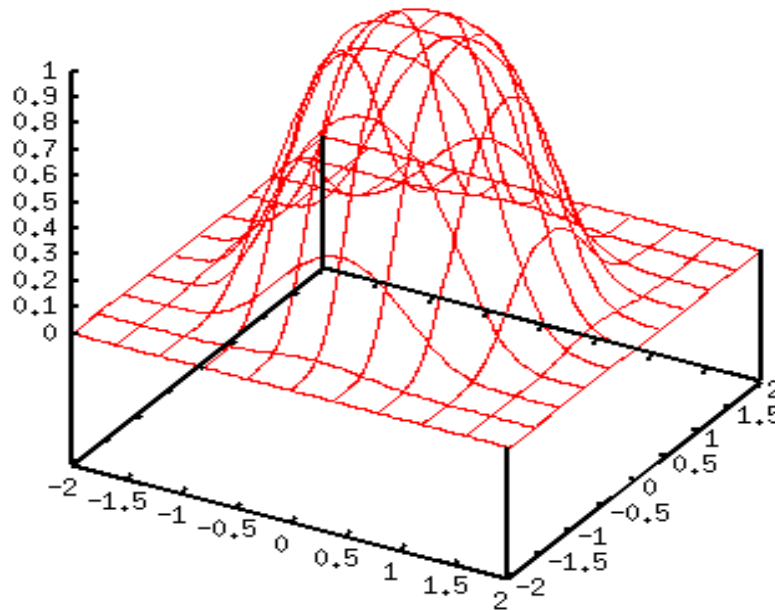
II.3.12.2-сурет. Радиалды нейрон құрылымы

II.3.12.3-суретте түрлі мәндер үшін s_I скаляр түрінде бірөлшемді радиалды функцияның графигі көрсетілген.



II.3.12.3-сурет. Бірөлшемді радиалды функциясын жоспарлау

II.3.12.4-суретте s_I түрлі мәндер үшін скаляр түрде екіөлшемді радиалды функцияның графигі көрсетілген.



II.3.12.4-сурет. Екіөлшемді кесте радиалды функциясы

Радиалды нейронның сигмоидалды нейрон мен перцептроннан негізгі айырмашылығы мынада: сигмоидалды нейрон кіріс сигналдарының көпөлшемді кеңістігін гипержазықтықпен, ал радиалды нейрон – гиперсферамен бөледі.

II.3.12-тапсырмалар.

1. Деректер кеңістігін радиалды және сигмоидалды нейрондармен бөлу тәсілдерін салыңыз.
2. Радиалды базистік функция үшін өрнекті жазыңыз.
3. Мультикватраттық функция үшін өрнекті жазыңыз.
4. Жіңішке қабыршық сплайн функциясы үшін өрнекті жазыңыз.
5. Кері мультикватраттық функция үшін өрнекті жазыңыз.

II.3.12-сұрақтар.

1. Радиалды нейрондар қандай функцияны жүзеге асырады?
2. Радиалды функцияны қалай өзгертуге болады?
3. Орталық айналасында радиалды функция қандай мәндерге ие болады?

4. Екіөлшемді радиалды функция графигі қандай кескінге графикаға ұқсас болады?

5. Радиалды нейронның сигмоидальқ нейрон мен перцептроннан түбегейлі айырмашылығы қандай?

II.3.12-тестілер.

1. Радиалды нейрондар қандай нейрондарға жатады?

А) Стандартты емес түрге жататын жасанды нейрон моделіне жатады және олар орталық нүктенің айналасында шар тәріздес кеңістік бөлігін жасайтын гиперсфера пішіндес болады, ол сигмоидальды нейронның табиғи толықтауышы болады;

В) Жасанды нейрондардың стандартты емес моделіне жатады және онда X екілік кіріс сигналдардың векторы құрамы 0-ге немесе 1-ге тең жад ұяшықтарының адресі ретінде қарастырылады;

С) Жасанды нейрондардың стандартты емес моделіне жатады және кіріске n -өлшемді X векторын ала отырып, шығыста x_i мәніне a_i мәні сәйкестік санына тең бүтін сан, яғни $x_i = a_i$ болатын I санын береді;

Д) Жасанды нейронның стандартты емес моделіне жатады және сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады;

Е) Жасанды нейронның стандартты емес моделіне жатады және нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді.

2. Радиалды базистік функция үшін қандай өрнек белгіленген?

А) $f(X, C) = \|X - C\|$;

В) $f(X, C) = \|X + C\|$;

С) $f(X, C) = \|X - C\| + 1$;

Д) $f(X, C) = \|X - C\| - 1$;

Е) $f(X, C) = \|X - C\| + X$.

3. $f(X, C) = \|X - C\|$ функциясында X қандай мәндер болады?

A) $X \in R^n$;

B) $X \in R$;

C) $X \in N^n$;

D) $X \in Z^n$;

E) $X \in N$.

4. Орталық айналасында радиалды функция қандай мәндерге ие болады?

A) нөлдік;

B) бүтін;

C) нақты;

D) тұрақты;

E) тек бір ғана мәнге ие.

5. Радиалды функцияның мысалдарына қандай функциялар жатады?

A) Мультиквадраттық функция, Сплайн жұқа қабыршық, Кері мультиквадраттық функция, Гаусс функциясы жатады;

B) Мультиквадраттық функция, Сигма пи функциясы, Кері мультиквадраттық функция, Гаусс функциясы жатады;

C) Мультиквадраттық функция, Ссплайн жұқа қабыршық, Кедергілі функция, Гаусс функциясы жатады;

D) Мультиквадраттық функция, Векторлы функция, Кері мультиквадраттық функция, Гаусс функциясы жатады;

E) Мультиквадраттық функция, Рационалды функция, Кері мультиквадраттық функция, Гаусс функциясы жатады.

II.3.13. Стохастикалық нейрондар

Бұл тақырыпта стохастикалық нейрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тесттестілер қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылады [5-38].

Жоғарыда жасанды нейронның детерминделген моделі сипатталған, онда шығысындағы жағдай кіру сигналдарының қосқыш жұмысымен анықталады. Сондай-ақ, жасанды нейронның стандартты емес моделіне жататын стохастикалық нейрондар бар, онда нейрондарды ауыстыру индукцияланған локалды өрістен тәуелді болатын ықтималдылықпен жүреді, яғни белсендіру функциясы келесі өрнекпен анықталады:

$$f(s) = \begin{cases} 1, & \text{егер } P(s) \\ 0, & \text{егер } 1-P(s) \end{cases},$$

мұнда P ықтималдықты тарату әдетте сигмоида түрінде беріледі:

$$\sigma(s) = \frac{A(T)}{1 + e^{(-s/T)}},$$

ал нормалдаушы A тұрақтысы ықтималдықты үлестіруді нормалдау шарты үшін енгізіледі:

$$\int_0^1 \sigma(s) ds = 1.$$

Осылайша, нейрон P ықтималдықпен белсендіріледі. T параметрі— температура аналогы және нейрондық желінің тәртіпсіздігін анықтайды. Егер T параметрі 0-ге ұмтылса, стохастикалық нейрон Хэвисайдтың табыстама функциялы кәдімгі нейронға ауысады.

Стохастикалық моделдің детерминделген моделден айырмашылығы, нейронның шығыстық күйі кіру сигналдарының салмақталған салмағынан ғана емес, сонымен қатар, әрбір жүзеге асыру кезінде мәндері $(0,1)$ аралығын таңдап алынатын қандай да бір кездейсоқ айнымалы шамадан тәуелді болады.

Нейронның стохастикалық моделінде y шығу сигналы келесі ықтималдылықпен ± 1 мәнін қабылдайды:

$$P(y_i = \pm 1) = 1 / (1 + e^{\pm 2\beta s_i}) ,$$

мұнда s_I – I -ші нейронның кіру сигналдарының салмақталған қосындысы, β – көбінесе 1-ге тең болатын оң тұрақты .

Стохастикалық моделде нейронды оқыту үдерісі келесі кезеңдерден тұрады:

1) желінің әр нейроны үшін салмақталған қосынды есептеледі:

$$2) s_j = \sum_{i=0}^n w_{ij} x_i ;$$

3) y_i шығыстық айнымалы ± 1 мәнін қабылдайтын P ықтималдылықты есептеу:

$$P(y_i = \pm 1) = \frac{1}{1 + e^{\pm 2\beta s_i}} ;$$

4) $R \in (0, 1)$ кездейсоқ айнымалының мәнін тудыру;

5) егер $R < P(y)$ болса y , қарсы жағдайда $-y$ шығу сигналын қалыптастыру;

6) қолданылатын ережелер бойынша w_{Ij} (y_I белгіленгенде) салмақтарының бейімделуі.

Уидроу-Хофф ережесі бойынша мұғаліммен оқыту кезінде салмақтардың бейімделуі келесі формуласы бойынша жүзеге асырылады:

$$\Delta w_i = \eta x_i (d_i - y_i)$$

мұнда η – мәні $(0, 1)$ аралығында жататын нейрондарды оқыту коэффициенті.

Бір-бірімен және сыртқы ортамен белгілі бір үлгіде біріктірілген нейрондардың жиынтығы нейрондық желіні құрайды. Қандай нейрондық желілер бар екендігін, олардың мүмкіндіктерін дұрыс бағалау үшін, нейрондардың және нейрондық желілердің бар моделдерін қалай жақсарту керектігін және бұл саладағы жаңа

әзірлеулер қандай болатындығын білу үшін, оларға сыныптау жүргізу керек.

Нейрондарды және нейрондық желілерді сыныптаудың көптеген тәсілдері бар. Келесі бөлімде нейрондық желілердің мүмкін болатын кейбір сыныптаулары қарастырылған.

II.3.13-тапсырма.

1. Нейронның стохастикалық моделі үшін белсендіру функциясының өрнегін жазыңыз.

2. Нейронның стохастикалық моделі үшін ықтималдықтарды тарату өрнегін жазыңыз.

3. Полярланған n кірістер үшін белсендіру өрнегін жазыңыз.

4. Әрбір нейрон үшін салмақталған қосындыны есептеу өрнегін жазыңыз.

5. Нейронның стохастикалық моделінің шығу сигналын алу ықтималдылығының өрнегін жазыңыз.

6. Нейронның стохастикалық моделін Уидроу-Хофф ережесі бойынша мұғаліммен оқыту кезінде салмақтарды бейімдеу өрнегін жазыңыз.

7. Нейрондық желілерді сыныптау үшін қойылатын талаптарды атаңыз.

II.3.13-сұрақтар.

1. Стохастикалық жасанды нейрондардың детерминделген нейрондардан айырмашылығы қандай?

2. Нейронның стохастикалық моделі үшін ықтималдықты тарату өрнегінде A нормалау тұрақтысы не үшін енгізіледі?

3. Нейронның стохастикалық моделі үшін ықтималдықты тарату өрнегінде T параметрі нені білдіреді?

4. Нейронның стохастикалық моделінің шығыс күйі тәуелді болатын кездейсоқ айнымалы шаманы әр жүзеге асыру кезінде, кездейсоқ айнымалы шамалардың мәні әрбір іске асырылу барысында қандай аралықты таңдап алынады?

5. Стохастикалық моделде нейрондарды оқыту үдерісі қандай кезеңдерден тұрады?

6. Бір-бірімен және сыртқы ортамен нақты бір үлгімен біріктірілген нейрондар жиынтығы нені құрайды?

7. Нейрондық желілер сыныптауы не үшін керек?

II.3.13-тестілер.

1. Стохастикалық нейрондарға қандай мінездеме беріледі?

A) Жасанды нейронның стандартты емес моделіне жатады және олар нейрондарды ауыстыру индукцияланған локалды өрістен тәуелді болатын ықтималдылықпен жүреді;

B) Жасанды нейрондардың стандартты емес моделіне жатады және онда X екілік кіріс сигналдардың векторы құрамы 0-ге немесе 1-ге тең жад ұяшықтарының адресі ретінде қарастырылады;

C) Жасанды нейрондардың стандартты емес моделіне жатады және кіріске n -өлшемді X векторын ала отырып, шығыста x_i мәніне a_i мәні сәйкестік санына тең бүтін сан береді;

D) Жасанды нейронның стандартты емес моделіне жатады және сызықтық және квадраттық белсендірілу функциялары бар нейрондардың жиынтығы болады;

E) Жасанды нейронның стандартты емес моделіне жатады және нейронның функционалды түрлендіргішіне кіру сигналының салмақталған квадраттық формасын береді.

2. Стохастикалық нейрондардың белсендіру функциясы қандай өрнекпен анықталады?

$$A) f(s) = \begin{cases} 1, & \text{егер } P(s) \\ 1, & \text{егер } 1-P(s) \end{cases};$$

$$B) f(s) = \begin{cases} 1, & \text{егер } P(s) \\ 1, & \text{егер } 1+P(s) \end{cases};$$

$$C) f(s) = \begin{cases} 1, & \text{егер } P(s)=1 \\ 1, & \text{егер } 1-P(s) \end{cases};$$

$$D) f(s) = \begin{cases} 1, & \text{егер } P(s) \\ 1, & \text{егер } 1 < P(s) \end{cases};$$

$$E) f(s) = \begin{cases} 1, & \text{егер } P(s) \\ 1, & \text{егер } 1 > P(s) \end{cases}.$$

3. Стохастикалық моделдің детерминделген моделден айырмашылығы неде?

A) Нейронның шығыстық күйі кіру сигналдарыдың салмақталған салмағынан ғана емес, сонымен қатар мәндері (0,1) аралығынан әрбір жүзеге асыру кезінде таңдап алынатын қандай да бір кездейсоқ айнымалы шамадан тәуелді болады;

B) Нейронның шығыстық күйі кіру сигналдарыдың салмақталған салмағынан ғана емес, сонымен қатар мәндері (0,1) аралығынан әрбір жүзеге асыру кезінде таңдап алынатын қандай да бір кездейсоқ айнымалы шамадан тәуелсіз болады;

C) Нейронның шығыстық күйі кіру сигналдарыдың салмақталған салмағынан ғана емес, сонымен қатар мәндері (0,10) аралығынан әрбір жүзеге асыру кезінде таңдап алынатын қандай да бір кездейсоқ айнымалы шамадан тәуелді болады;

D) Нейронның кірісітік күйі шығу сигналдарыдың салмақталған салмағынан ғана емес, сонымен қатар мәндері (0,1) аралығынан әрбір жүзеге асыру кезінде таңдап алынатын бір ғана айнымалыға тәуелді болады;

E) Нейронның шығыстық күйі кіру сигналдарыдың салмақталған салмағынан ғана емес, сонымен қатар бір ғана айнымалыға тәуелді болады.

4. Уидроу-Хофф ережесі бойынша мұғаліммен оқыту кезінде салмақтардың бейімделуі қандай формуламен жүзеге асырылады?

A) $\Delta w_i = \eta x_i (d_i - y_i)$;

B) $\Delta w_i = \eta x_i (d_i + y_i)$;

C) $\Delta w_i = \eta (d_i - y_i)$;

D) $\Delta w_i = \eta x_i (d_i - y_i) + 1$;

E) $\Delta w_i = \eta x_i (d_i + y_i) + 1$.

5. Нейронды $\Delta w_i = \eta x_i (d_i - y_i)$ формуласымен оқытқанда η коэффициентінің мәні қандай аралықта жатады?

A) (0,1);

B) (-1;1);

C) (-1;0);

D) $(-\infty; +\infty)$;

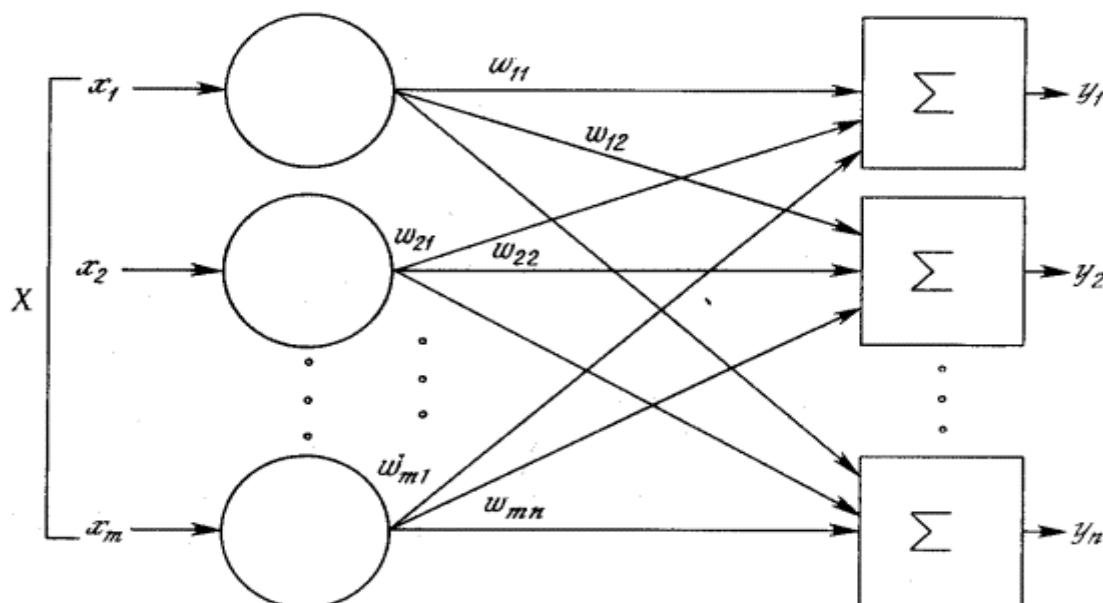
E) $(0; +\infty)$.

II.4. Нейрондық желілердің моделі

II.4.1. Бір қабатты нейрондық желілер

Бұл тақырыпта бір қабатты нейрондық желілер қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Бір нейрон танудың қарапайым процедурасын орындау қабілетіне ие болуына қарамастан, маңызды нейрондық есептеулер үшін нейрондарды желіге біріктіру керек. Нейрондар тобынан тұратын, бір қабатты қарапайым нейрондық желінің құрылымы II.4.1.1- суретте көрсетілген:



II.4.1.1-сурет. Бір қабатты нейрондық желі құрылымы

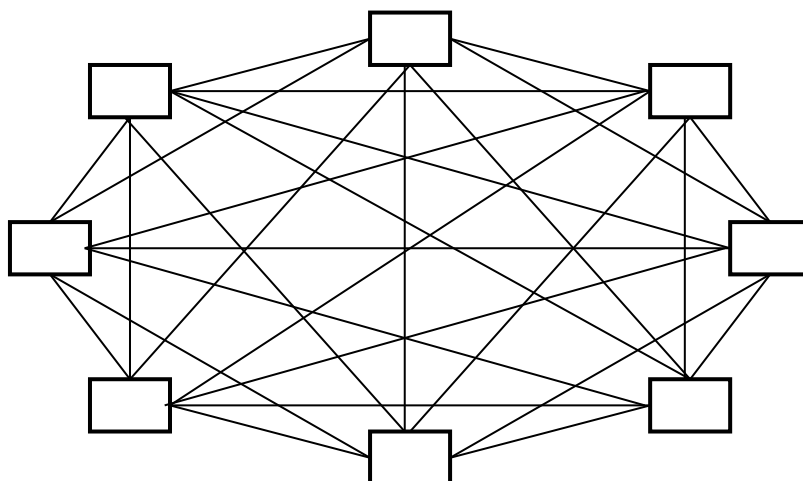
Мұнда төбелер–сол жағындағы шеңберлер–кіру сигналдарын тарату үшін қолданылады. Олар қандай да бір есептеулерді орындамайды және сол себепті қабат болып саналмайды. X кірістер векторының әрбір x_I ($I=1,2,\dots, n$) элементі жеке салмақпен әрбір j -ші ($j=1,2,\dots, m$) нейронмен қосылған. Әрбір нейрон салмақтанған қосындыны желіге береді. Салмақтар n жолдар мен m бағандардан тұратын W матрицасының элементтері болады, мұнда n – кірістер саны, ал m – нейрондар саны. Мысалы, w_{21} – бұл бірінші нейронмен екінші кірісті байланыстыратын салмақ.

Компоненттері $y_1, y_2, \dots, y_m$ нейрондар шығысы болатын Y шығу векторын есептеу, $S=XW$ матрицалық көбейтіндіге f белсендіру функциясын қолданғанға алып келеді, мұнда S, X – векторлар-жолдар.

Бірқабатты нейрондық желі байланыс түрі бойынша *толық байланысты* және *әлсіз байланысты (регулярлы)* болып бөлінеді.

Толық байланысты нейрондық желілерде барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналын басқа нейрондарға, соның ішінде өзіне де береді. Шығу сигналдары желінің бірнеше такті жұмыс істеуінен кейінгі нейрондардың барлық кейбір шығу сигналдары болуы мүмкін.

II.4.1.2-суретте толық байланысқан нейрондық желі құрылымының мысалы ұсынылған.

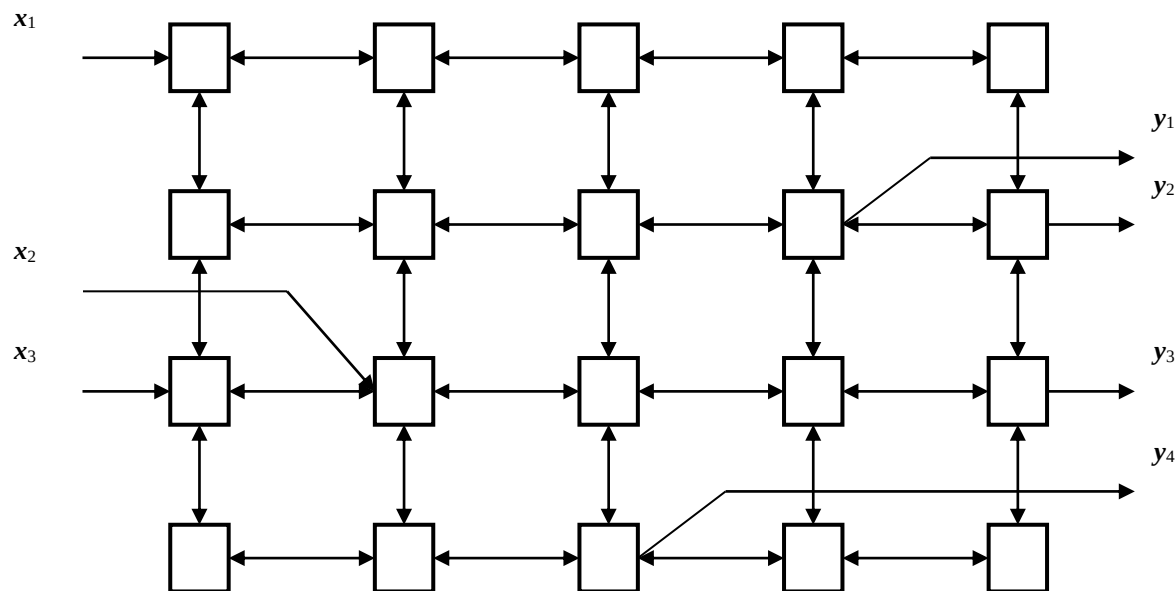


II.4.1.2-сурет. Толық байланысқан нейрондық желі құрылымы

Жүйелі (регулярлы байланысқан) нейрондық желілерде нейрондардың кез келген кірістік және шығыстық болуы мүмкін және әрбір нейрон байланыстардың тек нақты бір санына ие болады. Мұнымен қоса, нейрондық байланыстар саны желідегі оның орналасуы орнынан байланысты болады. Егерде әр нейрон 4, 6 және 8 көршілерімен байланысқан болса, онда ол сәйкесінше *фон Нейманның, Голейдің* және *Мурдың* аймағы болады.

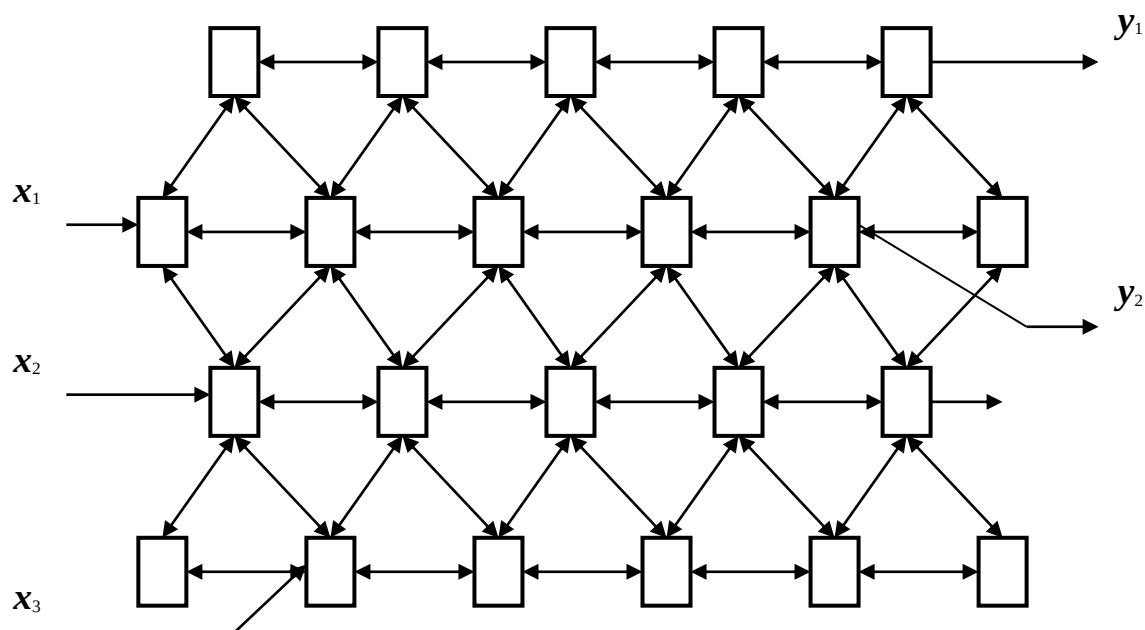
II.4.1.3-суретте регулярлы нейрондық желінің мысалы келтірілген, онда бұрыштық нейрондар (желінің бұрыштарында орналасқан) көлденең және тігінен орналасқан екі жақын

нейронмен ғана байланысқан, ал шекаралық нейрондар (сол жақтан және оң жақтан шетте тұрғандар) үш жақын нейрондармен: шекара бойымен орналасқан екі нейронмен, және ішкі бір нейронмен байланысқан.



II.4.1.3-сурет. Регулярлы нейрондық желі

II.4.1.4-суретте ішкі нейрондары алты байланыстан тұратын регулярлы гексогональді нейрондық желісінің мысалы берілген.



II.4.1.4-сурет. Регулярлы гексагоналды нейрондық желі

II.4.1-тапсырмалар.

1. Бір қабатты нейрондық желісіне анықтама беріңіз.
2. Бір қабатты нейрондық желісінің құрылысын салыңыз.
3. Бір қабатты нейрондық желісінің формуласын жазыңыз.

4. Егер барлық кіріс сигналдар барлық нейрондарға жіберілсе, және әр нейрон өзінің шығу сигналын қалған нейрондармен қатар өзіне жіберетін болса, онда бұл нейрондық желісінің қандай түріне жатанын анықтаңыз.

5. Толық байланысты нейрондық желісінің құрылысын сипаттаңыз.

6. Егер кез-келген нейрон кіріс және шығу нейроны бола алса және әр нейронның байланыс саны шектеулі болса, онда бұл нейрондық желісінің қандай түріне жатанын анықтаңыз.

7. Егер нейрон 4 көршілерімен байланысқан болса, оның аймағын атаңыз.

8. Егер нейрон 6 көршілерімен байланысқан болса, оның аймағын атаңыз.

9. Егер нейрон 8 көршілерімен байланысқан болса, оның аймағын атаңыз.

II.4.1-сұрақтар.

1. Бөлек жасанды нейронадарды белгілі желінің ішінде не себепті байланыстырамыз?

2. Бір қабатты нейрондық желі қандай элементтерден тұрады?

3. Кіру векторы X -тің ішіндегі әрбір x_i элементі немен байланысқан?

4. Персептронның орналасу қабаты қандай функцияны атқарады?

5. Бір қабатты нейрондық желілер қандай түрлерге жіктеледі?

6. Желінің функциялануының бірнеше такттарынан кейін қандай элементтер толық нейрондық желісінің шығу сигналдары бола алады?

7. Қандай нейрондық желісінде нейрон байланысының саны оның желіде орналасуына тәуелді болады?

8. Регулярлы жасанды нейрондық желісі дегеніміз не?

9. Регулярлы жасанды гексагоналды нейрондық желісі дегеніміз не?

II.3.13-тестілер.

1. Бір қабатты нейрондық желі неше түрге бөлінеді?

- A) 2-ге;
- B) 3-ке;
- C) 4-ке;
- D) 5-ке;
- E) 6-ға.

2. Толық байланысты нейрондық желінің қандай құрылымы бар?

A) Барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналдарын басқа нейрондарға, соның ішінде өзіне де береді;

B) Нейрондардың кез келген кірісі және шығысы болуы мүмкін, және әрбір нейрон байланыстардың тек нақты бір санына ие болады;

C) Бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

D) Олар қандай да бір есептеулерді орындамайды және нақты бір есептеулерді орындайтын төбелер-шаршылардан ерекшеленеді;

E) Практикада қарапайым есептермен шектеледі.

3. Егерде әр нейрон 4, 6 және 8 көршілерімен байланысқан болса, онда ол сәйкесінше қандай аймаққа жатады?

A) фон Нейманның, Голейдің және Мурдың аймағы болады;

B) Хопфилдтың, Розенблаттың және Мурдың аймағы болады;

C) фон Нейманның, Хопфилдтың және Розенблаттың аймағы болады;

D) фон Нейманның, Мурдың және Розенблаттың аймағы болады;

E) фон Нейманның, Мурдың және Кохонен аймағы болады.

4. Регулярлық нейрондық желі қалай анықталады?

А) Нейрондардың кез келген кірістік және шығыстық болуы мүмкін, және әрбір нейрон байланыстардың тек нақты бір санына ие болады;

В) Барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналдарын басқа нейрондарға, соның ішінде өзіне де береді;

С) Бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

Д) Олар қандай да бір есептеулерді орындамайды және нақты бір есептеулерді орындайтын төбелер-шаршылардан ерекшеленеді;

Е) Практикада қарапайым есептермен шектеледі.

5. Регулярлы гексагоналды нейрондық желі деген не?

А) Ішкі нейрондары бірнеше байланыстан тұрады;

В) Ішкі нейрондары байланыспайды;

С) Ішкі нейрондары тек бір ғана байланыстан тұрады;

Д) Ішкі нейрондары екі байланыстан тұра;

Е) Ішкі нейрондары шекара бойында орналасқан екі нейронның байланысынан тұрады.

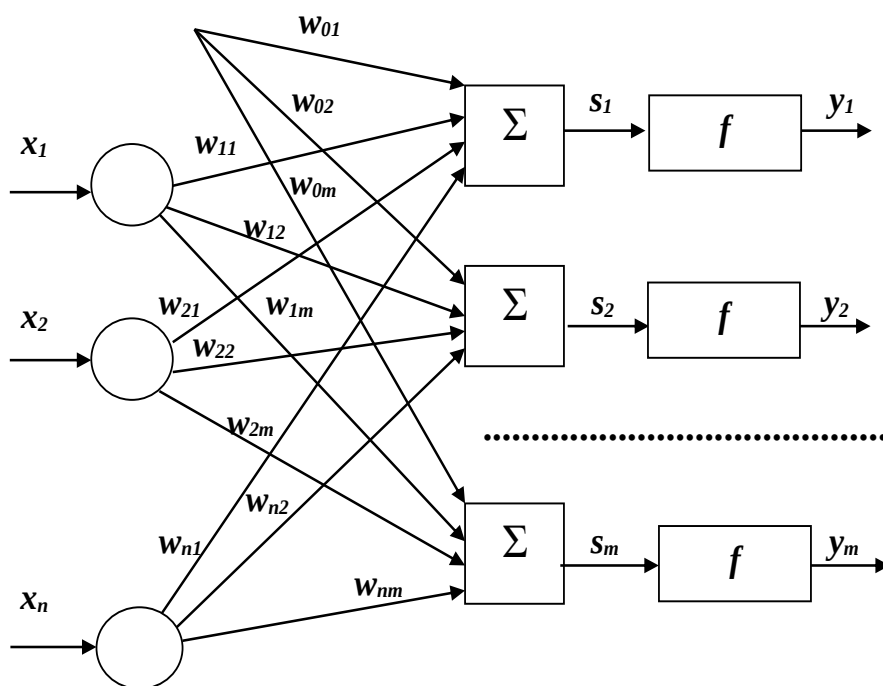
II.4.2 Бір қабатты персептрондар

Бұл тақырыпта бір қабатты персептрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Бір нейрондық персептрондар қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады. Бұл персептрондарда нейрондар бір немесе бірнеше қабаттар құрауы мүмкін, онда әрбір қабат синапстық байланыстардың салмақтарының жиынтығынан, одан кейін салмақталған сигналдарды қосатын нейрондардан тұрады. Мұндай персептрондардың ішінде ең қарапайымы бір қабатты персептрондар болады.

Әдетте бір қабатты персептрондар екі қабатпен бейнеленеді: нейрондардың бірінші қабаты *таратушы*, ал нейрондардың екінші қабаты – *өңдеуші* болады.

n кіру сигналдары бар және m нейрондардан тұратын бір қабатты персептрондардың құрылысы II.4.2-суретте көрсетілген.



II.4.2.-сурет. Бір қабатты персептрондардың құрылысы

Мұнда сол жағындағы төбелер-шеңберлер *таратушы қабаттар*. Олар қандай да бір есептеулерді орындамайды және нақты бір есептеулерді орындайтын төбелер-шаршылардан ерекшеленеді.

Таратушы қабат кіру сигналдарын, синапстық байланыстармен және белсендіру функциясымен сәйкесінше кіріс ақпараттарын түрлендіретін өңдеуші қабатқа жібереді. Мұнымен қатар, таратушы қабатта жеке салмақпен X кіріс векторларындағы әр элемент өңдеуші қабаттағы барлық нейрондармен қосылған, және бұл қосылу W салмақтар матрицасымен беріледі. Әрбір нейрон скалярлық функция (скалярлық көбейтінді) көмегімен ұсынылған s_j кірістерінің салмақталған қосындысын береді. s_j скалярлық функциялар жиынтығы қабаттың белсендіру функциясының m кірісінің m -элементтік векторымен қосылады. Нейрон қабаттарының шығысын y_j вектор-бағаны қалыптастырады.

Бір қабатты персептрон кіріс сигналдар элементтерінің салмақталған қосындысын есептейді, меженің (ығысудың) мәнін азайтады және нәтижені шығысы $+1$ -ге немесе -1 -ге теңесетін қатаң межелік функция арқылы өткізеді:

$$y_j = f(s_j) = f\left(\sum_{i=1}^n w_{ij} x_i - w_{0j}\right),$$

мұнда x_i – I -ші кіру сигналдары, j -ші нейронның (бинарлы–цифрлық немесе нақты–аналогтық) мәндер типімен; w_{ij} – таратушы қабаттың I -ші нейроны мен өңдеуші қабаттың j -ші нейроны арасындағы синапстық байланыстардың салмақтары; w_{0j} – кірістік қабаттың j -ші нейронның межесі (ығысуы); s_j – j -ші нейронның салмақталған қосындысы; y_j – j -ші нейронның шығу сигналы, f – сызықты емес белсендіру функциясы; n – кірістер саны (синапстық салмақтар); $I=1, 2, \dots, n$; m – нейрондар саны, $j=1, 2, \dots, m$.

Егер w_{ij} салмағының мәні оң (теріс) болса, онда I -ші синапстық байланыс қоздырушы (тежеуші) болады.

Мұнда n екілік кірісі бар нейрон, 0 және 1-ден тұратын кірістің (үлгілердің) 2^n әр түрлі мәндерін қабылдай алады. Әрбір үлгі екі әр түрлі (0 және 1) бинарлық шығыстарға сәйкес келетіндіктен, барлығы n айнымалыдан 2^{2^n} функциясы ғана болады.

n айнымалыдан сызықты-бөлінген функциялар саны II.4.2-кестеде берілген.

II.4.2-кесте. Сызықты-бөлінген функциялар саны.

<i>№</i>	2^{2^n}	<i>Сызықты-бөлінген функциялар саны</i>
1	4	4
2	16	14
3	256	104
4	65536	1 882
5	$4,3 \times 10^9$	94 572
6	$1,8 \times 10^{19}$	15 028 134

II.4.2-кестеден көріп отырғанымыздай, кездейсоқ түрде таңдап алынған функция сызықты-бөлінген болу ықтималдығы, тәуелсіз айнымалылардың қалыпты саны үшін де тым кішкентай. Осы себептен бір қабатты персептрондар практикада қарапайым есептермен шектелген.

Сыныптарды сызықты бөлу шешуші ережені құрастырудан тұрады:

$$f(W \cdot X) > 0, \Rightarrow X \in \text{бірінші сыныпқа};$$

$$f(W \cdot X) \leq 0, \Rightarrow X \in \text{екінші сыныпқа};$$

яғни, $W = (w_1, w_2, \dots, w_n)$ векторын табу, мұнда w_0 - барлық мүмкін болатын берілетін сигналдар үшін теңсіздікті дәл орындау мүмкін болатын шек.

Массалар орталығына бөліп – шешуші ережені құрастырудың қарапайым тәсілі. Бұл тәсілдің мәнісі келесі формуламен персептрондардың салмақтар векторын есептеуден тұрады

$$W = \left(\sum_{i=1}^n X_1^i - \sum_{j=1}^m X_2^j \right) / (n+m),$$

мұнда X_1^i – бірінші сыныпты жиындардағы векторлар, ал X_2^j – екінші сыныптағы жиындардағы векторлар.

Массалар орталығын бөлуге негізделіп құрастырылған сызықты шешуші ережелер, оқытушы іріктеуден алынатын мысалдарда қателесуі мүмкін. Дегенмен, бұл әдіс персептронды оқытатын алгоритм үшін салмақтар векторының бастапқы мәндерін (инициализациясын) анықтау құралы ретінде пайдалы.

Бір қабатты персептрон сыныптар саны көп болатын сыныптау есептерін шешу мүмкіндігін береді.

II.4.2-тапсырмалар.

1. Бір қабатты персептронға анықтама беріңіз.
2. Бір қабатты персептронның құрылысын сызыңыз.
3. Бір қабатты персептронның формуласын жазыңыз.
4. Сыныптарды сызықты бөлу үшін шешуші ережені жазыңыз.
5. Бір қабатты персептронның салмақтар векторын есептеу формуласын жазыңыз.
6. Персептронды белсендіру функциясының формуласын жазыңыз.
7. Персептронның сигмоида түріндегі белсендіру функциясының графигін салыңыз.

II.4.2-сұрақтар.

1. n екілік кірісі бар нейронның қанша әр түрлі үлгісі болады?
2. Бір қабатты персептрон деген не?
3. Персептронның таратушы қабаты қандай қызмет атқарады?
4. Сызықты-бөлу функцияларының саны қалай есептеледі?
5. Персептронды оқып үйрену үшін қандай ереже қолданылады?
6. Қосынды квадраттық қателерді минималдау үшін қандай әдіс қолданылады?

7.Топпен оқығанда бейімділік қадамының шамасы қалай анықталады?

II.4.2-тестілер.

1. Бір нейрондық персептрондар қандай болады?

А) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

В) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама аз болады;

С) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды таратудан алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

Д) барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналдарын басқа нейрондарға, соның ішінде өзіне де береді;

Е) нейрондардың кез келген кірістік және шығыстық болуы мүмкін, және әрбір нейрон байланыстардың тек нақты бір санына ие болады.

2. Әдетте бір қабатты персептрондар неше қабатпен бейнеленеді?

А) 2;

В) 3;

С) 1;

Д) 4;

Е) бірнеше.

3. Персептронның бірінші қабаты не болады?

А) таратушы болады;

В) өңдеуші болады;

С) кіру қабаты болады;

Д) шығу қабаты болады;

Е) түрлендіруші қабат болады.

4. Пекрсефронның екінші қабаты не болады?

А) өңдеуші болады;

В) таратушы болады;

С) кіру қабаты болады;

Д) шығу қабаты болады;

Е) түрлендіруші қабат болады.

5. Бір қабатты персефрон қандай есептерді шешуге мүмкіндік береді?

А) сыныптар саны көп болатын сыныптау есептерін шешуге;

В) сыныптар саны аз болатын сыныптау есептерін шешуге;

С) бірінші сыныпты жиындардағы векторларды табуға;

Д) екінші сыныпты жиындардағы векторларды табуға;

Е) кіріс сигналдар элементтерінің салмақталған айырмасын есептеуге.

II.4.3 Көп қабатты нейрондық желілер

Бұл тақырыпта көп қабатты нейрондық жүйелер қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Көп қабатты нейрондық желілерде нейрондар қабаттарға бірігеді. Қабат кіру сигналдарының бірыңғай векторы бар нейрондар жиынтығын қамтиды. Қабаттардағы нейрондардың саны әр түрлі болуы мүмкін және басқа қабаттардағы нейрондар санына тәуелді емес. Жалпы жағдайда желі солдан оңға қарай нөмірленген $q > 1$ қабатынан тұрады. Сыртқы кіру сигналдары кіру қабатындағы нейрондардың кірісіне беріледі, ал желі шығысы болып соңғы қабаттағы шығу сигналдары болады. Көп қабатты нейрондар желісінде кіріс және шығыс қабаттарынан басқа, бір немесе бірнеше жасырын қабаттар бар. Қандай да бір q қабатының шығу нейрондарынан келесі $(q+1)$ қабатының кіру нейрондарына байланысы *тізбекті* деп аталады. Бір қабаттың ішінде бір ғана белсендіру функциясы қолданылады. *Кіру қабаты* нөлмен нөмірленеді және желінің $x_1, x_2, \dots, x_n$ кіру сигналдарының векторын қабылдайды, шығу қабаты q -мен нөмірленеді және желінің $y_1, y_2, \dots, y_m$ шығу сигналдар векторын өндіреді, ал *жасырын қабаттар* k -мен $(k=1, 2, \dots, q-1)$ нөмірленеді. Бір қабатты нейрондардың белсендіру функциялары бірдей.

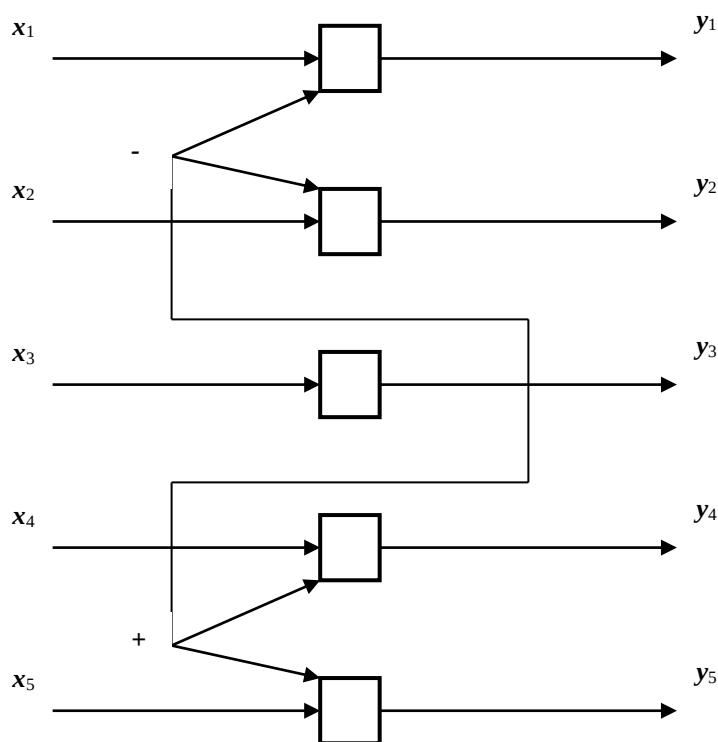
Көп қабатты нейрондық желі байланыс мәніне тәуелді монотонды болуы мүмкін. *Монотондық нейрондық желіде* әрбір қабат, шығу қабатынан басқа, екі блокқа бөлінген: *қоздырушы* және *тежеуші* блоктар. Блоктар арасындағы байланыстар да өзара қоздырушылар мен тежеушілер болып бөлінеді. Егер A блогының нейрондарынан B блогының нейрондарына тек қоздырғыш байланыстары жеткізсе, онда B блогының кез келген шығу сигналы A блогының кез келген шығу сигналының монотонды кемімейтін функциясы болып есептеледі. Егер дәл осы байланыстар тежелетін болып табылса, онда B блогының кез келген шығу сигналы A

блогының кез келген шығу сигналының монотонды өспейтін функциясы болады. Монотонды нейрондық желілер үшін нейрондардың шығу сигналдары кіру сигналдар параметрлерінен монотонды тәуелділігі қажет.

Жалпы жағдайда кез келген құрылымды көп қабатты нейрондық желілерде байланыстардың әр түрлі конфигурациясы бар болуы мүмкін. Байланыс түрлері бойынша көп қабатты нейрондық желілер: *латералды байланысты нейрондық желілер, тура байланысты нейрондық желілер, қиылыс байланысты нейрондық желілер, кері байланысты нейрондық желілер* болып бөлінеді.

Латералды (бүйірлі) байланысты көп қабатты нейрондық желілер жасырын қабаттарда суреттерді кездейсоқ фонда кереғарлылауға мүмкіндік береді және графикалық бейнелерді тану тиімділігін арттырады. Латералды байланыстар қоздырушы және тежеуші болуы мүмкін.

Көп қабатты нейрондық желілердің біріндегі латералды байланыстары бар бір қабат II.4.3.1-суретте көрсетілген.

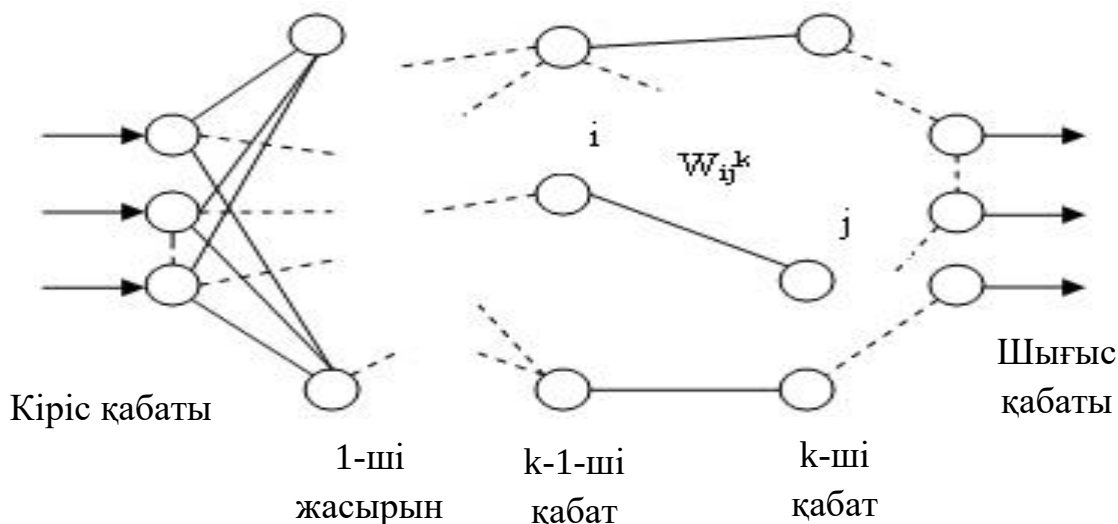


II.4.3.1-сурет. Латералды байланысты нейрондық желісінің қабаты

Бұл жерде орталық нейроннан екі жоғарғыға синапстық байланыстың теріс коэффициенті арқылы сигнал бере отырып, осы нейрондардың белсенділігінің төмендеуіне қол жеткіземіз. Орталық нейроннан екі төменгіге синапстық байланыстың оң коэффициенті арқылы сигнал бере отырып, осы нейрондардың белсенділігі жоғарылайды. Сондықтан да орталық нейрон мен екі төменгі нейрон айтарлықтай екі жоғарғы нейрондардан тиімдірек болады, бұлар кіру сигналдарының кереғарлығын арттырады.

Тура байланысты көп қабатты желілерде (сигналдардың тура таралуы) әр қабаттардың нейрондары өзара бір-бірімен байланыспаған және барлық байланыстар қатаң түрде кіру нейрондарынан шығу нейрондарына бағытталған, яғни олардың графтарында тұзақтар мен циклдер жоқ. Желіге келіп түсетін кіру сигналы кіру қабатының нейрондарына беріледі, барлық қабат арқылы өтіп және кіру сигналдары дәл кіріске берілген мезгілде шығу қабатына нейрондардың шығумен аяқталады. Желіде сигналдың таралауына қарай ол бастапқы мәніне, түрлендіруші функциялардың әсеріне және байланыстар салмақтарының көлеміне тәуелді өзгерістерге ұшырайды.

Тура байланыстары бар k -қабатты нейрондық желілердің сұлбасы II.4.3.2-суретте көрсетілген.



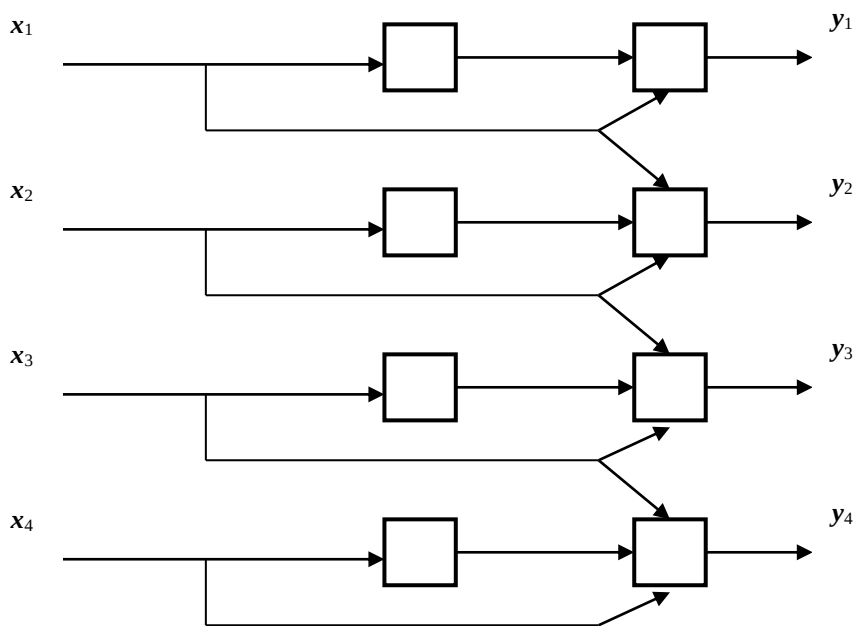
II.4.3.2-сурет Тура байланысты k -қабатты нейрондық желісінің сұлбасы

Тура байланысты нейрондық желілерде әрбір нейрон шығыс сигналдары элементтерінің салмақтанған қосындысын орындайды және кейін алынған нәтижелер бойынша белсендіру функциясының көмегімен нейрондық желілердің шығысы болып саналатын мәндерге сызықты емес түрлендіру орындалады.

Тура байланысты көп қабатты нейрондық желілердің арасында, I -қабатындағы әрбір нейронның шығысы $(I+1)$ -қабатындағы әрбір нейронның кірісімен байланысқан болғандарды толық байланыстылар деп ажыратады. Осындай тура байланысты нейрондық желілер уақытша тежеуші нейрондық желілер болады.

Айқыш байланысты нейрондық желілерде екінші дәрежеден басталатын, алдындағы қабаттардағы нейрондардың шығысымен, сондай-ақ желінің тікелей сыртқы кірісімен бапталатын синапстық байланыстары бар нейрондарды қамтиды.

Айқыш байланыстары бар екі қабатты нейрондық желінің мысалы II.4.3.3-суретте келтірілген.



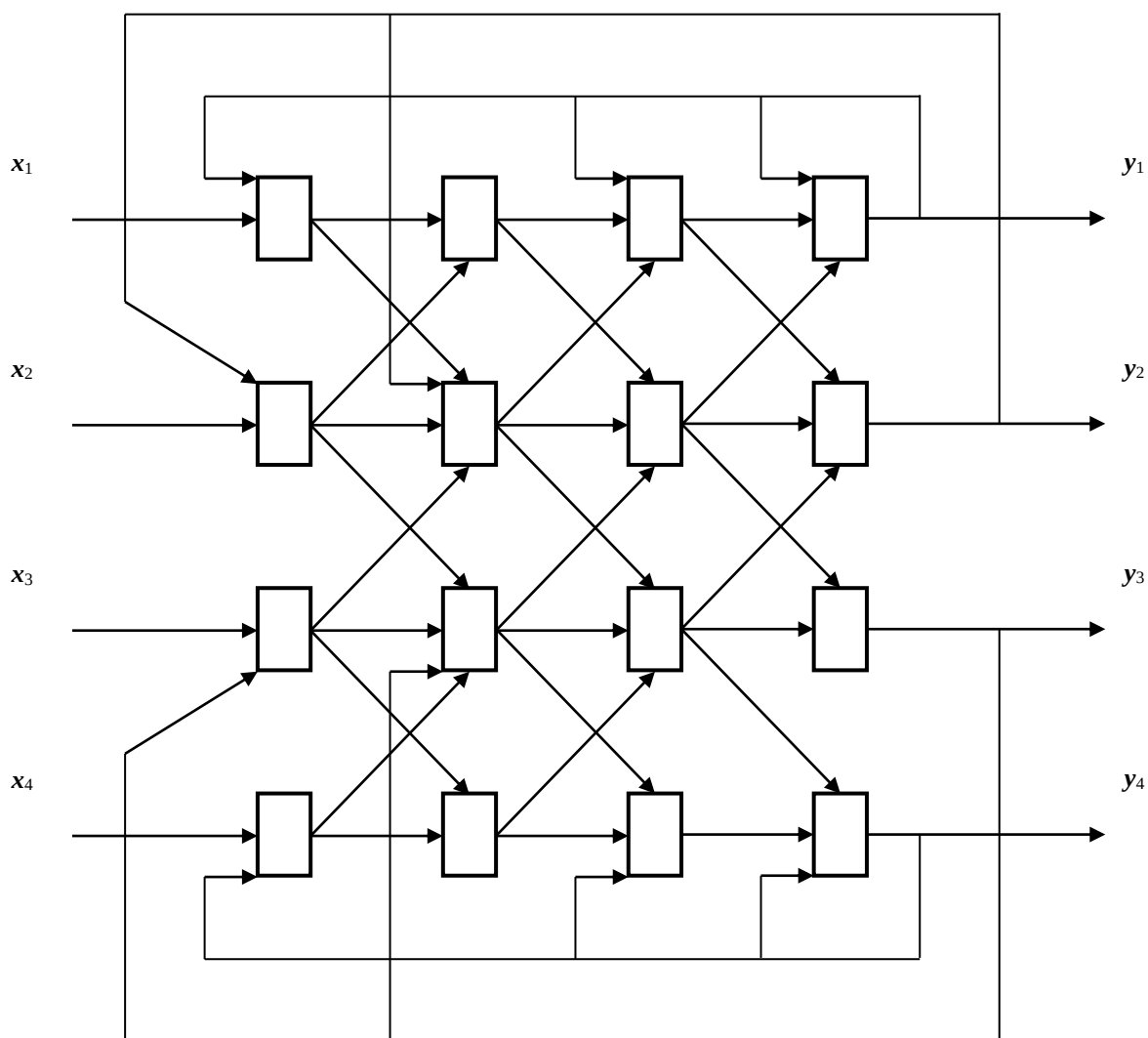
II.4.3.3-сурет. Айқыш байланысты екі қабатты нейрондық желі

Кері байланысты нейрондық желілерде әрбір нейрон кіріс ақпаратын сыртқы ортадан және басқа да нейрондардан, тіпті өзінен де алатындай (кері байланыспен) етіп ұйымдастырылған. Мұндай кері байланыспен жеке қабатты да, бірнеше қабатты да

және барлық желіні де қамтып алуға болады, яғни кері байланысты нейрондық желілер графы тұзақтардан немесе циклдерден тұрады.

Кері байланысты нейрондық желілер динамикалық нейрондардан тұрады, олардың іс әрекеттері бірінші реттік дифференциалдық теңдеулермен немесе рекуррентті айырымдық қатынастарымен сипатталады. Сондықтан да, оларды кейде динамикалық немесе рекурренттік нейрондық желілер деп атайды.

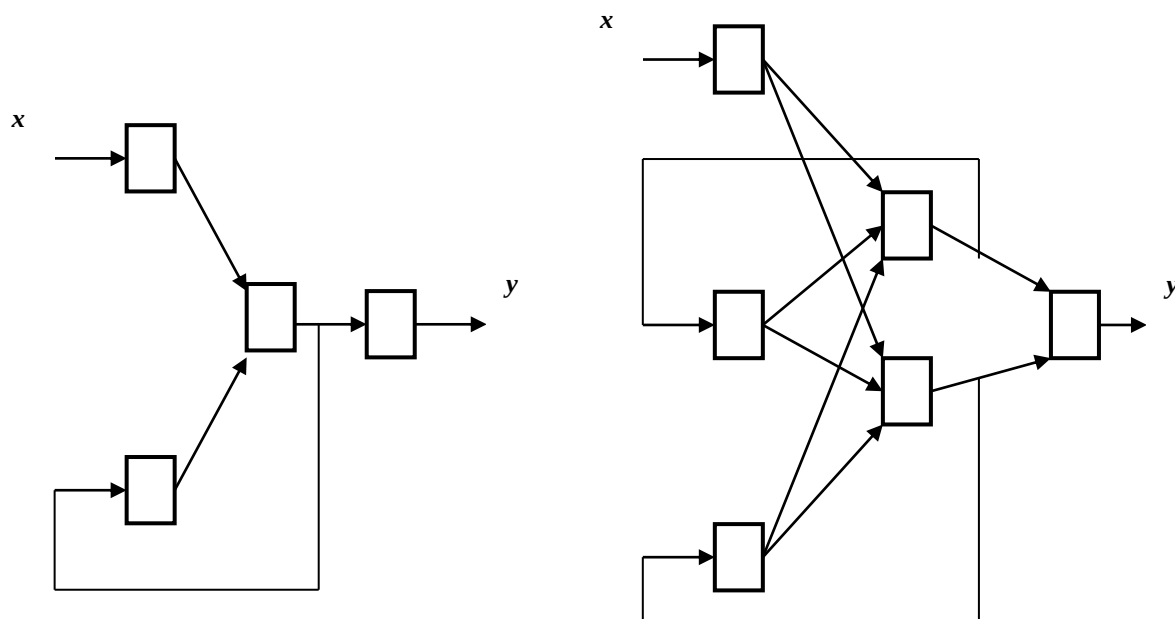
Төрт қабатты рекуррентті нейрондық желі ІІ.4.3.4-суретте көрсетілген.



ІІ.4.3.4-сурет. Төрт қабатты рекуррентті нейрондық желі

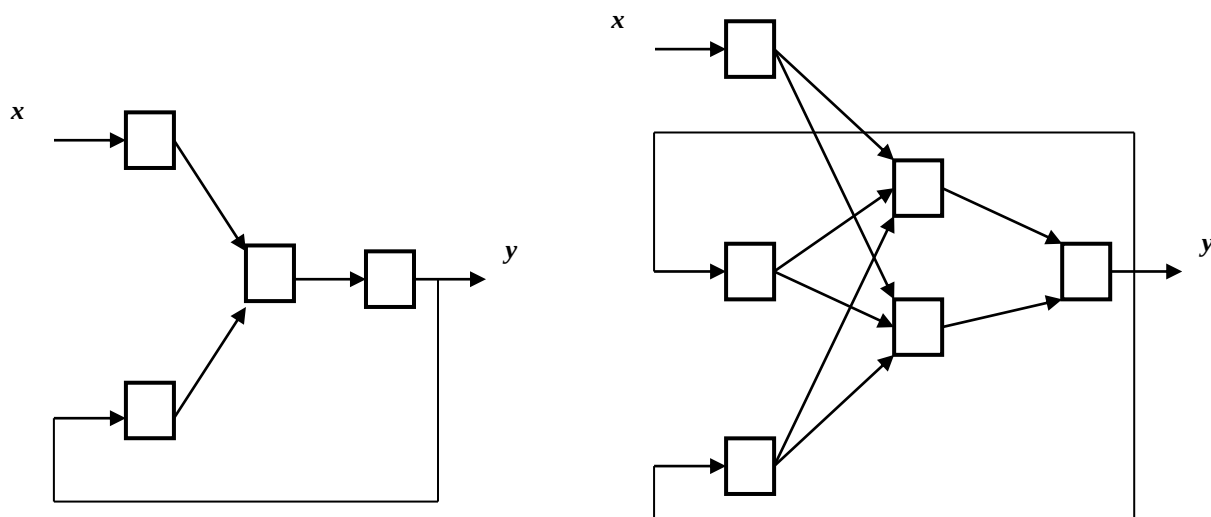
Кері байланысты нейрондық желілердің арасында Элманның және Жорданның ішінара-рекурренттік нейрондық желілер жиі қолданылады.

Элманның ішінара-рекуррентті нейрондық желісі II.4.3.5-суретте көрсетілген.



II.4.3.5-сурет. Элманның ішінара-рекуррентті нейрондық желісі

Жорданның ішінара-рекуррентті нейрондық желісі II.4.3.6-суретте көрсетілген.



II.4.3.6-сурет. Жорданның ішінара-рекуррентті нейрондық желісі

Жеке жағдайда рекуррентті желілер екі бағытты желілер болады. Мұндай желілерде қабаттар арасында кіру қабатынан шығу қабатына бағытталған және кері бағытталған байланыстар бар. Кері байланыстары бар нейрондық желілер қабатты-циклдық,

қабатты-толық байланысты, толық байланысты-қабатты болып бөлінеді.

Қабатты-циклды желілер қабаттардың сақина тәрізді тұйықталғандығымен ерекшелінеді. Соңғы қабат өзінің шығу сигналдарын біріншісіне береді. Барлық қабаттар бірдей бағытталған және кіру сигналдарын алған тәрізді шығу сигналдарын да бере алады.

Қабатты-толық байланысты желілер өзара толық байланысты желіні көрсететін қабаттардан тұрады, ал сигналдар қабаттан қабатқа және қабаттың ішіне де беріле алады, әрбір қабатта жұмыс циклі үш бөлікке бөлінеді: алдыңғы қабаттан сигналдарды қабылдау, қабаттың ішінде сигналдармен алмасу, шығу сигналын өңдеп шығару және келесі қабатқа сигналды жіберу.

Толық байланысты-қабатты желілер өзінің құрылымы бойынша қабатты-толық байланысты желіге ұқсас келеді, бірақ басқаша жұмыс атқарады. Берілген желіде қабаттың ішінде алмасу және келесіге сигнал беру фазаларға бөлінбейді, әрбір тактіде барлық қабаттағы нейрондар өз қабатындағы және келесі қабаттағы нейрондардан сигнал алады.

Желінің мүмкіндіктері желідегі нейрондар санының, олар арасындағы байланыс тығыздығының және қабат санының үлкеюімен өседі. Бірнеше синапстардың түрлерін енгізу нейрондық желі қуатының артуына әкеледі. Кері байланысты қосу желінің мүмкіндіктерінің артуымен қатар нейрондық желінің динамикалық тұрақтылығы туралы сұрақ көтереді.

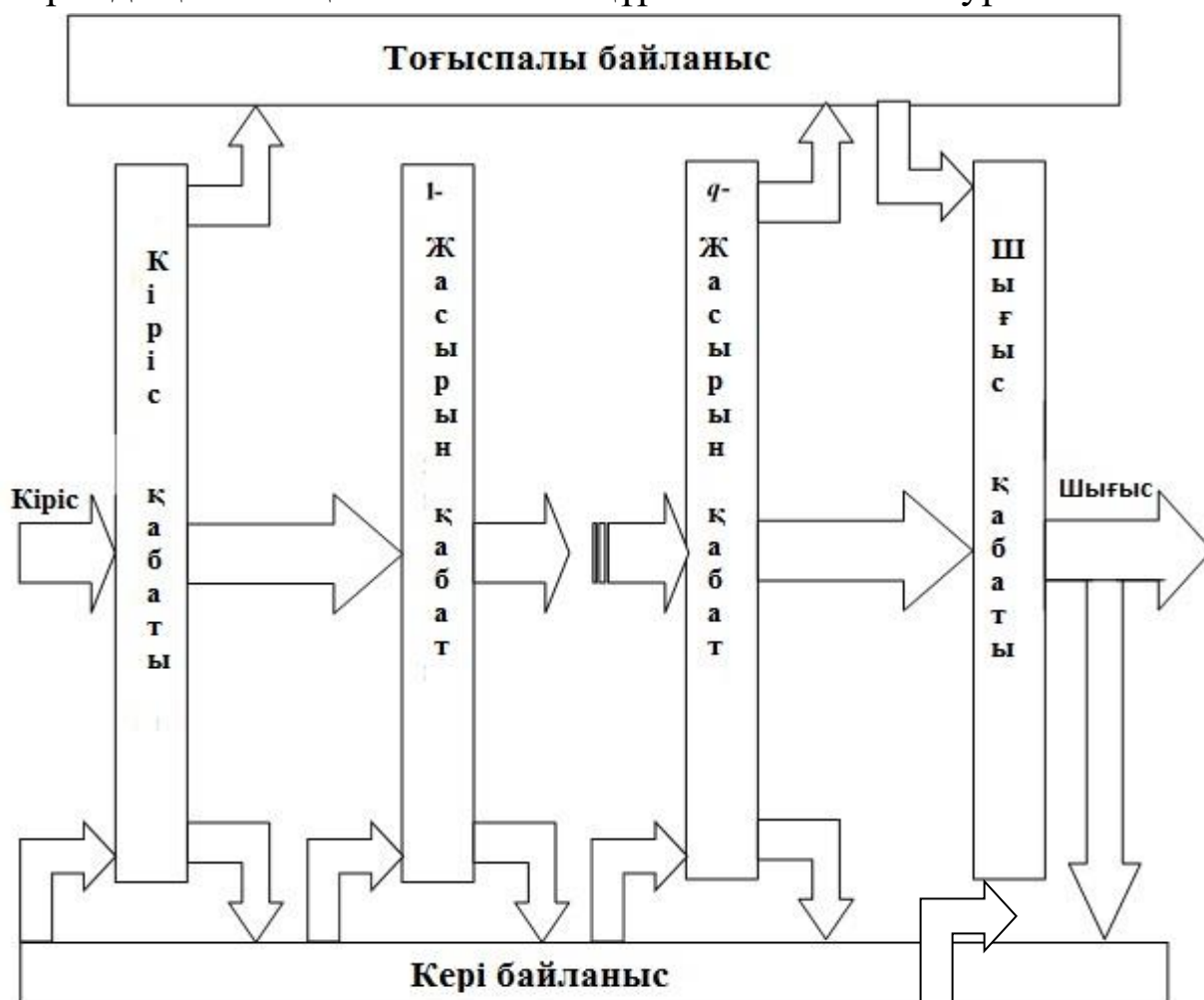
Шындығында бір жеткілікті үлкен жасырын қабаттың көмегімен еркін дәлдікті кіру деректерінен кез келген үздіксіз функцияны алуға болады. Екі қабаттың болуы кезінде тіпті үздікті функцияларды да ұсынуға мүмкіндік береді. Көп қабатты нейрондық желі күрделілігі кез келген дәрежедегі функцияны моделдей алады және бұл жерде қабаттар саны мен әрбір қабаттағы элементтер саны функцияның күрделілігін анықтайды.

Нейрожелілік моделдеудің теориялық негізделуі $[0;1]$ кесіндідегі кез келген үздіксіз көпөлшемді функцияны ақырлы санды бір өлшемді қосынды түрінде ұсынуға болады дегенді дәлелдеген А.Н.Колмогоров теоремасында көрсетілген:

$$f(x_1, x_2, \dots, x_n) = \sum_{j=1}^{2n+1} g\left(\sum_{i=1}^n \lambda_i \varphi_j(x_i)\right),$$

мұнда g және φ_j үздіксіз және бірөлшемді функциялар, $\lambda_i = \text{const}$. Осыдан, кез келген көп қабатты нейрондық желінің көмегімен барлығы екі өңдейтін қабатпен кез келген дәлдікпен бірлік кесіндідегі көпөлшемді функцияны жуықтап алуға (аппроксимациялауға) болады.

Қабаттардың ішкі байланыстары ескерілмеген көп қабатты нейрондық желінің жалпыланған құрылымы II.4.3.7-суретте.



II.4.3.7-сурет. Көп қабатты нейрондық желінің жалпыланған құрылымы

II.4.3- тапсырмалар.

1. Көп қабатты нейрондық желісінің қалыптасу тәсілін сипаттаңыз
2. Көп қабатты нейрондық желілердегі нөмірлену тәсілін сипаттаңыз.
3. Көп қабатты нейрондық желісіндегі кіріс қабаттағы, жасырын қабаттағы және шығыс қабаттағы нөмірлену тәсілін келтіріңіз.
4. Көп қабатты нейрондық желінің түріне бөлінген критерийді атаңыз.
5. Нейрожелілік моделденудің теориялық негіздемесін келтіріңіз.

II.4.3-сұрақтар.

1. Көп қабатты нейрондық желінің әр қабатында қанша нейрон бар?
2. Нақты қабатта нейрондардың саны басқа қабаттардағы нейрондар санына байланысты ма?
3. Көп қабатты нейрондық желіде бір қабаттың ішінде белсендіру функциясы қанша?
4. Қандай функцияларды көп қабатты нейрондық желінің көмегімен алуға болады?

II.4.3-тестілер.

1. Көп қабатты нейрондық желілер қалай сипатталады?

А) нейрондар қабаттарға бірігеді, қабат кіру сигналдарының бірыңғай векторы бар нейрондар жиынтығы құрайды, қабаттардағы нейрондардың саны әр түрлі болуы мүмкін және басқа қабаттағы нейрондар санына тәуелді емес;

В) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

С) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама аз болады;

Д) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды таратудан алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

Е) барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналдарын басқа нейрондарға өзіне де береді.

2. Көп қабатты нейрондық желі солдан оңға қарай нөмірленген қандай қабатынан тұрады?

A) $q > 1$;

B) $q < 1$;

C) $q = 1$;

D) $q > 0$;

E) $q < 0$.

3. Монотондық нейрондық желі неше блоктан тұрады?

A) 2;

B) 3;

C) 1;

D) 4;

E) 5.

4. Байланыс түрлері бойынша көп қабатты нейрондық желілер нешеге бөлінеді?

A) 4;

B) 2;

C) 5;

D) 3;

E) 6.

5. Айқыш байланысты нейрондық желілердің сипаттамасы қандай?

А) екінші дәрежеден басталатын, алдындағы қабаттардағы нейрондардың шығысымен, сондай-ақ тікелей желінің сыртқы кірісімен бапталатын синапстық байланыстары бар нейрондардан тұрады;

В) әр қабаттардың нейрондары өзара бір-бірімен байланыспаған және барлық байланыстар қатаң түрде кіру нейрондарынан шығу нейрондарына бағытталған;

С) жасырын қабаттарды кездейсоқ фонда суреттерді қарама-қарсы етуге мүмкіндік береді және графикалық бейнелерді тану тиімділігін арттырады;

Д) әрбір қабаты, шығу қабатынан басқа бөлінген екі? *қоздырушы және тежеуші* блоктан тұрады;

Е) әрбір нейрон кіріс ақпаратын сыртқы ортадан және басқа да нейрондардан, тіпті өзінен де алатындай (кері байланыспен) етіп ұйымдастырылған.

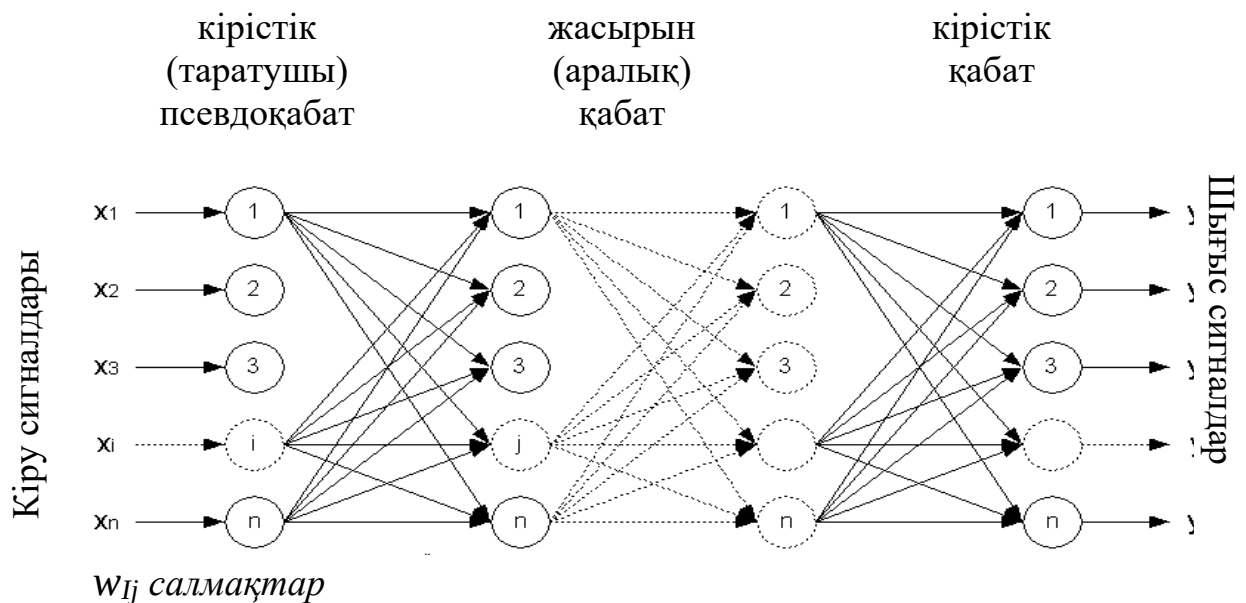
II.4.4 Көп қабатты персептрондар

Бұл тақырыпта көп қабатты персептрондар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Көп қабатты персептрондар тура байланысты (тікелей тарату сигналы бар) көп қабатты нейрондық желілер болады және нейрондардың бірнеше қабатынан тұрады: *кіру қабатынан, бірнеше жасырын қабаттардан және шығу қабатынан.*

Персептронның кіру қабаты ешқандай есептеуді орындамайды және $X = (x_1, x_2, \dots, x_n)$ векторын қабылдайды, ол X векторын жай ғана келесі қабаттарға таратып, ақпаратты қабылдауыш қызметін атқарады. Әрбір жасырын элементті персептрон деп қарастыруға болады, ол өзімен бірге үлкен кеңістікте жұмсақ межелі функцияны бейнелейді. Әрбір S -элементі бірінші аралық қабаттағы ассоциативті элементтер (A -элементтер) жиынтығымен байланысқан, ал соңғы қабаттың A -элементтері жауап қайтаратын R -элементтермен байланысқан.

Тікелей байланысты көп қабатты персептронның құрылысы II.4.4-суретте көрсетілген.



II.4.4-сурет. Көп қабатты персептронның құрылымы

Егер жасырын қабаттар екіден көп болса, онда барлық жасырын қабаттарда нейрондар саны бірдей болады. Жасырын

элементтері көп болған жағдайда әртүрлі жерлерде орналасқан әртүрлі өлшемдегі денелерді құруға мүмкіндік туады.

Кіру сигналдары кіру қабатының нейрондарымен қабылданады және бірінші жасырын қабаттағы нейрондарға беріледі, содан әрі қарай барлық нейрондық желінің шығу сигналдарын шығаратын шығыс қабаттарына дейін беріледі. k -ші қабаттың әрбір шығу сигналы тізбекті байланыстар құрай отырып, $(k+1)$ -ші қабаттың барлық нейрондарының кірісіне беріледі. Бірақ k -ші қабат $(k+q)$ еркін қабатымен қосылғанда кенеттен секіру болуы мүмкін, мұнда $k > 1$ & $q \in N$.

Таңбалаулар енгіземіз: W^k – $(k-1)$ -ші қабаттың нейрондарын k -ші қабаттың нейрондарымен жалғайтын синапстық байланыстар салмақтарының матрицасы; w_{ij}^k – $(k-1)$ -ші қабаттың I -ші нейронын k -ші қабаттың j -ші нейронымен жалғайтын байланыстардың вектор-бағаны $(y_j^{k-1} \in y^{k-1}, y_j^k \in y^k)$.

Осыған ұқсас x_j^1 – бірінші қабаттың кіріс вектор-бағаны. Сонда көп қабатты персептронның тікелей жұмыс істеуі келесі қатынастармен сипатталады:

$$s_j^k = f\left(\sum_{i=1}^{n(k-1)} w_{ij}^k y_i^{k-1}\right), \quad y_j^k = f(s_j^k + w_{0j}^k), \quad x_{ij}^{k+1} = y_j^k$$

мұнда I – кіріс нөмірі, j – қабаттағы нейрон саны, k – қабат саны;
 x_{ij}^k – k қабатындағы j -ші нейронының I -ші кіру сигналы;
 w_{ij}^k – k қабатындағы j -ші нейронының I -ші кіріс синапсының салмағы;
 s_j^k – k қабатындағы j -ші нейронының салмақталған сомасы;
 w_{0j}^k – k қабатындағы j -ші нейронының межесі.

Әрбір қабат алдыңғы қабат сигналдарының сызықты комбинацияларынан сызықты емес түрленуін есептейді. Көп қабатты нейрондық желі шығысында қабаттар санын, сигналдардың өзгеру диапазонын және нейрондардың

параметрлерін лайықты таңдаған кезде еркін көпөлшемді функцияны қалыптастыра алады. Көп қабатты нейрондық желілер функцияларды әмбебап аппроксиматорлары болады. Көп қабатты персептронды сипаттайтын жалпы формула келесі түрде анықталады:

$$f(s) = f \left(\underbrace{\sum_{i_K} w_{i_K j_K}^K \dots f \left(\sum_{i_2} w_{i_2 j_2}^2 f \left(\underbrace{\sum_{i_1} w_{i_1 j_1} + w_{j_1 2}}_{1 \text{ qabat}} \right) + w_{j_2 2} \right)}_{2 \text{ qabat}} \dots + w_{j_K K} \right)_{K \text{ qabat}}$$

Сызықтық комбинацияларды және сызықты емес түрлендірулерді кезекпен есептеу есебінен еркін көпөлшемді функциялар желі параметрін лайықты таңдаған кезде аппроксимацияға жетеді. Персептрондар бөлінетін біршама күрделі қабаттар құруға мүмкіндік береді, соның салдарында көбірек таралған.

Көп қабатты персептрондардың үш айрықша белгісі бар. Әрбір нейрон сызықты емес белсендіру функциясына ие. Бұл сызықты емес функция Розенблатт персептронында қолданылатын қатаң межелік функцияға қарағанда тегіс (барлық жерде дифференциалданатын) болады. Осы талапты қанағаттандыратын функцияның ең танымалы логистикалық функциямен анықталатын сигмоидалық түрі болады:

$$f(s) = \frac{1}{1+e^{-s}}, (1)$$

Сызықты еместіктің болуы өте маңызды рөл атқарады, себебі, кері жағдайда желінің «кіріс-шығыс» бейнеленуін қарапайым бір қабатты персептронға келтіруге болады. Онымен қоса, логистикалық функцияны қолдану биологиялық тұрғыдан

негізделген, себебі оның ішінде нағыз нейронды қайта қалпына келтіру фазасы ескеріледі.

Көп қабатты желілердің мүмкіндіктері бұрыннан белгілі болғанына қарамастан, көп жылдар шамасында олардың салмағын баптаудың теориялық негізделген алгоритмі мүлдем болмаған. Ары қарай біз көп қабатты оқытатын алгоритмімен анығырақ танысамыз, бірақ қазір осы кезеңде мәселені жеткілікті түсіну және зерттеулердің нақты белгілі нәтижелерге алып келгенін түсінуіміз керек.

Тура байланысты көп қабатты нейрондық желілер мұғаліммен оқытылатын дельта-ережелер көмегімен оқытылады және ол келесі түрде жүзеге асады:

$$w_{ij}(t+1) = w_{ij}(t) + \eta x_i(d_j - y_j),$$

мұнда η – параметр (оқыту қадамы); d – шығу элементінің эталонды мәні (талап етілетін).

Сөйтіп, байланыстар күшінің өзгеруі ($d - y$) шығу сигналдарының қатесіне және x кіру элементінің белсендірілу дәрежесімен сәйкес өтеді. Қателіктің кері таратылуы (Back – propagation) деп аталатын дельта-ережені жалпылау кез-келген қабат саны бар желілерге қолданыла береді. Мұндай жағдайда желілердің оқытылуы келесі қадамдардан тұрады:

1. Кезекті оқытылатын (x, d) жұпты таңдау. Желі кірісіне кіру векторын беру.
2. y желісінің шығысын есептеу.
3. Талап етілетін шығыс пен желінің шығысы арасындағы айырмашылықты (қатені) анықтау.
4. Қатені неғұрлым төмендетіндей етіп желі салмақтарына түзетулер жасау.
5. Қателік қажетті дәрежеге жеткенге дейін, әрбір оқытушы жұп үшін 1-ден 4-ке дейінгі қадамдарды қайталау.
6. Желілердің жұмыс істеу қателері әдетте келесі түрде анықталады.

$$E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j - y_j)^2,$$

мұнда $y_j = y_j^k$ – желі шығысы.

Бұл қатені өзгерту үшін желінің салмағын мына ереже бойынша өзгерту керек:

$$w^k(t+1) = w^k(t) - \eta \cdot \frac{dE}{dw^k},$$

Шығу қабаты үшін j -ші нейронның δ_j қатесін анықтау оңай:

$$\delta_j = (d_j - y_j) \cdot f'(s_j) = (d_j - y_j) \cdot C \cdot f(s_j) \cdot (1 - f(s_j))$$

және бір қабатты желілерді оқыту кезінде қолданылатын мадақтау-жазалау жүйесін еске түсіреді. Әрбір алдыңғы қабаттың қателері келесі қабаттардың қателері арқылы рекурсивті түрде анықталады:

$$\delta_j = \left(\sum_{i=1}^m \delta_i w_{ij} \right) \cdot f'(s_j),$$

яғни әрбір j -ші нейрон үшін келесі қабаттың қателері сәйкес салмақтарынан өтіп, оған қайтадан таралады.

Бұл кері тарату механизмі оңтайландырудың көптеген градиенттік әдістері үшін дәстүрлі болатын ең қысқа түсу векторын бағалау, қадам шамаларын еңістің құлдылығына пропорционал түрде өзгерту және т.б. процедуралармен толықтырды.

Егер тікелей жұмыс істеу кезінде кіру сигналы желі бойынша кіру қабатынан шығу қабатына қарай таралса, онда салмақтарды баптау кезінде желі қателігі шығу қабатынан кіру қабатына қарай таралады.

II.4.4-тапсырмалар.

1. Көп қабатты персептронның жұмыс жасауын сипаттаңыз.
2. Көп қабатты персептронның құрылымын сипаттаңыз
3. Көп қабатты персептронның тікелей жұмыс жасауы сипатталатын қатынасты жазыңыз.
4. Көп қабатты персептронның жалпы формуласын жазыңыз

5. Көп қабатты персептронды белсендіру функциясын жазыңыз.
6. Көп қабатты персептронды оқыту қатесін есептеу формуласын жазыңыз.

II.4.4-сұрақтар.

1. Көп қабатты персептрондар қандай қабаттардан тұрады?
2. Егер персептрондағы қабаттар саны екі есеге көп болса, нейрондар саны қалай өзгереді?
3. Көп қабатты персептрондағы k -қабатының әрбір шығу сигналы қайда жіберіледі?
4. Көпөлшемді персептрон қалай зерттеледі?
5. Көп қабатты персептронда қателіктер қалай таралады?

II.4.4-тестілер

1. Көп қабатты персептрондар қандай қабаттардан тұрады?
 - A) кіру қабатынан, бірнеше жасырын қабаттардан және шығу қабатынан;
 - B) бірнеше кіріс қабаттардан, бір жасырын қабаттан және шығу қабатынан;
 - C) кіру қабатынан, бірнеше жасырын қабаттардан және бірнеше шығыс қабаттардан;
 - D) кіру қабатынан, жасырын қабаттан және шығу қабатынан;
 - E) кіру қабатынан және шығу қабатынан.
2. Қателіктің кері таратылуы (Back – propagation) деп аталатын дельта-ережені жалпылау кез-келген қабат саны бар желілерге қолданылса, мұндай жағдайда желілердің оқылуы неше қадамнан тұрады?
 - A) 6;
 - B) 5;
 - C) 2;
 - D) 3;
 - E) 4.
3. Желілердің жұмыс істеу қателері әдетте қандай формуламен анықталады?

$$A) E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j - y_j)^2;$$

$$B) E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j - y_j);$$

$$C) E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j + y_j)^2;$$

$$D) E = \sum_{j=1}^{n(k)} (d_j - y_j)^2;$$

$$E) E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j)^2.$$

4. Желілердің жұмыс істеу қателері әдетте $E = \frac{1}{2} \cdot \sum_{j=1}^{n(k)} (d_j - y_j)^2$ өрнегімен анықталады. Мұнда $y_j = y_j^k$ нені білдіреді?

- A) желі шығысы;
- B) желі кірісі;
- C) желі салмағын;
- D) шығу қабаты үшін нейрон қатесін;
- E) коэффициенттің мәні.

5. Көп қабатты персептрондардың неше айрықша белгісі бар?

- A) 3;
- B) 2;
- C) 4;
- D) 5;
- E) 6.

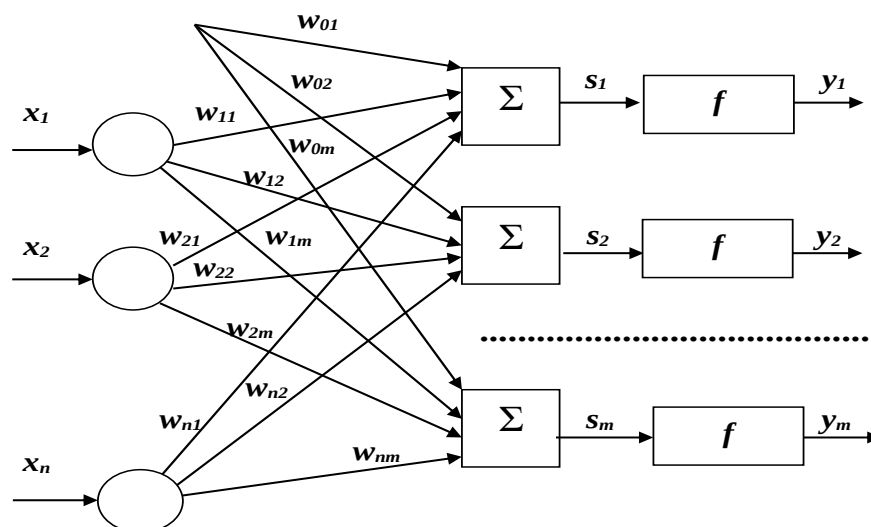
II.4.5. Кохоненнің нейрондық желілері

Бұл тақырыпта Кохоненнің нейрондық желілері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Кохоненнің нейрондық желілері нейрондық желілердің жеке сыныбын көрсетеді және өздігінен ұйымдастырылатын нейрондық желілерге жатады: *векторлық кванттау желісі, өздігінен ұйымдастырылатын картасы.*

Кохоненнің базалық нейрондық желісі бір қабатты желі болады, оның әрбір нейроны n -өлшемді кіру векторының барлық компоненттерімен байланысқан. Кіру векторы – бұл кластерлеуге жататын объектілердің бірінің сипаттамасы. m нейрондар саны желіні анықтайтын сыныптар санына сәйкес келеді. Кохоненнің желісінде нейрондар түрінде салмақталған сызықтық қосқыштар қолданылады. Желінің $X=(x_1, x_2, \dots, x_n)$ кіру векторы келесі түрде болады. Әрбір j -ші нейрон n -өлшемді кеңістікте $W_j=(w_{1j}, w_{2j}, \dots, w_{nj})$ синапстық байланыстар салмақтары векторымен сипатталады.

Кохоненнің базалық желісінің құрылымы II.4.5-суретте көрсетілген.



II.4.5- сурет. Кохоненнің базалық нейрондық желісінің құрылымы

Кохонен желісі m параллель жұмыс істеп тұрған сызықтық элементтерден тұрады. Барлығы бірдей n кіріс санына ие және өз кірістеріне бірдей $X=(x_1, x_2, \dots, x_n)$ кіріс сигналдар векторын алады. Ал j -ші сызықтық элементтің шығу сигналы былай анықталады:

$$y_j = w_{0j} + \sum_{i=1}^n w_{ij} x_i,$$

мұнда w_{0j} – межелік коэффициент; w_{ij} – j -ші нейронның I -ші кірісінің салмақтық коэффициенті; $I=1, \dots, n$ – кіру сигналының нөмірі; $j=1, m$ – сызықтық элемент нөмірі.

Кохонен желілерінде мұғалімсіз оқыту қолданылады. Оқыту үшін желінің *бәсекелестік механизмдері* қолданылады. Кірісіне X векторының мәндерін берген кезде, кіру векторынан ең аз дәрежеде ерекшеленетін салмақтар векторы жеңіске жетеді. Жеңімпаз-нейрон үшін келсі қатынас орындалады

$$d(x, W_j) = \min_{1 \leq j \leq m} d(x, W_j),$$

мұнда j – жеңімпаз-нейрон нөмірі, m – нейрондар саны, $d(x, W_j)$ – x және W_j векторлар арасындағы қашықтық.

Қашықтықты өлшеу ретінде W_j нүктесіне дейінгі $\rho_j(x)$ евклидік өлшемі қолданылады

$$\rho_j(x) = d(x, W_j) = \|x - W_j\|^2 = \|x\|^2 - 2 \sum_{j=1}^m x_j \cdot w_{ij} + \|W_j\|^2$$

Бұл жерде $\|y\|$ – вектордың Евклидік ұзындығы: $\|y\|^2 = \sum_{j=1}^m y_j^2$.

Бірініші қосылғыш барлық нейрондарға бірдей, сондықтан жақын арадағы нүктені табу үшін ол қажет болмайды.

Сигналдар сызықтық элементтердің қабаттарынан өткеннен кейін «жеңімпаз бәрін алады» ережесі бойынша өңделуге жіберіледі. Шығу сигналдары арасында ең үлкені ізделеді, оның нөмірі анықталады. Шығысында табылған нөмірмен сигнал бірге

тең, ал қалғандары–нөлге тең. Егер максимум бірмезгілде бірнешеуі үшін жетсе, онда:

- не барлық сигналдар сәйкесінше бірге тең мән қабылдайды;
- не тек тізімдегі бірінші сигнал (келісім бойынша) бірге тең мән қабылдайды.

Есеп сызықтық функция мәндерінен ең үлкен мәннің нөмірін іздеуге алып келеді, яғни:

$$j_{\max} = \arg \max_j \{y_j\} = \arg \max_j \left\{ \sum_{j=1}^m x_j \cdot w_{ij} - \frac{1}{2} \|W_j\|^2 \right\}.$$

Сөйтіп, $W_j = (w_{1j}, w_{2j}, \dots, w_{nj})$ нүктелер координатасы Кохонен сызықтық қабатының салмақтарымен, сонымен қоса межелік коэффициент мәні $w_{0j} = \frac{1}{2} \|W_j\|^2$ болады.

Нейрон-жеңімпаздың айналасында шеңберлену (*neighborhood*) немесе оқыту радиусы (*radius of learning*) қалыптасады. Оқыту радиусы берілген итерацияда нейрон-жеңімпаздан басқа қанша нейрон оқытуға қатысатынын анықтайды. Берілген жағдайдағы радиус деп нейрондардың салмақтар векторларының кеңістігіндегі арақашықтық қарастырылады. Яғни, салмақ векторынан нейрон-жеңімпаз салмақ векторына дейінгі қашықтықтағы кез келген нейрон, оқыту радиусынан кіші болса, берілген итерация кезінде салмақ түзетілуіне қатыса алады. Берілген итерацияда оқыту радиусы максималды және итерация санының артуымен азаяды, сөйтіп, оқытудың соңында өз салмағын тек нейрон-жеңімпаз жөңдейді.

Кохонен ережесі бойынша нейрон-жеңімпаздың және оқыту радиусында жататын барлық нейрондардың салмақтары оқытуға ұшырайды. Мұнда Кохонен қабаттары деп аталатын реттелмеген нейрондары бар желілер және өздігінен ұйымдастырылатын Кохонен картасы (*Self-Organizing Maps – SOM*) деп аталатын реттелген нейрондармен құралған желілер болып бөлінеді.

Кохонен қабаты бейімді сызықтық қосқышлардан (сызықтық формалды нейрондардан) тұрады. Әдетте, Кохонен қабатының шығу сигналдары «Жеңімпаз бәрін алады» ережесі бойынша өңделеді: ең үлкен сигнал бірлікке айналады, ал қалғандары нөлге тең болады.

Кохонен нейрондарын кез келген кіру векторы үшін біреуі ғана жанатын электр шамдарының жиынтығы ретінде қабылдауға болады.

Векторлық кванттау желісі. X кіріс векторларының берілген жиынтығы үшін W_j кодтық векторлармен векторлық кванттау есебі, кодтау кезінде, яғни X -тегі әр векторды сәйкесінше кодтық векторға алмастыру кезінде бұрмалауды барынша азайту есебі ретінде қойылады. Кохонен желілерінің базалық нұсқасында ең кіші квадраттар әдісі пайдаланылады және D бұрмалау мына формула бойынша есептеледі:

$$D = \sum_{i=1}^n \sum_{x \in K_j} \|x - W_j\|^2,$$

мұнда $K_j - W_j$ басқаларынан қарағанда W_i -ға ($i \neq j$) жақын $x \in X$ нүктелерінен тұрады. Басқаша айтқанда, K_j кодтық векторы W_j -мен кодталатын $x \in X$ нүктесінен тұрады.

Егер X жиынтығы берілсе және ол жадта сақталса, онда Кохонен желісіне сәйкес келетін оқытудағы стандартты таңдау K -орташа тәсілі болады. Бұл ыдырату тәсілі:

- Берілген W_j кодтық векторларды (олар салмақтық векторлар желісі) таңдауда D минималдауы ретінде K_j жиынын табамыз, олар W_i -басқаларынан қарағанда W_j -ға жақын $x \in X$ нүктелерінен тұрады.

- Берілген X -ті K_j жиындарына бөлу кезінде D минимизация ретінде ең кіші квадраттар тәсілі бойынша бағалау үшін W_j

кодтық векторлардың оңтайлы позициясын табамыз, бұл қарапайым арифметикалық орташа:

$$W_j = \frac{1}{|K_j|} \sum_{x \in K_j}^n x,$$

мұнда $|K_j|$ – K_j -дағы элементтер саны.

Бұдан әрі итерациялаймыз. Бұл ыдырату әдісі қадамдардың ақырлы санында жинақталады және бұрмалаудың локалдык минимумын береді.

Егер де, мысалы, X жиынтығы алдын ала берілмесе немесе қандайда бір себеппен жадта сақталмаса, онда онлайн тәсілі кеңінен қолданылады. x кіру сигналдарының векторлары бір-бірден өңделеді. Олардың әрқайсысы үшін ең жақын $W_j(x)$ кодтық векторы («барлығын алатын» «жеңімпаз») орналасақан. Осыдан кейін кодтық вектор келесі формула бойынша есептеледі:

$$W_{j(x)}^{new} = W_{j(x)}^{old}(1 - \theta) + \theta x,$$

мұнда $\theta \in (0,1)$ – оқыту қадамы. Қалған кодтық векторлар бұл қадамда өзгеріске ұшырамайды.

Тұрақтылықты қамтамасыз ету үшін жылдамдығы өшпелі оқыту онлайн тәсілі қолданылады. Егер T – оқыту қадамының саны болса, онда $\theta = \theta(T)$. Ал $\theta(T) > 0$ функциясын $T \rightarrow \infty$ кезде

монотонды болатындай және $D = \sum_{T=1}^n \theta(T)$ қатарлары

таратылатындай етіп таңдайды, мысалы, $\theta(T) = \frac{\theta_0}{T}$.

Векторлық кванттау кластеризацияға қарағанда біршама ортақ операция болады, себебі кластерлер өзара екіге бөлінулері тиіс, алайда W_j әртүрлі кодтық векторлары үшін K_j жиынтықтары жеке кластерлерден тұруы міндетті емес. Басқаша қарағанда, ажыратушы кластерлер болған кезде векторлық кванттаулар оларды таба алады және әр түрде кодтайды.

Қосқыштардың кіріс салмақтарын баптау тәсілдері және шешілетін есептер бойынша Кохоненнің желілерінің көптеген түрлері болады. Олардың ішінде ең әйгілері:

- кластерлік талдаудың қарапайым базалық алгоритмдерімен тығыз байланысқан сигналдарды векторлық кванттау желісі (динамикалық ядро немесе K-орташа әдісі);
- мұғаліммен оқытылатын векторлық кванттау желісі (LearnIng Vector QuantIzatIon).

Кохоненнің өздігінен ұйымдастырылатын картасы (Self OrganIzIng Maps – SOM) көпөлшемді векторларды бейнелеу және кластерлеу міндетін атқаратын мұғалімсіз оқытатын жарыспалы нейрондық желіден тұрады.

Мұндай алгоритмдер мысалы, *k*-жақын орташа (*k*-means) алгоритм бола алады. SOM алгоритмінің ерекшелігі, ондағы барлық нейрондар (түйіндер, сыныптар орталығы) қандай да бір құрылыммен (әдетте екіөлшемді тормен) реттелген болады. Бұл ретте оқыту үдерісінде нейрон-жеңімпаз ғана емес, оның көршілері, бірақ аз дәрежеде түрленеді. Осының есебінен SOM-ды көпөлшемді кеңістікті өлшемі аса аз (көп жағдайда, екіөлшемді) кеңістікке проекциялау әдістерінің бірі деп санауға болады, сондай-ақ моделдеу, болжау және т.б. есептерін шешу үшін қолданылады.

Бұл алгоритмді қолдану кезінде бастапқы кеңістікте ұқсас векторлар бір-бірінің жанында орналасады және алынған картада бейнеленеді.

Кохонен картасында сигнал бірден барлық нейрондарға түседі, синапстарға сәйкес салмақтар түйіндердің орналасу координаттары ретінде қарастырылады, және шығыс сигнал «жеңімпаз бәрін алады» принципі бойынша қалыптасады – яғни нөлге тең емес шығу сигналының, кірісіне берілетін объектіге жақын (синапстар салмағы мағынасында) нейрондары болады. Оқыту үдерісінде синапстық байланыстардың салмақтары тордың түйіндері деректердің локалдық қоюланған орындарында «орналасатындай», яғни деректер бұлтының кластерлік құрылысын сипаттайтындай

етіп бапталады, басқа жағынан, нейрондар арасындағы байланыстар белгілер кеңістігіндегі тиісті кластерлер арасындағы көршілес қатынастарға сәйкес келеді.

Мұндай карталарды көпөлшемді кеңістікте орналастырылған түйіндердің екіөлшемді торы ретінде қарастырған ыңғайлы. Бастапқыда өздігінен ұйымдастырылатын карта өзара байланыстармен біріктірілген түйіндерден тұратын торды білдіреді. Кохонен түйіндердің қосылуының екі нұсқасын – тік бұрышты және гексагоналды торды қарастырған – олардың айырмашылықтары, тік бұрышты торда әрбір түйін 4 көршісімен, ал гексагоналды – 6-жақын түйіндермен қосылған. Екі осындай торлар үшін Кохонен желісін құру үдерісі, берілген түйінге жақын жатқан көршілер таңдап алынатын орындарда ғана ерекшеленеді. Деректер кеңістігіне торларды бастапқы салу еркін түрде таңдап алынады. SOM_PAK авторлық пакетте кеңістікте түйіндерді кездейсоқ бастапқы орналастыру нұсқасы және жазықтықта түйіндерді орналастыру нұсқасы ұсынылады. Осыдан кейін түйіндер келесі алгоритмге сәйкес орын ауыстыра бастайды:

1. x деректер нүктесі кездейсоқ тәртіппен таңдап алынады.

2. Картаның x -ке жақын түйіні анықталады (BMU – Best MatchIng UnIt).

3. Бұл түйін x бағыты бойынша берілген қадамға жылжиды. Алайда, ол жалғыз ауыспайды, картаның қандай да бір аймағынан жақын тұрған түйіндерді өзімен бірге ерте кетеді. Барлық жылжитын түйіндердің арасында ең жақсы жылжитыны ол орталық – деректер нүктесіне жақын түйіндер, ал қалғандары BMU алыс болған сайын, соншалықты нашар жылжиды. Картаны баптау кезінде екі этапты қарастырады – баптаудың дөрекі кезеңі (orderIng) және нәзік кезең (fine-tunIng). Бірінші кезеңде аймақтағы үлкен мәндер таңдап алынады және түйіндер қозғалысы топтық сипатқа ие,–нәтижесінде карта «жайылады» және дөрекі түрде деректер құрылысын бейнелейді. Баптаудың нәзік кезеңінде аймақтар радиусы 1-2 түйінге тең және түйіндердің жеке орналасуы бапталады. Мұнымен қоса, ығысу шамасы уақыт өте келе

бірқалыпты өше бастайды, яғни ол әрбір моменттің басында үлкен болғанмен, соңында нөлге тең болады.

4. Алгоритм белгілі бір дәуірді қайталап отырады (қадамдар саны есепке байлынысты қатты өзгеріп отыруы мүмкін екендігі түсінікті).

Өздігінен ұйымдастырылатын Кохонен картасы деректерді бастапқы талдау және визуализациялау үшін қолданылады.

II.4.5-тапсырмалар.

1. Кохонен нейрондық желісінің құрылымын салыңыз.
2. Кохонен нейрондық желісінің j -ші сызықтық элементінің шығу сигналын есептейтін формуланы жазыңыз.
3. Кохонен нейрондық желісінің түрлерін атап шығыңыз.
4. Кохонен қабатының құрамын сипаттаңыз.
5. Кохонен қабатындағы сызықтық элементтің шығу сигналы үшін формуланы жазыңыз.
6. Кохонен қабатында параллельді әсер ететін сызықтық элементтер санын атаңыз.
7. Сызықтық функциядағы ең үлкен мәннің нөмірін іздеу формуласын жазыңыз.

II.4.5-сұрақтар.

1. Кохонен нейрондық желілері қандай типке жатады?
2. Кохонен нейрондық желісіндегі нейрондар саны неге тең?
3. Қай ереже бойынша үлкен сигнал бірлікке, ал қалғандары нөлге айналады?
4. Қандай парадигмамен Кохонен нейрондық желілері оқытылады?
5. Кохонен желісін оқыту үшін қандай механизм қолданылады?
6. Векторлық кванттау желісі мен өздігінен ұйымдастырылатын Кохонен картасының арасында қандай айырмашылық бар?

7. Кохонен желісінде шығу сигналдары қандай ереже бойынша өңделеді?

II.4.5-тестілер.

1. Кохоненнің нейрондық желісі қалай сипатталады?

А) олар өзімен бірге нейрондық желілердің жеке сыныбын көрсетеді және өздігінен ұйымдастырылатын нейрондық желілерге жатады;

В) нейрондар қабаттарға бірігеді, қабат кіру сигналдарының бірыңғай векторы бар нейрондар жиынтығы құрайды, сонымен қоса қабаттардағы нейрондардың саны әр түрлі болуы мүмкін және басқа қабаттардағы нейрондар санына тәуелді емес;

С) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама күшті болады;

Д) қарапайым есептеулерді орындай алатындығына қарамастан, бірнеше нейронды біріктіруден алынған персептрондарда есептеу қуаттылығы біршама аз болады;

Е) барлық кіру сигналдары барлық нейрондарға беріледі және әрбір нейрон өзінің шығу сигналдарын басқа нейрондарға, соның ішінде өзіне де береді.

2. Кохонен желісі қандай элементтерден тұрады?

А) m параллель жұмыс істеп тұрған желілік элементтерден;

В) n кірістерден;

С) m параллель жұмыс істеп тұрған бір ғана желілік элементтен;

Д) m параллель жұмыс істеп тұрған желілік екі элементтен;

Е) n кірістерден және m параллель жұмыс істеп тұрған бір ғана желілік элементтен.

3. Кохонен желілерінде қандай оқыту қолданылады?

А) мұғалімсіз;

В) мұғаліммен;

С) мұғалімсіз немесе мұғаліммен;

Д) тек бір ғана мұғаліммен;

Е) бірнеше мұғалімдермен.

4. Кохонен қабаты неден тұрады?

A) бейімді сызықтық қосқышлардан («сызықтық формалды нейрондардан»);

B) кіру қабатынан, бірнеше жасырын қабаттардан және шығу қабатынан;

C) бірнеше жасырын қабаттардан;

D) кіру қабатынан және шығу қабатынан;

E) бейімді сызықтық қосқышлардан, кіру қабатынан және шығу қабатынан.

5. Кохонен түйіндердің қосылуының неше нұсқасын қарастырған?

A) 2;

B) 3;

C) 4;

D) 5;

E) 6.

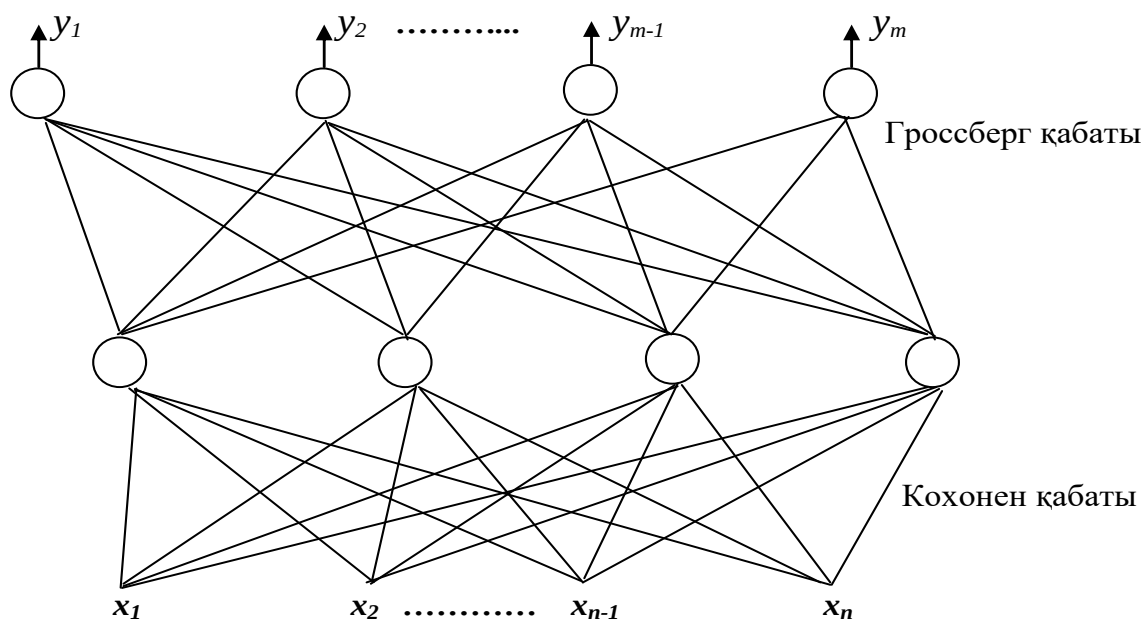
II.4.6 Аралас нейрондық желілер

Бұл тақырыпта аралас нейрондық желілер қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған.

Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Аралас нейрондық желілер әртүрлі типтегі нейрондық желілердің мүмкіндіктерін біріктіреді. Практикада кеңінен пайдаланылатын аралас желілердің бірі болып қарама-қарсы тарату желісі болады (Counter Propagation Network), оның бірінші қабаты Кохонен желісі, ал екінші- Гроссберг желісі.

Қарама-қарсы тарату желісінің құрылымы II.4.6.1-суретте көрсетілген.

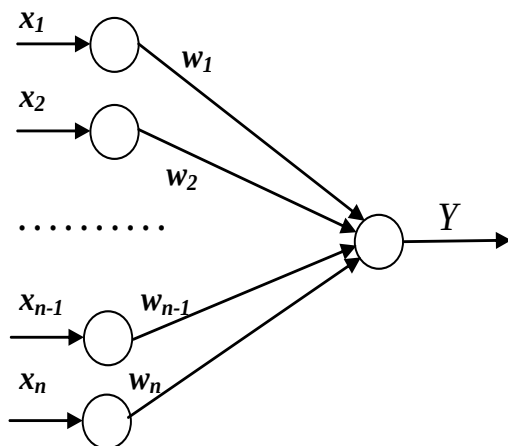


II.4.6.1-сурет. Қарама-қарсы тарату желісінің құрылымы.

Гроссберг қабаты Гроссберг кіру және шығу жұлдыздарының конфигурациясын білдіреді.

Кіру жұлдызы нысанындағы нейронның $W=(w_1, w_2, \dots, w_n)$ салмақтық коэффициенттері және кірістердің салмақталған қосындысы болатын бір шығыс Y сәйкес келетін n кірісі болады.

Гроссберг кіру жұлдызының құрылымы келесі II.4.6.2-суретте көрсетілген.



II.4.6.2-сурет. Кіріс жұлдызының құрылымы.

Салмақтық коэффициенттерді есептеу келесі формуламен есептеледі:

$$W_i(t+1) = W_i(t) + \mu(X_i - W_i(t)),$$

мұнда $W_i(t)$ – t -ші оқыту тактісіндегі I -ші кіру жұлдызының салмақтық векторы; X_i – кіру векторы; μ – оқыту үдерісінде біртіндеп азая беретін және бастапқы мәні $0,1 \dots 0,2$ шамасындағы оқыту жылдамдығы.

Гроссбергтің шығу жұлдызы қарама-қарсы функция қызметін атқарады – кіріске сигнал түскенде нақты вектор беріледі. Мұндай типті нейрон бір кірістен және $W = (w_1, w_2, \dots, w_m)$ салмағы бар m шығыстан тұрады. Салмақтар мына формуламен бапталады:

$$W_i(t+1) = W_i(t) + \eta(Y_i - W_i(t)),$$

мұнда $W_i(t)$ – t -ші оқыту тактісіндегі I -ші шығу жұлдызының салмақтық векторы, Y_i – кіру векторы, η – оқыту жылдамдығы.

Оқытуды бірлік шегінде η мәнімен біртіндеп нөлге жақын мәнге дейін азайта отырып бастау ұсынылады.

Қарама-қарсы тарату желісін оқыту екі фазадан тұрады:

1. Кохонен қабатын оқыту

Кохонен қабатында нейрондар әрбір оқытылу қадамында жеңімпаз нейрондарды анықтауы бар бәскелестік оқытуды пайдаланады. Өлшем ретінде эвклидік қашықтық қолданылады:

$$D=|X-W_i|=\sqrt{(x_1-w_{1i})^2+(x_2-w_{2i})^2+\dots+(x_n-w_{ni})^2},$$

$$D_j=\min_i |X-W_i| - j \text{ нөмерімен нейрон-жеңімпаз анықталады.}$$

Мұнда I – кіру векторының нөмірі, j – жеңімпаз нейронның нөмірі.

Кохонен қабатында нейрондардың шығу мәні осылайша есептеледі:

$$y_j = \begin{cases} 1, & D_j = \min_i |X - W_i| \\ 0, & \text{әйтпесе} \end{cases}$$

Келесі талаптарға сәйкес нейрон-жеңімпаздың салмақтық коэффициенттерінің жетілдірілуі мыналармен сәйкес жүргізіледі:

– дұрыс сыныптау кезінде нейрон-жеңімпаздың салмақтық коэффициенті артады:

$$W_{ij}(t+1)=W_{ij}(t)+\gamma(X_i-W_{ij}(t));$$

– қате сыныптау кезінде нейрон-жеңімпаздың салмақтық коэффициенті азаяды:

$$W_{ij}(t+1)=W_{ij}(t)-\gamma(X_i-W_{ij}(t)).$$

Шығу қабаты мұғаліммен бірге Кохонен қабатының шығу нейрондарына және V_i Гроссберг қабатындағы I -ші кіру нейрондарын желінің жақсы үлгіде $f(X_i)$ функциясын аппроксимирлеуге байланысты оңтайлы салмақ коэффициенттерін табу арқылы жүзеге асады. Бұл үшін ең алдымен қателердің қайтымды тарату алгоритмі қолданылады және төмендетілген квадраттар түріндегі тәсілмен қателіктер орындалады:

$$E_0=\frac{1}{2}\cdot[V_i-\alpha(V_i(t)-f(X_i))],$$

мұнда α – оқыту жылдамдығы: $\alpha \in [0,1]$, $f(X_i)$ – берілген кіру векторы үшін функцияның эталондық мәні.

Оқытудың қателігі тек бір вектор үшін емес, белгілі нейронға сәйкес келетін кластер құрамына кіретін векторлар қатары үшін есептеледі.

2. Гроссберг қабатын оқыту

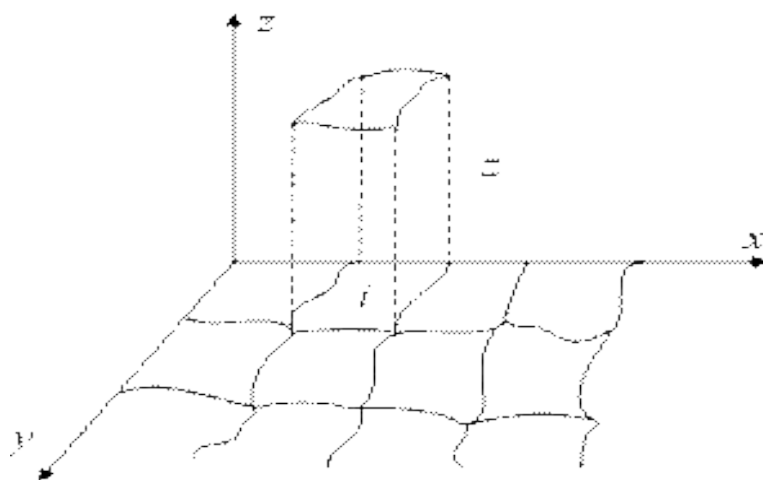
Гроссберг қабатында нейрондарды мұғалімдермен оқыту үшін және сызықтық белсендіру функциясы пайдаланылады, ал нейрондардың шығу мәні келесі формула бойынша анықталады:

$$z_i = \sum_{i=1}^m V_i y_i ,$$

мұнда m – шығу нейрондарының саны, V_i – Кохонен қабатындағы шығу нейрондарының және Гроссберг қабатындағы нейронның i -ші кірісінің салмақтық коэффициенті.

Ендеше, аралық қабаттағы әрбір i -ші кластердегі қарама-қарсы тарату нейрондық желісін V_i салмақтық коэффициентіне сәйкес қояды. Бұл функцияның аппроксимациясын орындау үшін қолданылуы да мүмкін.

Мысалы, Кохонен картасы (x y жазықтығы) кіру кеңістігінің кластерлерге бөлінуіне сәйкес келеді, ол II.4.6.3-суретіне көрсетілген. Сонда функцияның аппроксимациясы биіктікке сәйкес келеді $z: z=f(x,y)$.



II.4.6.3-сурет. Кіріс кеңістігінің кластерлерге бөлу.

Сызықты емес функцияның аппроксимациясын іске асыру үшін шығу қабатының нейрондарын белсендіру функциясы ретінде сызықты емес (мысалы, сигмоидалық) функцияны алған жөн.

Ескерту. Желілерді оқыту және жұмыс істету процедуралары параллель түрде жасалуы мүмкін. Бұл үшін алдымен желіде сақталатын үлгіден кіру үлгісінің ауытқуын рұқсат етілген дәрежеге дейін сипаттайтын ρ қырағылық коэффициентін енгіземіз. Әрбір кіру үлгілері үшін D_j қашықтығы есептелінеді, бұл кіру үлгісінің желідегі үлгіден айырмашылығын сипаттайды. Егер $D_j > \rho$ болса, онда жаңа сынып құрылады, әйтпесе, нейрондық желіні оқыту басталады.

II.4.6-тапсырмалар.

1. Қарама-қарсы таралаудың нейрондық желісіндегі бірінші қабат пен екінші қабат элементінің типін атаңыз.
2. Қарама-қарсы таралудың нейрондық желісінің құрылымын сызыңыз.
3. Кіру жұлдызының құрылымын сызыңыз.
4. Кіру жұлдызы үшін салмақ коэффициентін құруға арналған формуланы жазыңыз.
5. Шығу жұлдызы үшін салмақ коэффициентін құруға арналған формуланы жазыңыз.

II.4.6-сұрақтар.

1. Аралас нейрондық желідер қалай құралады және қандай мүмкіндіктерге ие?
2. Қарама-қарсы таралудың нейрондық желісіндегі Гроссберг қабаты деген не?
3. Гроссбергтің шығу сигналы қандай қызмет атқарады?
4. Қарама-қарсы таралудың нейрондық желісіндегі Кохонен қабаты қалай оқытылады?

5. Қарама-қарсы таралудың нейрондық желісіндегі Кохонен қабатын оқыту үшін қандай метрика қолданылады?

II.4.6-тестілер.

1. Практикада кеңінен қолданылатын аралас желілердің бірі болып қандай желі табылады?

- A) қарама-қарсы тарату желісі;
- B) параллельді тарату желісі;
- C) Кохонен желісі;
- D) Гроссберг желісі;
- E) Көп қабатты нейрондық желі.

2. Қарама-қарсы тарату нейрондық желісін оқыту неше фазадан тұрады?

- A) 2;
- B) 3;
- C) 4;
- D) 6;
- E) 1.

3. Гроссберг қабатында нейрондардың шығу мәні қандай формула бойынша анықталады?

- A) $z_i = \sum_{i=1}^m V_i y_i$;
- B) $E_0 = \frac{1}{2} \cdot [V_i - \alpha(V_i(t) - f(X_i))]$;
- C) $D_j = \min_i |X - W_i|$;
- D) $z_i = \sum_{i=1}^m V_i y_i + 1$;
- E) $z_i = \sum_{i=1}^m V_i$;

4. Кохонен қабатындағы нейрондардың шығу мәні қалай есептелінеді?

$$A) y_j = \begin{cases} 1, D_j = \min_i |X - W_i|; \\ 0, \text{ айтпесе} \end{cases}$$

$$B) y_j = \begin{cases} 1, D_j = \min_i |X + W_i|; \\ 0, \text{ айтпесе} \end{cases}$$

$$C) y_j = \begin{cases} 1, D_j = \max_i |X - W_i|; \\ 0, \text{ айтпесе} \end{cases}$$

$$D) y_j = \begin{cases} 1, D_j = \min_i |X - W_i|; \\ 0, j > 1 \end{cases}$$

$$E) y_j = \begin{cases} 1, D_j = \min_i |X - W_i|; \\ 0, j < 1 \end{cases}$$

5. Гроссберг қабаты нені білдіреді?

A) Гроссбергтің кіру және шығу жұлдыздарының конфигурациясын;

B) Гроссбергтің кіру және шығу жұлдыздарының қосындысын;

C) Кірістердің салмақтаған қосындысын;

D) t -ші оқыту тактісіндегі I -ші кіру жұлдызының салмақтық векторын;

E) берілген кіру векторы үшін функцияның эталондық мәнін.

III. НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ОҚЫТУ

III.1. Нейрондық желілерді оқытудың парадигмалары

III.1.1. Нейрондық желілерді оқытудың негізгі түсініктері

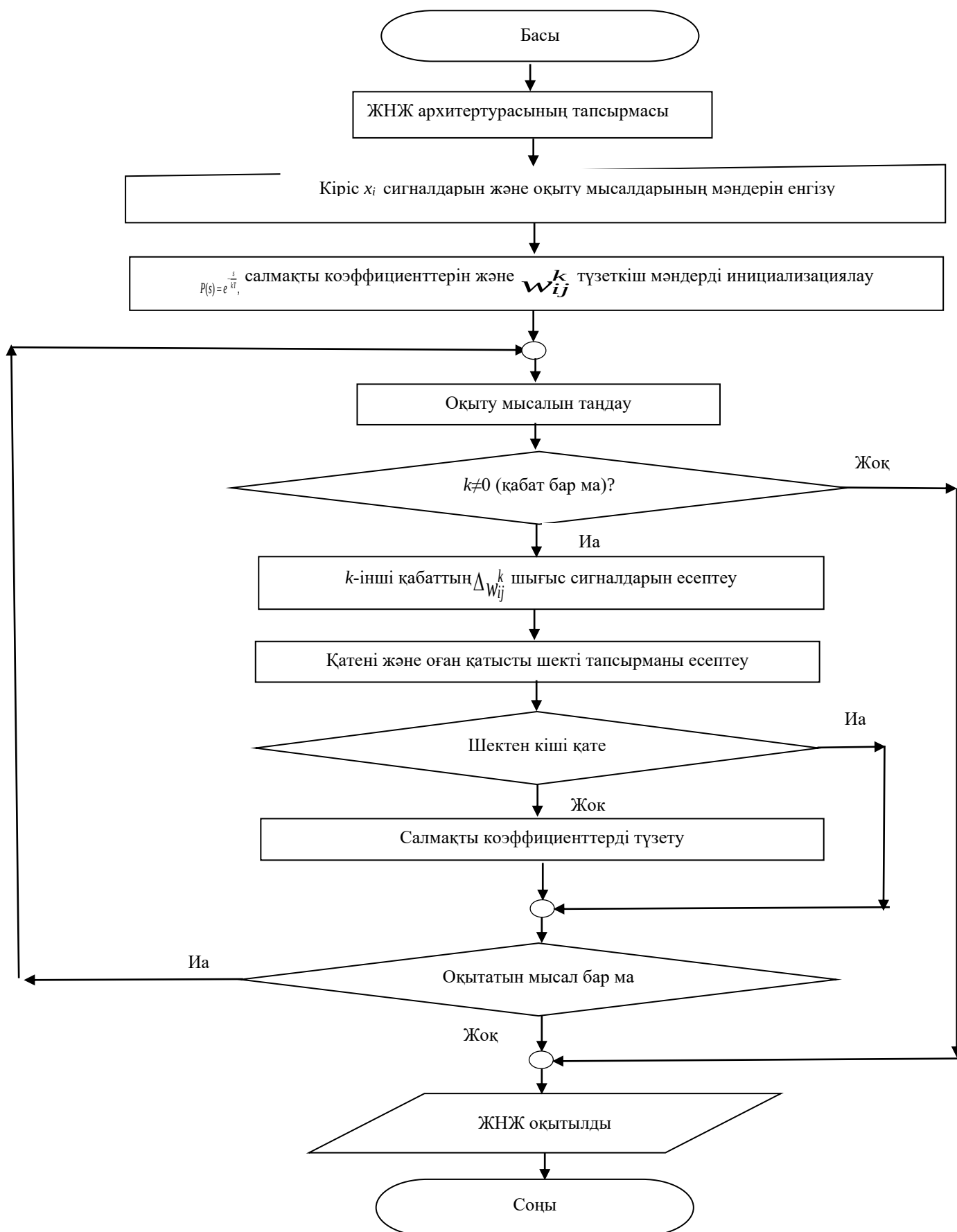
Бұл тақырыпта жасанды нейрондық желілерді оқытудың түсініктері мен парадигмалары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Өзінің ұйымдастырылуы және функционалдық қызметі бойынша бірнеше кірісі мен шығысы бар жасанды нейрондық желілер (ЖНЖ) кіру сигналдарын–сыртқы орта жайлы сенсорлы ақпарат–басқарушы шығу сигналдарына түрлендіреді. Бұл жағдайда түрлендірілетін сигналдар саны ЖНЖ-дің кіріс саны n -ге, ал шығу сигналдары m шығыстар санына сәйкес келеді. Барлық мүмкін n өлшемді *кіріс* векторлар жиынтығы *кіріс белгілік кеңістік* деп аталатын X векторлық кеңістігін құрайды. Сәйкесінше, шығыс векторлары осылайша, Y *шығыс белгілік кеңістікті* құрайды. Енді ЖНЖ-н аргументі кіріс белгілік кеңістікке, ал мәні шығыс белгілік кеңістікке тиісті болатын көпөлшемді $F: X \rightarrow Y$, *функция* ретінде қарастыруға болады. Осы функцияның түрі нейрондардың түрі мен саны және синапстық байланыстары (синапстары), сондай-ақ сипатикалық байланыстар коэффициентінің (салмағының), ығысудың және белсендіру функциясының мәндері берілетін архитектурасы бойынша анықталады.

ЖНЖ көмегімен белгілі бір есепті шешу үшін ЖНЖ-н *оқыту сатысын* өту керек, онда әрбір кіру векторы үшін оған жұп болатын және қажетті шығысты анықтайтын мақсатты вектор таңдап алынады. Олар бірге *оқытатын мысалы* деп аталады.

Әдетте, ЖНЖ есепке қатысты барлық ақпарат қамтылған көптеген *оқытатын мысалдарында* оқытылады. Сондықтан ЖНЖ оқыту сапасы оқытатын мысалдар санына және есептің осы мысалдар арқылы толықтай анықталуына байланысты. Толық жаттығу үшін бірнеше жүз мысалдар қажет деп есептелінеді.

ЖНЖ оқыту үдерісі III.1.1-суретте көрсетілген.



III.1.1-сурет. ЖНЖ оқыту үдерісі

Оқыту үдерісін бастамай тұрып, оқытылатын ЖНЖ-н құру қажет, оның көмегімен шешілетін есептерді келесі түрде анықтап алу керек:

- бірінші кезеңде оның құрылымы, қабаттар саны және нейрондардың саны көрсетілетін ЖНЖ-нің архитектурасы анықталады;

- екінші кезеңде күтілетін нәтижеге ең аз қатемен әсер ететін кіріс мәлеметтерін іріктеу жүзеге асырылады;

- үшінші кезеңде ақпаратты ұсыну тәсілі таңдалады, мәселенің типі мен сипатын ескере отырып, бастапқы деректерді түрлендіру орындалады және ЖНЖ-нің шығу айнымалылары анықталады;

- төртінші кезеңде ЖНЖ тудырылады және берілген мысалдарда конструктивті және деструктивті әдіс негізінде оқыту жүзеге асырылады.

ЖНЖ оқыту үдерісінде шығыс кезеңінде қажет нәтижеге жету үшін архитектурасы беріледі, кіріс векторларының мәні енгізіледі, оқытатын мысалы мен осы ЖНЖ үшін барлық салмақтардың оңтайлы мәндері таңдалынады. Бұл жағдайда ЖНЖ *оқытылуға қабілетті*, яғни *оқытылымдық қасиетіне* ие шешу кезінде алгоритмделмейтін, проблемалық болып табылатын немесе көп еңбекті қажет ететін есептер үшін ЖНЖ таптырмас құрал болады.

Енді оқытылуы *оқытушы деректердің жиынында*, ол *жұмысының дұрыстығын тексерілуі*—оқыту кезінде қолданылмайтын *тестілік деректер жиынында* жүргізілетін ЖНЖ-ні оқыту технологиясын сипаттауға болады. Оқыту технологиясы 2 кезеңнен тұрады:

- 1) ЖНЖ-н оқыту кезеңі, онда ЖНЖ-не оқытуды тоқтату шарты орындалғанша оқытушы жұптарының (мысалдардың) жиыны ұсынылуы:

- қатені есептеу функциясы берілгеннен кіші болғанда (яғни ЖНЖ жұмыс істеуі синапстық байланыстардың салмақтарын итерациялық баптау барысында жақсарады) немесе итерацияның нақты бір санында өзгеру тоқтағанда (ЖНЖ-н оқыту аяқталады);

– итерацияның берілген саны аяқталғаны бойынша (әдетте бұл сан 10^3 және 10^8 аралығында жатады);

2) ЖНЖ жұмысының дұрыстығын тексеру кезеңі, онда ЖНЖ-не көптеген тесттер ұсынылады және мұнымен қоса, егер жалпылау қателігі берілген жалпылау қатесінен артық болса, онда келесі амалдардың бірі орындалады:

- не итерация қадамын азайту (санын арттыру);
- не оқытатын мысалдарының санын арттыру;
- не ЖНЖ архитектурасын жетілдіру (құрамын және құрылысын өзгерту).

Оқытуды бастамас бұрын барлық салмақтарға кездейсоқ түрде таңдап алынған бастапқы кішігірім мәндері берілуі керек. Бұл ЖНЖ-нде салмақтардың үлкен мәндерімен қанығуы болмайтындығын, сондай-ақ басқа да патологиялық жағдайлар қатарының алдын алуға кепілдік береді. Мысалы, барлық салмақтарға бастапқы бірдей мәндерді беретін болсақ, ал қажетті жұмыс істеу үшін тең емес мәндері қажет болған жағдайда, ЖНЖ оқытылмай қалуы мүмкін.

Программалық жүзеге асыруда оқыту үдерісіндегі қателіктер шамасы (барлық шығыс бойынша қателіктер квадраттарының қосындысы) біртіндеп кемитініне көз жеткізуге болады. Қателік шамасы нөлге немесе ең аз мәнге жеткенде оқытылу тоқтатылып, ЖНЖ тануға дайын болады.

Оқыту үдерісінің қаншалықты сапалы болуынан, ЖНЖ-нің жұмыс істеу барысында оның алдына қойылған есептерді дұрыс шешу қабілеттілігі тәуелді болады.

Салмақтар мен межелік деңгейлер кездейсоқ мәндермен басталады. Осылайша жасалған ЖНЖ шешілетін есепке мүлдем жеткіліксіз және шығысында тек шуды ғана шығарады. Егер кішігірім оқытатын мысалдар жиынын қолданатын болсақ, онда ЖНЖ-н оқытуда оқытатын мысалдарды (артықоқытуға) тым жақын еліктеп, деректер бейненің құрылымын емес, ондағы шуды қабылдайтын болады.

Сондықтан оқытудың басында қателік өте үлкен және осыған орай параметрлерді көп түзету маңызды болады. Оқытудың соңына қарай қателік мәні кемуі түседі, сәйкесінше түзетулер кіші болуы керек. Параметрлер бойынша қадамдар ұзындығын өзгерту үшін *оқыту кестесін* (learning schedule) қолданады. t оқыту уақытынан тәуелді $\eta(t)$ оқыту жылдамдығын таңдаймыз. Әдетте жылдамдық уақыт өскен сайын монотонды кемиді. Алгоритмнің жинақталуы

$$\text{үшін қажет: } \eta(t) \xrightarrow{t \rightarrow \infty} 0, \quad \int_0^{\infty} \eta(t) dt = +\infty.$$

Оқыту кестесі бар алгоритмдер кестесі жоққа қарағанда тез жинақталады, себебі басында үлкен түзетулер қолданып, оқытудың соңында параметрлерді нақты баптау есебінде дәл нәтижелер береді.

ЖНЖ-нің оқытатын мысалдар жиынында үйрену ғана емес, жаңа деректерде жақсы нәтижелерді көрсету қабілеттілігі *экстраполяция* (жалпылау) деп аталады. Деректерді жалпылау сапасын көптеген тесттерден есептелінген қателік ауқымын бақылау арқылы анықтауға болады. Бұл қателік берілгеннен кіші болуы керек. Егер бірнеше оқыту итерацияларынан кейін, жалпылау қателігі алғашында кеміп, кейіннен қайтадан арта бастағанда, қателік нөлге дейін төмендесе, бұл артық оқытудың салдары және жалпылау қателігі өскен жағдайда оқытуды тоқтату қажет. Оқытуды аяқтаған соң ЖНЖ жұмысын үшінші жиында – *деректерді айқындау* жиынында тексеру керек.

Оқыту үдерісін анықтау үшін, біріншіден, оқыту парадигмасын анықтап, ЖНЖ-не қолжетімді ақпаратты білетін сыртқы орта моделін иелену керек. Екіншіден, оқыту тәсілдерімен берілетін баптау барысын басқарудың қандай оқыту ережелерін қолдануды, сонымен қатар ЖНЖ салмақтарын қалай түрлендіруді түсіну қажет.

Жасанды нейрондық желілердің сыйымдылығы ол онда қандай функциялар мен шешімдерді қабылдау шекаралары қалыптасатындығын, сондай-ақ оның жұмыс істеу үдерісін ұйымдастыру үшін қанша қосымша айнымалы қажет

болатындығын тануды (есте сақтауды) үйренуге қабілеті бар шығысына ұсынылатын бейнелер саны. ЖНЖ сыйымдылығы бейнелердің түпкілікті жіктелуін жүзеге асыратын талап етілген ЖНЖ шығу қабатының қуатымен тығыз байланысты.

Бейнелердің күрделілігі ЖНЖ жинақтауға жету қабілеттеріне қажетті оқытатын мысалдарының санын анықтайды. Өте аз мысалдарда ЖНЖ «аса оқытылуды» тудыруы мүмкін, ол оқытатын таңдамалар мысалдарында жақсы, ал жаңа (тест түріндегі) статистикалық тарату мысалдарында нашар жұмыс істейді және оқыту қабілетін жоғалтады. Есептеу күрделілігі есептің дұрыстау шешімін алу үшін ЖНЖ-н қанша қадам санында оқытылатынын анықтайды. ЖНЖ оқытылуының есептеу күрделілігі бірнеше факторларға тәуелді: нейрон моделдері және нейрон сандары, ЖНЖ моделі мен қабат саны, белсендіру функциясы, оқыту тәсілдері. ЖНЖ-н оқыту әртүрлі парадигмаларға және әдістерге сәйкес әртүрлі алгоритмдермен орындалуы мүмкін.

ЖНЖ-н оқыту парадигмасы қажетті ақпараттың қолжетімділігімен айқындалады. Оқытудың үш негізгі парадигмасы бар: *мұғаліммен оқыту (бақыланатын оқыту)*; *мұғалімсіз оқыту (бақылаусыз оқыту – өз бетінше оқып-үйрену)*; *аралас оқыту*.

Жасанды интеллектті жасау табиғи бейнелерді көшіру жолында дамып жатқандықтан, мұғаліммен оқыту парадигмасы табиғи немесе толығымен жасанды ма деген тақырыпқа дау әлі тоқтамады. Алайда, адам өмірінде сыртқы әсер етулерді үйлестіретін мұғалімдердің аз емес екендігін жоққа шығаруға болмайды. Әйтседе, осы екі концепцияны жақтаушыларының таласы немен бітсе де, олардың екеуі де жойылмауға құқылы.

ЖНЖ оқыту әдістері салмақтарды баптауға арналған оқыту ережесі анықталатын алгоритмдерді құруды қарастырады. Оқытудың келесі әдістері бар: *қателерді кері тарату әдісі*, *Больцманның оқыту әдісі*, *Хеббтің оқыту әдісі*, *Кохоненнің оқыту әдісі*, *жарыстыру арқылы оқыту әдісі*, *қателерді түзету әдісі*.

III.1.1-мысалдар.

1. Екі қабатты ЖНЖ үшін C_d сыйымдылығы былай болады:

$$\frac{N_w}{N_y} < C_d < \frac{N_w}{N_y} \log \left(\frac{N_y}{N_y} \right),$$

мұнда N_w – салмақтардың саны, N_y – шығу қабатындағы нейрондардың саны.

2. Кіріс бейнелерінің жиынын екі сынып бойынша бөлу үшін бір шығу нейроны жеткілікті. Бұл ретте әрбір 0 немесе 1 логикалық деңгей жеке сыныпты бейнелейтін болады. Межелі белсендіру функциясы бар екі шығу нейроны арқылы төрт сыныпты кодтауға болады. Сыныптау нақтылығын арттыру үшін мүмкіндігінше, шығу қабатында әрбір сыныпқа бір нейронды немесе әрбірі өзінің нақтылық дәрежесімен сыныпқа бейненің тиістілігін анықтауға үйрететін бірнеше нейронды бөлу жолымен артықты енгізу керек.

3. ЖНЖ көмегімен 26 латын әріптерін тану кезінде, бізге 26 оқыту жұбы болуы керек, оның әрқайсысының кіріс бөлігі әріптердің екілік бейнесінен тұратын болады. Мысалы, "А" –латын алфавитінің бірінші әрпін тану үшін кіру векторының мәні мына түрде болады: $X^1 = (010001111)$.

III.1.1-тапсырмалар.

1. ЖНЖ оқыту үдерісін сипаттаңыз;
2. Оқытылатын ЖНЖ құру кезеңдерін сипаттаңыз;
3. Салмақ саны 16-ға тең және шығу қабатындағы шығу нейрондарының саны 4-ке тең екі қабатты ЖНЖ үшін сыйымдылық мәндерінің аралығын анықтаңыз.

III.1.1-сұрақтар.

1. Қойылған есептерді ЖНЖ-нің дұрыс шешу қабілеті неге тәуелді болады?
2. ЖНЖ-ны оқыту аяқталғаннан кейін оның жұмысы қандай деректер жиынында тексеріледі?
3. ЖНЖ оқыту кезіндегі ЖНЖ сыйымдылығы, бейнелердің күрделілігі, есептеу күрделілігі деген не?

III.1.1-тестілер.

1. X векторлы кеңістігін құрайтын барлық мүмкін n өлшемді кіріс векторлар жиынтығы қалай аталады?

- A) кіріс белгілік кеңістік;
- B) шығыс белгілік кеңістік;
- C) сыныптар кеңістігі;
- D) қатарлар кеңістігі;
- E) топтар кеңістігі.

2. U векторлы кеңістігін құрайтын барлық мүмкін n өлшемді шығыс векторлар жиынтығы қалай аталады?

- A) шығыс белгілік кеңістік;
- B) кіріс белгілік кеңістік;
- C) сыныптар кеңістігі;
- D) қатарлар кеңістігі;
- E) топтар кеңістігі.

3. Қадамдар ұзындығын параметрлері бойынша өзгерту үшін не қолданылады?

- A) оқыту кестесі;
- B) есептер кестесі;
- C) сыныптар кестесі;
- D) қатарлар кестесі;
- E) топтар кестесі.

4. ЖНЖ-нің оқытатын мысалдар жиынтығын үйретумен қатар, жаңа деректерде жақсы нәтижелерді көрсету қабілеттілігі қалай аталады?

- A) экстраполяция;
- B) корреляция;
- C) дисперсия;
- D) айқындау;
- E) жинақтау.

5. ЖНЖ-н оқытудың қандай парадигмалары бар?

- A) мұғаліммен, мұғалімсіз және аралас;
- B) бағдарламашымен, өздігімен, аралас;
- C) кеңесшімен, кеңесшісіз және аралас;
- D) бақылаушымен, бақылаушысыз және аралас;
- E) білім алушымен, жеке және аралас.

III.1.2. Нейрондық желілердің мұғаліммен оқыту парадигмасы

Бұл тақырыпта жасанды нейрондық желілердің (ЖНЖ) мұғаліммен оқыту мәселесі қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Жалпы жағдайда, мұғаліммен оқытуда оқытылатын жүйенің күйін бағалау мен жетілдіру үшін оқытатын мысалдарды қолданады. ЖНЖ жағдайында, барлық оқытатын мысалдары кіру векторының (іріктеу) мәнін және қажет шығыстарды (дұрыс жауаптарды) қамтиды, ал синапстық байланыс салмақтары дұрыс жауапқа ең жақынды шығаратындай етіп жетілдіріледі, яғни әрбір оқытатын мысал мына жұптан тұрады: *кіру векторы* және *мақсатты (қалаулы шығу) векторы*.

Оқытатын мысалын берген кезде нақты шығу векторы есептеледі және тиісті мақсатты вектормен салыстырылады. Мақсатты вектормен есептелген шығу векторының айырмасы (қате) кері байланыс көмегімен ЖНЖ-нің кірістеріне беріледі және салмақ мақсатты шығысқа барынша жақын болатындай шығыстарды алу үшін өзгертіледі. Әдетте ЖНЖ кейбір оқытатын мысалдар жиынында оқытылады. Оқытатын мысалдары бірінен соң бірі ұсынылып, қателіктер есептелінеді және әрбір кіру векторы үшін салмақтар қателік барлық оқытылатын жиын бойынша тиімді деңгейге жеткенше жетілдіріледі.

Оқыту үдерісі ЖНЖ-нің барлық оқытатын мысалдар жиынында бұрын оқытылғандар қалып қоймайтындай және ЖНЖ-ға синапстық байланыстардың салмақтарының түзетілуі орындалмай тұрып барлық оқытатын мысалдары берілетіндей болуы тиіс. Салмақтардың қажетті өзгертулері барлық мысалдар жиынында есептелуі тиіс, ал бұл қосымша жадыны талап етеді. Осындай оқытушы циклдің қатарынан кейін салмақтар минималды қатеде жинақталады. Егер ЖНЖ әрдайым өзгертін сыртқы ортада болса, онда бұл әдіс пайдасыз болуы мүмкін, сөйтіп екінші рет дәл

сол вектор қайталанбауы мүмкін. Бұл жағдайда оқыту үдерісі мүлдем жинақталмауы да мүмкін.

Мейлі X кіріс белгілері кеңістігіндегі кіру векторын Y шығыс белгілері кеңістігінің шығу векторына айналдыратын $F: X \rightarrow Y$ түрлендіруді орындайтын ЖНЖ берілсін, яғни F ЖНЖ-нің берілген архитектурасына сай келетін барлық мүмкін болатын функциялар жиыны. Сонда берілген есепті ЖНЖ арқылы шешу дегеніміз $f \in F$ нақты функциясын құру және ЖНЖ-н оқыту– E қателік функциясы (x_i, y_i) , $f_i(x_i) = y_i$, $x_i \in X$, $y_i \in Y$, $i = 1, 2, \dots, K$ барлық оқыту жұптары үшін минималды болатындай етіп, синапстық байланыстар мен ығысулардың салмақтар мәнін іріктеу үшін итерациялық процедураны құру болады.

E қателік функциясы әрбір f_i функциясы үшін f функциясына жақындық дәрежесін көрсететін болады.

W күйлер кеңістігінің w күйіндегі ЖНЖ-нің толық қателігі былай болады:

$$E(w) = \sum_{i=1}^K [f_i(x_i w) - y_i]^2.$$

Толық қателіктің екі қасиетін айта кетейік. Біріншіден, $E(w)$ қателігі W күйіндегі кеңістігінде анықталған w күйдегі функция болады. Екіншіден, ЖНЖ кейбір w^* оқытылған күйде оқытушы іріктелулерде қателіктер жасамайды, бұл функция нөлдік мәнге ие болады. Демек, оқытылған күй енгізілген $E(w)$ функциясының *минимум нүктесі* болады.

E қателік функциясы еркін түрде болуы мүмкін. Егер оқытатын мысалдарының жиыны мен қателік функциясының есептеу тәсілдері таңдалса, онда ЖНЖ оқытылуы көпөлшемді оңтайландыру есебіне айналады, оны шығару үшін келесі әдістер қолданылуы мүмкін:

– бірінші ретті туындыны есептей отырып, локалды оңтайландыру (градиенттік әдіс, антиградиентке бағытталған мақсатты функцияның бір өлшемді және екі өлшемді

оңтайландыру әдісі, түйіндес градиенттер әдісі, алгоритмнің бірнеше қадамындағы антиградиентке бағытталғанын ескеру әдісі);

– бірінші және екінші ретті туындыны есептей отырып, локалды оңтайландыру (Ньютон әдісі, Гессе матрицаларымен оңтайландыру әдісі, квазиньютон әдісі, Гаусса-Ньютон әдісі, Левенберг-Маркардт әдісі);

– стохастикалық оңтайландыру (кездейсоқ бағытта іздеу әдісі, босандату имитациясы, Монте-карло әдісі);

– глобалды оңтайландыру (мақсатты функция тәуелді болатын айнымалылардың мәнін теру әдісі).

Көп қабатты ЖНЖ-нде ереже бойынша, оңтайлы шығыс мәндері соңғы қабаттан басқа барлық қабатта белгісіз. Егер ЖНЖ-нің қабаттар саны үштен артық болса, онда шығу қабатындағы қателердің шамаларымен ғана оқыту мүмкін емес. Бұл есептің бірден-бір шешімі—әрбір қабат үшін кіріс сигналдарға сәйкес тиісті шығу сигналдарының жиынтығын әзірлеу, бұл күрделі амал және әрдайым орындала бермейді.

Басқа шешім —салмақтардың динамикалық бапталуы, оның барысында әлсіз синапстық байланыс таңдалады да, оның шамасын сол немесе оң жаққа қарай өзгертіледі және тек барлық ЖНЖ шығысындағы қателерді азайтатын өзгертулер ғана сақталатын болады, бұл өте көп кішігірім есептеулерді талап етеді.

Үшінші шешім—әдеттегі жұмыс тәртібінде сигналдардың тура таралуына кері бағытында қателіктер сигналының ЖНЖ-нің шығысынан кірісіне қарай таралуы.

Осылайша ЖНЖ-нің үйрету мәселесі күйлер кеңістігіндегі қателіктер функциясының минимумын іздеу болады және оны шешу үшін оңтайландыру теориясының стандартты тәсілдері қолданылуы мүмкін. Бұл есеп көпфакторлы есептердің сыныбына жатады, мысалы бір қабатты N кірісі бар және M шығысы бар персептрон үшін ол $N \times M$ өлшемді кеңістігіндегі минимумды іздеу болады.

Практикада аз шамадағы қатесі бар күйдегі қателер функциясының дәл минимумы болып табылмайтын ЖНЖ-лер

қолданылуы мүмкін. Басқаша айтқанда, шешім ретінде үйретілген W^* күйінің аймағындағы кейбір күй қабылданады. Қателіктің рұқсат етілетін дәрежесі нақты қолданбалы есептердің ерекшеліктерімен анықталады және де пайдаланушының оқытуға жұмсайтын қаражатына сай болатындай.

Ең қарапайым қателіктер функциясы K оқытатын мысалдарға орташаланған қажетті шығыс пен әрбір I үйрететін мысалы үшін нақты алынған y_I шығысының арасындағы айырмасының квадраттар сомасы болатын орташа квадраттық қателік:

:

$$E(w) = \frac{1}{K} \cdot \sum_{i=1}^K (d_i - y_i)^2$$

Мұғаліммен бірге оқытудың күшейтілген нұсқасы мадақтау арқылы оқыту деп аталады, ол жерде қажетті шығыстың нақты мәні көрсетілмейді, жүйенің жақсы немесе жаман жұмыс істегеніне баға беріледі.

III.1.2-тапсырмалар.

1. Оқытатын мысалдардың неден тұратындығын көрсетіңіз.
2. Оқытатын мысалдарды жариялау тәсілдерін суреттеңіз.
3. Көп қабатты ЖНЖ-да шығыстарындағы мәндерді тиімділеу есептерін шешу нұсқаларын көрсетіңіз.
4. ЖНЖ-ны оқытудың толық қателігінің формуласын жазыңыз.
5. Ең қарапайым қателіктер функциясын атаңыз.

III.1.2-сұрақтар.

1. Мұғаліммен оқыту нені білдіреді?
2. Оқытатын мысалды ұсынғанда не болады?
3. Қателіктер функциясы дегеніміз не?
4. Қай есеп нейрондық желіні оқыту есебі болады?
5. Мұғаліммен күшейтілген оқыту кезінде не көрсетіледі?

III.1.2-тестілер.

1. Әрбір оқытатын таңдама қандай жұптан тұрады?

- A) кіру векторы және мақсатты (қалаулы шығыс) векторы;
- B) шығу векторы және мақсатты (қалаулы кіріс) векторы;
- C) сыныптар векторы және мақсатты (қалаулы сынып) векторы;
- D) қатарлар векторы және мақсатты (қалаулы тізім) векторы;
- E) есептер векторы және қалаулы есептің векторы.

2. w күйде болатын ЖНЖ-нің толық қателігінің формуласы қайсы?

- A) $E(w) = \sum_{i=1}^K [f_i(x_i w) - y_i]^2$;
- B) $E(w) = \sum_{i=1}^K [f_i(x_i) - y_i]^5$;
- C) $E(w) = \sum_{i=1}^K [f_i(x_i w)]^3$;
- D) $E(w) = \sum_{i=1}^K [(x_i w) - y_i]^2$;
- E) $E(w) = \sum_{i=1}^K [f_i(x_i w)]^2$.

3. K үйрететін мысалдарға орташаланған d_I қажетті шығыс пен нақты алынған y_I шығысының арасындағы квадраттар қосындысының айырмасының үйрететін әрбір I мысал үшін қате функциясы қандай?

- A) $E(w) = \frac{1}{K} \cdot \sum_{i=1}^K (d_i - y_i)^2$;
- B) $E(w) = \frac{1}{K} \cdot \sum_{i=1}^K (d_i)^2$;
- C) $E(w) = \sum_{i=1}^K [f_i(x_i w)]^2$;
- D) $E(w) = \sum_{i=1}^K [f_i(x_i w) - y_i]^2$;

$$E) E(w) = \frac{1}{K} \cdot \sum_{i=1}^K (y_i)^3.$$

4. Мұғаліммен бірге оқытудың күшейтілген нұсқасы қалай аталады?

- A) мадақтау арқылы оқыту;
- B) жарысу арқылы оқыту;
- C) бәсекелестік оқыту;
- D) есептер көмегімен оқыту;
- E) тереңдетіп оқыту.

5. ЖНЖ-нің үйрету мәселесі неде?

- A) күйлер кеңістігіндегі қателіктер функциясының минимумын іздеу;
- B) күйлер кеңістігіндегі қателіктер функциясының максимумын іздеу;
- C) күйлер кеңістігіндегі кез келген функцияның экстремумын іздеу;
- D) кез келген кеңістікте қателіктер функциясының максимумын іздеу;
- E) кез келген кеңістікте функцияларды зерттеу.

III.1.3. Нейрондық желілерді мұғалімсіз оқытудың парадигмасы.

Осы тақырыпта жасанды нейрондық желілерді (ЖНЖ) мұғалімсіз оқыту сұрақтары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Мұғалімсіз оқытудың басты белгісі болып оның «дербестілігі» болады. Мұғаліммен оқыту кезіндегі сияқты оқыту үдерісі синапстық желілердің салмақтарын сәйкес орналастыруда болады. Мұғалімсіз оқытуды оқыту үлгілерінің барлық мысалдары үшін дұрыс жауап белгісіз болған жағдайда қолданылады. Олай болса үлгілерді сыныптар бойынша тарату мақсатымен деректер желісіне түсетін ішкі құрылысын анықтау әрекеті қарастырылады, яғни алдын-ала анықталған дұрыс жауаптармен салыстыру және мақсатты шығыстар талап етілмейді.

ЖНЖ-да кіру векторының кейбір мәндері жарияланады және оларды түрлендіру кезінде оларда нейрондардың саны және олардың өзара қарым-қатынасы өзгереді, оларда ЖНЖ-ны қандай да бір есепті шеше алатындай жағдайға жеткізетін үдерістер орын алады. Бұл жерде ЖНЖ-лардың салмақтары үйлесімді шығыстар пайда болатындай орналасады, яғни жақын шығыс векторларының жариялануы бірдей шығыстарды беретіндей болуы қажет. Бұл жағдайда белгілерді сыныптар бойынша сыныптауға мүмкіндік беретін үлгілер арасындағы корреляция немесе кіріс векторларының структурасы ашылады, яғни оқыту үдерісі кіру векторы мәндерінің көпшілігінің статистикалық қасиеттерін бөледі және ұқсас кіріс векторларын сыныптарға біріктіреді. Берілген сыныптағы ЖНЖ-нің кіру векторын жариялау белгілі бір шешімді береді бірақ оқытуға дейін кіру векторының берілген сыныбында қандай шығыстың орындалатынын болжау мүмкін емес. Сондықтан ЖНЖ шығыстары оқыту үдерісі анықтаған белгілі бір түсінікті формаға түрлену керек. Бұл мәселе болып табылмайды. Әдетте ЖНЖ орнатқан кіріс пен шығыс арасындағы байланысты идентификациялау қиындық тудырмайды.

Синапстық желілердің салмақтарын оңтайластыру нейрондағы қолжетімді дерек, яғни оның жағдайы және бар синапстық байланыстардың салмақтары негізінде ғана жүзеге асырылуы мүмкін.

Жауаптардың түрлері кіріс бейнелерінің әрбір сыныбы үшін алдын-ала белгісіз екенін айта кету қажет және инициализациялау кезіндегі салмақтарды кездейсоқ топтастыруға негізделген нейрондардың шығыс қабаттарының ерікті үйлесімділігін танытады. Сонымен қатар ЖНЖ ұқсас бейнелерді оларды бір сыныпқа жатқызу арқылы жалпылай алады. Үйретілген ЖНЖ-ны тексеру сыныптардың шығу қабатындағы топологиясын анықтауға мүмкіндік береді. Үйретілген ЖНЖ-нің жауаптарын көрсетуді ыңғайлы түрге келтіру үшін желіні бір қабатпен толықтыруға болады және ол бір қабатты персептронды үйрету алгоритмі бойынша оны ЖНЖ-нің шығыс реакцияларын қажетті үлгілерді көрсетуге мәжбүрлеу керек.

Кейбір мұғалімсіз оқыту алгоритмдері ЖНЖ-нің құрылысында, яғни нейрондардың санын және олардың өзара қарым-қатынасын өзгертеді. Мұндай түрлендірулерді *дербес ұйымдастырылу (self-organizing)* кеңейтілген терминмен атаған дұрысырақ, бірақ ағылшыннан бұл аударылу сәтті емес, өйткені бұл жерде ЖНЖ-нің құрылысының өзгеруі айтылып жатқан жоқ, тек нейронда бар, яғни оның күйі және бар синаптикалық байланыстар салмақтары туралы дерекке сүйену арқылы салмақтардың құрастырылуы айтылып жатыр.

q қабатынан j нейронын минималды ара-қашықтықпен таңдағаннан кейін:

$$D_j = \sqrt{\sum_{i=0}^{p-1} (y_i^{(q-1)} - w_{ij})^2}, \quad (\text{III.1.3.1})$$

оқыту (салмақтарды реттеу) төмендегі тәсілдердің біреуімен жүзеге асырылады:

$$1. \text{ Хебб тәсілі: } w_{ij}(t) = w_{ij}(t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)}, \quad (\text{III.1.3.2})$$

2. Кохонен тәсілі: $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)} - w_{ij}(t-1)]$,
(III.1.3.3)

мұнда j – q қабатындағы нейрон индексі, I – $(q-1)$ қабаттағы нейрондардың қосылу индексі; p – p -өлшемді кеңістігіндегі векторлар өлшемі;

$y_i^{(q-1)}$ – $(q-1)$ қабатындағы I нейронының шығу мәні;

$y_i^{(q)}$ – q қабатындағы I нейронының шығу мәні;

$w_{ij}(t)$ және $w_{ij}(t-1)$ – сәйкесінше t және $t-1$ итерацияларында осы нейрондарды қосатын синапс салмағы;

η – оқыту жылдамдығының коэффициенті;

I – $(q-1)$ қабатының нейрондарды қосу индексі;

$(q-1)$ қабатындағы нейрондардың шығысы q қабаты үшін кіріс мәндері болады. (III.1.3.1) формуласында түбір алу міндетті емес, өйткені әртүрлі D_j салыстырмалы бағалары ғана маңызды.

Жоғарыда келтірілген (III.1.3.3) формуласынан үйрету $y_i^{(q-1)}$ алдыңғы қабатындағы нейрондардың шығысынан және оның синапстық салмақтарынан келіп түсетін нейронның кіру сигналдары арасындағы айырмашылықты минималдауға келіп тіреледі.

Сонымен қатар Хебб оқыту тәсілінің дифференциалды әдісі бар $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - y_i^{(q-1)}(t-1)] \cdot [y_j^{(q)}(t) - y_j^{(q)}(t-1)]$, (III.1.3.4)

мұнда $y_i^{(q-1)}(t)$ және $y_i^{(q-1)}(t-1)$ – t және $t-1$ итерацияларындағы $n-1$ қабатының I нейронының шығу мәні;

$y_j^{(q)}(t)$ и $y_j^{(q)}(t-1)$ – q қабатының j нейроны үшін де дәл солай .

(III.1.3.) формуласынан көрініп тұрғандай шығыстары үлкею жағына қарай динамикалық түрде өзгерген нейрондарды қосатын синапстар күштірек үйретіледі.

Оқытудың жоғарыда келтірілген формулаларды қолданатын толық алгоритм осы түрде болады :

1. Инициализация кезінде барлық салмақтық коэффициенттерге аз кездейсоқ мәндер беріледі.

2. ЖНЖ кірісіне кіріс бейнесі беріледі және қозу сигналдары классикалық тік ағымды (feedforward) желілердің принциптеріне сәйкес барлық қабаттарда таралады, яғни әрбір нейрон үшін келесіде нейронды белсендіру функциясы қолданылатын оның кірістерінің салмақталған сомасы есептеледі, нәтижесінде оның $y_i^{(q)}$, $I=0...M_I-1$ шығу мәні алынады, бұл жерде M_I – I қабатындағы нейрондар саны, $n=0...N-1$, ал N – ЖНЖ-гі қабаттар саны.

3. Нейрондардың алынған шығыс мәндерінің негізінде (III.1.3.2) немесе (III.1.3.3) формулалары арқылы, синапстық байланыстардың салмағының өзгеруі жүргізіледі.

4. Жасанды нейрон жүйелерінің (ЖНЖ) шығыс мәндері берілген дәлдікпен тұрақтанбайынша, цикл 2-қадамнан бастап қайталанады. ЖНЖ үшін қолданылған кері таратудан ерекшеленетін, оқытуды аяқтаудың бұл жаңа тәсілін қолдану, синапстық байланыстар салмақтарының реттелген мәндері іс жүзінде шектелмейтіндігіне негізделген.

Циклдің екінші қадамында кіріс жиынтығынан барлық бейнелер кезекпен ұсынылады.

Кіріс бейнелердің әрбір сыныбына жауап түрлерінің алдын-ала белгілі болмайтынын және инициализация кезеңінде салмақтарды кездейсоқ таратуға негізделген шығу қабатының нейрондар жағдайының еркін үйлесуін беретіндігін айта кеткен жөн. Сонымен қатар, желі ұқсас бейнелерді бір сыныпқа біріктіре отырып жалпылай алады. Оқытылған ЖНЖ-ні тестілеу, шығу қабатында сынып топологиясын анықтауға мүмкіндік береді. Оқытылған ЖНЖ-ң жауаптарын ыңғайлы ұсынуға алып келу үшін желіні бір қабатпен толықтыруға болады.

Оқытудың толық алгоритмі Хебб әдісіне ұқсас құрылымға ие, бірақ, 3-қадамда барлық қабаттан синапстық байланыстар салмақтарының мәні кіріс бейнесіне максималды тұрғыдан сәйкес келетін нейрон таңдап алынады, және (III.1.3.4) формула бойынша салмақтарды реттеу тек ол үшін ғана орындалады. Бұл

аккредитация қабаттың басқа барлық нейрондарын тежеумен және таңдалған нейронға қанығуды енгізумен қоса жүргізіледі. Мұндай нейронды таңдау, мысалы, кіріс мәндер векторымен, синапстық байланыстар салмақтары векторының скалярлық көбейтіндісін есептеумен іске асырылуы мүмкін. Максималды көбейту жеңген нейронды береді.

Бұл жағдайда ең кіші қашықтықтағы нейрон «жеңеді». Нейрон қабаттарының «құқықтарын теңестіру» үшін, кейде жиі аккредитация алатын нейрондар күшпен қарастырудан шығып қалады. Мұндай алгоритмнің қарапайым түрі, жаңа ғана жеңген нейронның тежелуімен түйінделеді.

Кохонен алгоритмін пайдалану кезінде, кіріс бейнелерінің бір қалыпқа түсу тәжірибесі, сонымен қатар, баптау кезеңінде және нормалдаудың синапстық байланыс салмағының бастапқы мағынасында жүреді:

$$x_i = \frac{x_i}{\sqrt{\sum_{j=0}^{n-1} x_j^2}}, \quad (\text{III.1.3.5})$$

мұнда, x_i – кіріс бейнесінің i -ші компонент векторы немесе синапстық байланыс салмағының векторы, ал n – оның өлшемі. Бұл оқыту үдерісінің ұзақтығын қысқартуға мүмкіндік береді.

Синапстық байланыс салмағын кездейсоқ мағынада баптау, кіріс үлгілері тығыз орналастырылып сәйкестенсе, әртүрлі сыныптар қосылады, немесе керісінше, сол сыныптың жақын бейнелері қосымша сыныптарға біріктіріледі.

Осы жағдайды болдырмау үшін дөңес біріктіру әдісін пайдаланамыз. Оның мәні, нормаланған кіріс бейнелерінің түрленуінде:

$$x_i = \alpha(t) \cdot x_i + (1 - \eta(t)) \cdot \frac{x_i}{\sqrt{n}}, \quad (\text{III.1.3.6})$$

мұнда x_i – кіріс бейнесінің i -ші компоненті, n – оның компонентінің барлық саны, $\eta(t)$ – оқыту үдерісі кезінде 0-ден

бастап соңына дейін өзгертін, нәтижесінде бірдей бейнелер ЖНЖ-ң кірісін беретін, ал сол уақыт ішінде барлығы шығысқа кірістірілетін оқыту коэффициенті болады. Синапстық байланыс салмағының мағынасы тең ауқымды баптау қадамында орналастырылады:

$$w_0 = \frac{1}{\sqrt{n}}, \quad (\text{III.1.3.7})$$

мұнда n – бапталған қабаттағы салмақ векторының өлшемі.

Жоғарыда қарастырылған әдістер бойынша, «өзін-өзі ұйымдастыратын құрылымдар» деп аталатын ерекше түрдегі нейрондық желілер құрылады (self-organizing feature maps). j нейронының q қабатынан таңдалғаннан кейін, олар үшін D_j (III.1.3.1) минималды арақашықтығы (III.1.3.2) осы формула бойынша оқытылады, және тек (III.1.3.3) бұл нейрон емес, сонымен қоса оның R шекараларында орналасқан көршілері де оқытылады. R шамасы бірінші итерацияда өте үлкен, сондықтан барлық нейрондар оқытылады, бірақ уақыт өте келе итерация 0-ге дейін төмендейді. Бұл оқытудың соңы жақындаған сайын, бейнелер сыныптарының әрқайсысына жауап беретін нейрондар тобы анықталатындығын білдіреді.

III.1.3-тапсырмалар.

1. Мұғалімсіз оқыту әдісінде пайдаланылатын минималды арақашықтық формуласын жазыңыз.
2. Хебб әдісімен нейрондық желінің салмақтарын реттеу формуласын жазыңыз.
3. Кохонен әдісімен салмақты реттеу формуласын жазыңыз.
4. Хебб нейрондық желісін дифференциалдық оқыту әдісінің формуласын жазыңыз.
5. Кіріс бейнелерінің қалыпқа келу формуласын жазыңыз.
6. Синапстық байланыстың бастапқы салмақ мағынасының қалыпқа келу формуласын жазыңыз.

7. Синапстық байланыс салмағы мағынасын орналастыру формуласын жазыңыз.

III.1.3-сұрақтар.

1. Қай кезде жасанды нейрондық желілерді мұғалімсіз оқыту қолданылады?

2. Мұғалімсіз оқытуда қандай мүмкіндік қабылданады?

3. Мұғалімсіз оқыту негізінде, ерекше түрдегі қандай ЖНЖ құрылады?

4. Синапстық байланыс салмағы қашан орналастырылады?

5. Кездейсоқ мағынадағы синапстық байланыстың салмағын баптау неге әкеледі?

6. «Өздігінен ұйымдастырылатын» құрылымдар ненің негізінде құрылады?

7. «Өздігінен ұйымдастырылатын» құрылымдар деп аталатын ерекше түрдегі нейрондық желілер ненің негізінде құрылады?

III.1.3-тестілер.

1. Хебб тәсілі қайсысы?

A) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)};$

B) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)} - w_{ij}(t-1)];$

C) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - y_i^{(q-1)}(t-1)] \cdot [y_j^{(q)}(t) - y_j^{(q)}(t-1)];$

D) $w_{ij}(t) = w_{ij}(t-1);$

E) $w_{ij}(t) = [y_i^{(q-1)} - w_{ij}(t-1)];$

2. Кохонен тәсілі қайсысы?

A) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)} - w_{ij}(t-1)];$

B) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)};$

- C) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - y_i^{(q-1)}(t-1)] \cdot [y_j^{(q)}(t) - y_j^{(q)}(t-1)]$;
 D) $w_{ij}(t) = w_{ij}(t-1)$;
 E) $w_{ij}(t) = [y_i^{(q-1)} - w_{ij}(t-1)]$

3. Хебб оқыту тәсілінің дифференциалды әдісі қайсысы?

- A) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - y_i^{(q-1)}(t-1)] \cdot [y_j^{(q)}(t) - y_j^{(q)}(t-1)]$;
 B) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)}$;
 C) $w_{ij}(t) = w_{ij}(t-1) + \eta \cdot [y_i^{(q-1)} - w_{ij}(t-1)]$;
 D) $w_{ij}(t) = w_{ij}(t-1)$;
 E) $w_{ij}(t) = [y_i^{(q-1)} - w_{ij}(t-1)]$

4. Мұғалімсіз оқытудың басты белгісі не болады?

- A) дербестілігі;
 B) анықтылығы;
 C) жинақтылығы;
 D) жалпылығы;
 E) нәтижелілігі.

5. Кейбір мұғалімсіз оқыту алгоритмдері нейрондардың санын және олардың өзара қарым-қатынасын өзгертетін түрлендірулерді қалай атайды?

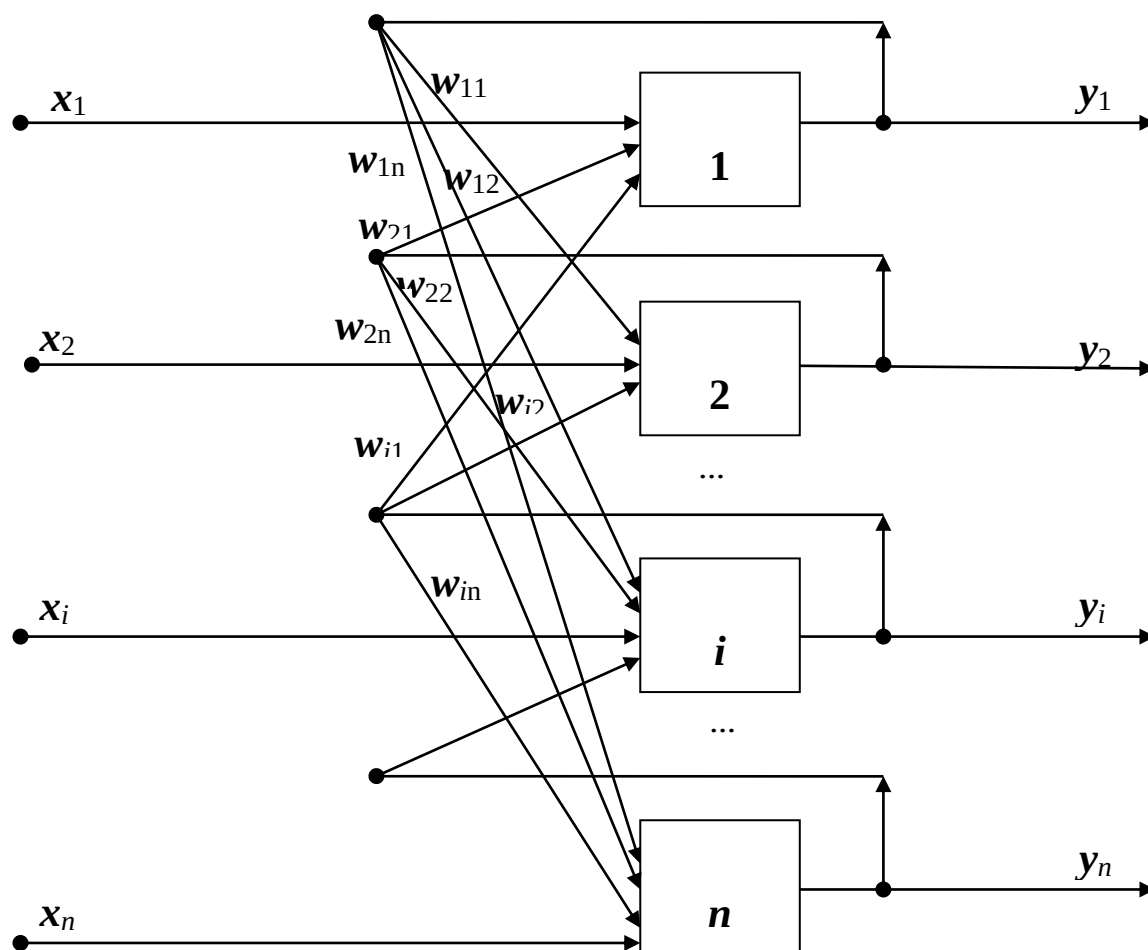
- A) дербес ұйымдастырылу;
 B) жалпы ұйымдастырылу;
 C) дербес тарату;
 D) дербес жинақтау;
 E) жылдам ұйымдастырылу.

III.1.4. Нейрондық желілерді аралас оқыту парадигмасы

Бұл тақырыпта аралас ЖНЖ-ны оқыту парадигмасын қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Аралас оқыту мұғаліммен оқыту және мұғалімсіз оқыту жиынтығынан тұрады, яғни салмақ бөлігі мұғаліммен оқыту құралы арқылы анықталса, келесі бөлігі өзіндік оқыту алгоритмі арқылы жүзеге асырылады.

Аралас оқытудың демонстрациясы үшін Хопфилдтың толық байланысты нейрондық желісін қарастырамыз (III.1.4.1-сурет).



III.1.4.1-сурет. Хопфилдтың толық байланысты нейрондық желісі

Суретке сәйкес кіру векторы $X=(x_1, x_2, \dots, x_n)$, n – ЖНЖ-ғы нейрондар саны және кіріс, шығыс векторларының өлшемі. Әрбір x_i элемент $+1$ немесе -1 -ге тең. k -шы вектордың мысалы – X^k , ал оның компоненттері – x_i^k , $k=1, \dots, m$, m – мысалдар саны. Егер бейне танылса, онда ЖНЖ шығысы $Y=X^k$ -ға тең, мұнда Y – ЖНЖ шығыс мағынасының векторы: $Y=\{y_i: i=1, \dots, n\}$.

Хопфилд нейрондық желісін баптау (инициализациялау).

ЖНЖ синапстық байланыс салмағы төмендегідей анықталады:

$$W_{ij} = \begin{cases} \sum_{k=1}^m x_i^k x_j^k, & i \neq j \\ 0, & i = j \end{cases},$$

мұнда i және j – нейрондардың пресинапстық және постсинапстық индекстері; x_i^k – i -ші және x_j^k – j -ші вектор элементінің k -шы мысалы.

Хопфилд нейрондық желісінің жұмыс істеу алгоритмі.

1-қадам. ЖНЖ кірісінде белгісіз сигнал беріледі. Шын мәнінде оның берілуі аксондардың мағынасын орнату болады. $y_i(0)=x_i$, $i=1, \dots, n$.

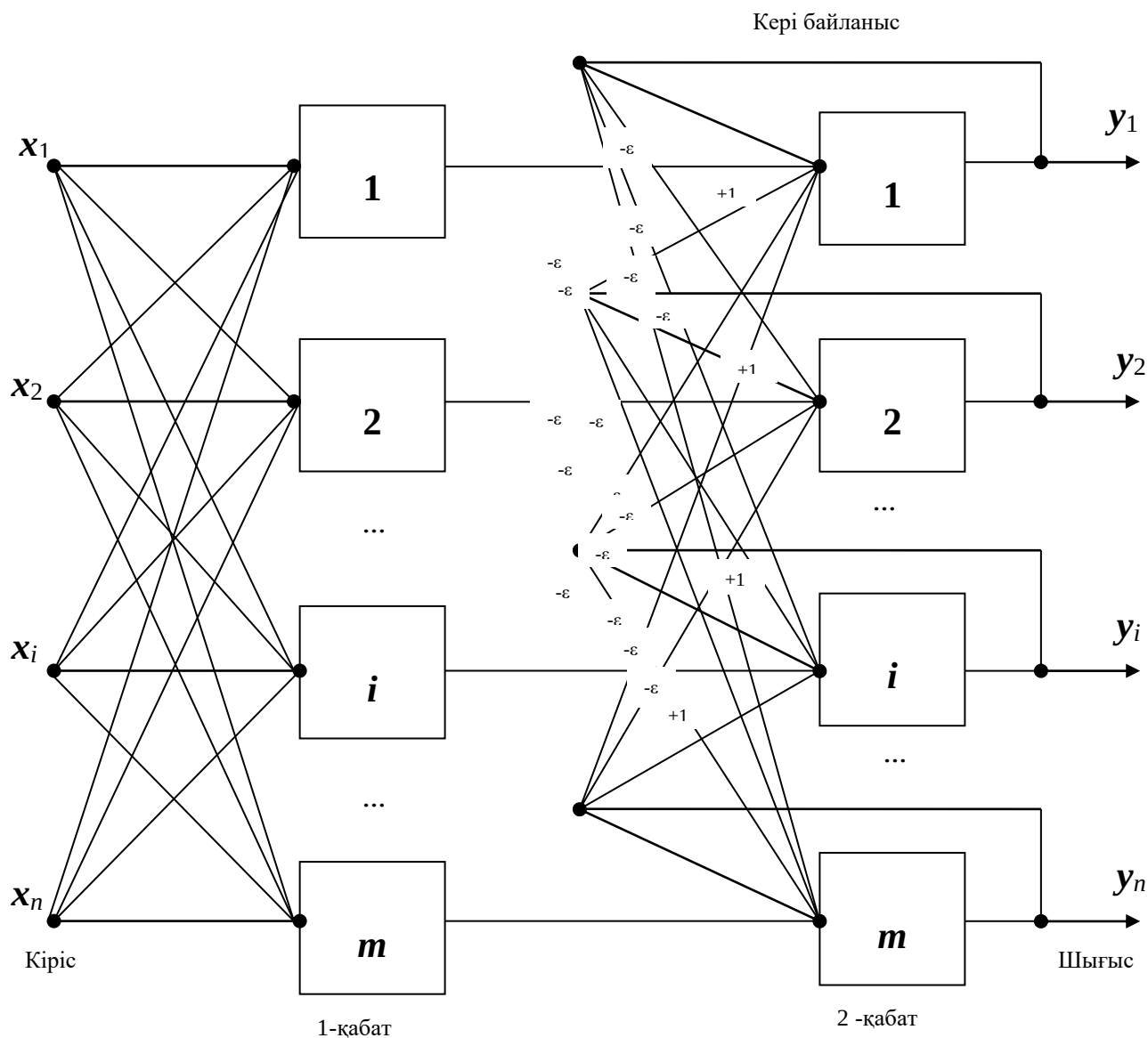
2-қадам. Аксон мағыналарының және нейрондардың жаңа күйлері есептеледі: $s_j(p+1)=\sum_{i=1}^n w_{ij} y_i(p)$; $y_j(p+1)=f(s_j(p+1))$, мұнда p – итерация нөмірі, f – бірқалыпты секіруді белсендіру функциясы;

$$Y = \begin{cases} 0, & X < T \\ 1, & X \geq T \end{cases}.$$

3-қадам. Соңғы итерацияда шығыс аксондарының мағынасы өзгерген не өзгермегендігі тексеріледі. Егер өзгерсе, 2-қадамға көшеді, ал олай болмаса – аяқталады.

Енді ІІІ.1.4.2-суретте көрсетілген Хемминг нейрондық желісін қарастырайық.

Мұндай нейрондық желіні оқыту кезінде Хемминг қашықтығының минималды жолынан белгісіз кіру сигналына өту жолын таңдау қажет. Еске сала кететін жайт, Хемминг қашықтығы дегеніміз екі бинарлық вектордың бірдей өлшемдерінің мағыналары сәйкесінше әртүрлі болатын сандар позициясы (лауазымы). Көптеген жағдайларда Хемминг қашықтығы кез-келген q -ші әліпбидің бірдей өлшем векторлары үшін қолданылады және 1 өлшемді объектінің метрикасының айырмасы (метрикалық кеңістіктегі қашықтық анықтау функциясы) үшін қызмет атқарады.



ІІІ.1.4.2-сурет. Хемминг нейрондық желісі

Хэмминг нейрондық желісін баптау (инициализациялау).

ЖНЖ-ың 1-қабатының салмағы және белсендіру функциясының сатысы келесі теңдеулерге тең:

$$W_{ik} = \frac{x_i^k}{2}, \quad i=1, \dots, n-1, \quad k=1, \dots, m-1$$

$$T_k = \frac{n}{2}, \quad k=1, \dots, m-1,$$

мұнда x_i^k – k -шы мысалдың I -ші элементі.

Синапстық салмақтың тежелуі 2-қабатта ε шартын қанағаттандырады $0 < \varepsilon < 1/m$. Аксонымен байланысқан нейрон синапсы +1 салмағын қабылдайды.

Хэмминг нейрондық желісінің жұмыс істеу алгоритмі.

1. ЖНЖ кірісінде белгісіз $X = (x_1, x_2, \dots, x_{n-1})$, векторы беріледі, және олардан нейронның бірінші қабатының күйі есептелінеді:

$$y_j^{(1)} = s_j^{(1)} = \sum_{i=0}^{n-1} w_{ij} x_i + T_j, \quad j=0, 1, \dots, m-1,$$

Мұнда жоғары индекс жақша ішіндегі қабат номерін көрсетеді.

Осыдан кейін, екінші қабаттың аксон мағынасын алынған мағынамен баптаймыз: $y_j^{(2)} = y_j^{(1)}, \quad j=0, 1, \dots, m-1$.

2. 2- қабат нейрондарының жаңа күйінің есептелуі және олардың аксондарының мағынасы:

$$s_j^{(2)}(p+1) = y_j^{(1)}(p) - \varepsilon \sum_{k=0}^{m-1} y_k^{(2)}(p), \quad k \neq j, \quad j=0, 1, \dots, m-1;$$

$$y_j^{(2)}(p+1) = f(s_j^{(2)}(p+1)), \quad j=0, 1, \dots, m-1.$$

f белсендіру функциясы шек түрінде беріледі, кез-келген мүмкін аргументті қанықтыруға әкелмес үшін, оның мағынасы жеткілікті түрде үлкен болуы керек.

3. Соңғы итерацияға дейін 2-қабаттағы нейрон шығыстары өзгерген-өзгермегендігін тексеру. Егер өзгерсе 2-қадамға көшеді, олай болмаса – аяқталады.

III.1.4-тапсырмалар.

1. ЖНЖ аралас оқытудың мәнісін жазыңыз.
2. Хопфилд желісінің жұмыс істеу алгоритмін сипаттаңыз.
3. Хэмминг желісінің жұмыс істеу алгоритмін сипаттаңыз.

III.1.4-сұрақтар.

1. Нейрондық желілерді аралас оқыту қалай орындалады?
2. Хопфилд желісі қалай түсіндіріледі?
3. Хэмминг желісі қалай түсіндіріледі?

III.1.4-тестілер.

1. Аралас оқытуды көрсету үшін қандай желі қолданылады?

- A) Хопфилдтың толық байланысты нейрондық желісі;
- B) Хеббтың толық байланысты нейрондық желісі;
- C) Кохоненнің толық байланысты нейрондық желісі;
- D) Гроссберг қабаты;
- E) Кохонен қабаты.

2. Хопфилд нейрондық желісін баптауда (инициализациялауда)

ЖНЖ синапстық байланыс салмағы қалай анықталады?

$$A) w_{ij} = \begin{cases} \frac{\sum_{k=1}^m x_i^k x_j^k}{m}, & i \neq j \\ 0, & i = j \end{cases};$$

$$B) s_j(p+1) = \sum_{i=1}^n w_{ij} y_i(p);$$

$$C) y_j(p+1) = f(s_j(p+1));$$

$$D) Y = \begin{cases} 0, & X < T \\ 1, & X \geq T \end{cases};$$

$$E) w_{ik} = \frac{x_i^k}{2}, \quad i=1, \dots, n-1, \quad k=1, \dots, m-1.$$

3. Аксон мағыналарының және нейрондардың жаңа күйлері қалай есептеледі?

$$A) s_j(p+1) = \sum_{i=1}^n w_{ij} y_i(p); y_j(p+1) = f(s_j(p+1));$$

$$B) w_{ij} = \begin{cases} \sum_{k=1}^m x_i^k x_j^k, & i \neq j \\ 0, & i = j \end{cases};$$

$$C) y_j^{(2)}(p+1) = f(s_j^{(2)}(p+1)), \quad j=0,1,\dots,m-1;$$

$$D) Y = \begin{cases} 0, & X < T \\ 1, & X \geq T \end{cases};$$

$$E) w_{ik} = \frac{x_i^k}{2}, \quad i=1,\dots,n-1, \quad k=1,\dots,m-1.$$

4. Хэмминг нейрондық желісін баптауда (инициализациялауда) ЖНЖ-ың 1-қабатының салмағы және белсендіру функциясының сатысы қандай теңдеулермен есептеледі?

$$A) w_{ik} = \frac{x_i^k}{2}, T_k = \frac{n}{2}, \quad i=1,\dots,n-1, \quad k=1,\dots,m-1;$$

$$B) s_j(p+1) = \sum_{i=1}^n w_{ij} y_i(p); y_j(p+1) = f(s_j(p+1));$$

$$C) y_j^{(2)}(p+1) = f(s_j^{(2)}(p+1)), \quad j=0,1,\dots,m-1;$$

$$D) Y = \begin{cases} 0, & X < T \\ 1, & X \geq T \end{cases};$$

$$E) w_{ij} = \begin{cases} \sum_{k=1}^m x_i^k x_j^k, & i \neq j \\ 0, & i = j \end{cases}.$$

5. Өлшемдері бірдей екі бинарлық вектордың сәйкес мәндері әртүрлі болатын сандар позициясы қалай аталады?

A) Хемминг қашықтығы;

B) Хеббтың қашықтығы;

- C) Кохоненнің қашықтығы;
- D) Гроссберг қашықтығы;
- E) Хопфилдтың қашықтығы.

III.2. Нейрондық желілерді оқыту әдістері

III.2. Қателерді кері тарату әдісімен нейрондық желілерді оқыту

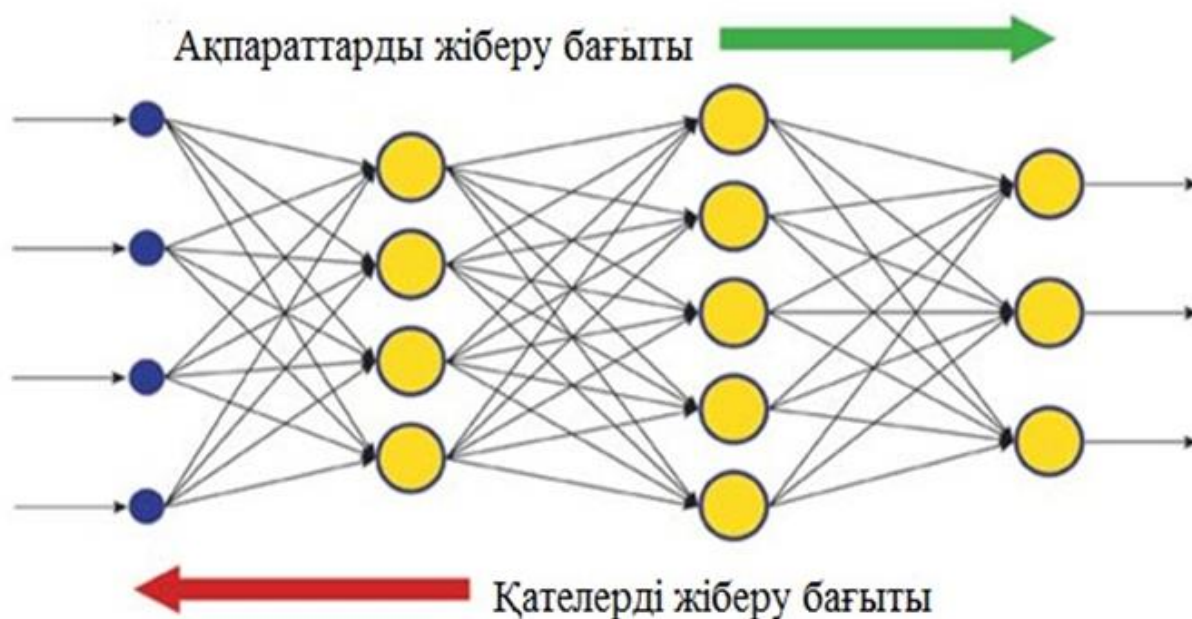
Бұл тақырыпта ЖНЖ-нің қателерді кері тарату әдісімен оқыту мәселелері қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

ЖНЖ оқыту әдістері оқытумен байланысты үш іргелі қасиеттерін: *сыйымдылықты, бейнелердің күрделілігін және есептеу күрделілігін* қарастырады.

Қателерді кері тарату әдісін (Error back propagation) 1986 жылы Румельхартпен әзірленген және тікелей байланыстармен көп қабатты ЖНЖ мұғаліммен оқыту үшін арналған.

Әдістің идеясы мыналардан тұрады:

1. кіріс және қаланатын шығыс мәндерінің векторларынан тұратын оқытушы жұптардың жиынын анықтау;
2. қаланатын шығу мәні мен нақты шығу мәні арасындағы айырма мәніне тең қатені есептеу;
3. бұл қатені кері бағытқа— деректер ағынына қарама-қарсы тарату және қосынды қатені минимизациялайтындай етіп нейрондар арасындағы синапстық байланыстар салмақтарының өзгеру жолымен қайта таратылады (III.2.1.1-сурет).



III.2.1.1-сурет Қателердің тарату бағыты

Сонымен қатар, қателерді кері тарату әдісі, дельта-ережені жалпылау әдісі деп те аталады.

Ендеше, қатені тарату оқытудың итерациялық алгоритмін береді:

1. Кездейсоқ түрде кезекті (x_i, y_i) оқыту жұбын таңдау, x_i —кіру векторының компоненттері, y_i —шығу векторының компоненттері, $i = 1, 2, \dots, K$; $j = 1, 2, \dots, N$; η —оқыту жылдамдығының коэффициенті, $0 < \eta < 1$.

2. ЖНЖ қалыпты жұмыс Істеу режиміндегі ЖНЖ шығысын есептеу;

3. $y_{ij}^{(Q)}$ нақты шығыс және $d_{ij}^{(Q)}$ қажетті шығыс (мақсаттық вектор) арасындағы айырмашылықты есептеу және шығу қабаты үшін дельта қатені келесі формула бойынша есептеу:

$$\delta_{ij}^{(Q)} = (y_{ij}^{(Q)} - d_{ij}^{(Q)}) \cdot \frac{dy_{ij}}{ds_{ij}}$$

мұнда $\frac{dy}{ds} = 1 - s$ гиперболалық тангенс жағдайында және

$\frac{dy}{ds} = s(1 - s)$ сигмоид жағдайында.

4. Қатені азайту үшін ЖНЖ-нің Q шығару қабаты үшін синапстық байланыстар салмақтарының өзгеруін есептеу.

$$\Delta w_{ij}^{(Q)} = -\eta \cdot \delta_{ij}^{(Q)} \cdot y_{ij}^{(Q-1)}$$

5. Формула бойынша барлық басқа қабаттар үшін қатені есептеу.

$$\delta_j^{(q)} = \left[\sum_k \delta_k^{(q+1)} \cdot w_{jk}^{(q+1)} \right] \cdot \frac{dy_j^{(q)}}{ds_j^{(q)}}$$

6. ЖНЖ-нің барлық синапстық байланыстар салмақтарын түзету.

$$w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \Delta w_{ij}^{(q)}(t),$$

$$w_{ij}^{(q)}(t) = -\eta \cdot (\mu \cdot \Delta w_{ij}^{(q)}(t-1) + (1-\mu) \cdot \delta_j^{(q)} \cdot \delta_i^{(q-1)}),$$

мұнда μ —инерциялылық коэффициенті. Бұл формула қателіктің ауытқуын төмендетуге алып келетін салмақтарды «тегіс» өзгертеді. Соңғы формуланы қателік берілгенге жақын болған кезде, оқыту соңында қолдану дұрыс болады.

7. Қателік барлық жиында қолайлы деңгейге жеткенге дейін оқытылатын жиынның әр вектор үшін 1-ден 6-ға дейінгі қадамдарды қайталау.

8. 1, 2 және 3 қадамдарымен орындалатын амалдар, оқытылған ЖНЖ жұмыс істеуінде орындалатын амалдарға ұқсас, яғни кіру векторы беріледі және нәтиже беретін алынған шығыс есептеледі. Есептеулер қабат бойынша орындалады. Қателіктің критеріі (мақсаттық функция) ретінде, шығу қабатының барлық нейрондары бойынша және ЖНЖ өңделетін барлық бейнелер бойынша орташа квадраттық қателік қолданылады:

$$E(w) = \frac{1}{2} \cdot \sum_{j=1}^N \sum_{i=1}^K (y_{ij}^{(q)} - d_{ij}^{(q)})^2, \quad (\text{III.2.1.1})$$

мұнда $y_{ij}^{(q)}$ — j -ші оқылатын мысалдағы (бейнедегі) i -ші нейрон үшін ЖНЖ q -ші шығу қабатының нақты шығысы, $d_{ij}^{(q)}$ — осы нейронның қалаулы шығысы, M — оқытатын мысалдар саны, K — кірістердің саны, N — шығыстар саны, $q = 1, 2, \dots, Q$ — ЖНЖ қабаттар саны.

9. Осындай функционалды минимизациялай отырып, біз ең кіші квадраттар әдісі бойынша шешімді аламыз.

Нәтижесінде, әрбір нейрон өзінің салмақтарының желі қосындысының қатесіне қосатын әрқайсысының үлесін анықтауға қабілетті.

Оқытудың қарапайым ережесі, синапстық байланыстар салмақтарының өзгеруі олардың жалпы қатеге қосатын үлесіне пропорционал болатын ең жылдам түсу әдісіне сәйкес келеді, яғни ЖНЖ қатесінің синапстық байланыстар салмағының өзгеруіне сезімталдығы есептеледі. Ол үшін салмақтар бойынша қатеден дербес туындыны есептеу керек және минимумды табу үшін градиенттік түсу әдісін қолдану керек, ол оқытудың әр қадамында бұл салмақтардың өзгеруін келесі формула бойынша қамтамыз етеді:

$$\Delta w_{ij}^{(q)} = -\eta \cdot \frac{\partial E}{\partial w_{ij}^{(q)}}, \quad (\text{III.2.1.2})$$

мұнда $\Delta w_{ij}^{(q)}$ - синапстың байланыстар салмақтарының түзетілуі, $w_{ij}^{(q)}$ - $(q-1)$ қабатындағы i -ші нейронмен q -ші қабаттың j -ші нейронының синапстық салмағы, η – оқыту жылдамдығының коэффициенті, $0 < \eta < 1$; $q = Q-1, Q-2, \dots, 2$.

Осылайша, ол ЖНЖ барлық синапстық салмағы бойынша E мақсаттық функцияның дербес туындысын анықтау талап етіледі. Күрделі функцияны дифференциалдау ережесіне сәйкес келесідегідей:

$$\frac{\partial E}{\partial w_{ij}^{(q)}} = \frac{\partial E}{\partial y_j^{(q)}} \cdot \frac{dy_j^{(q)}}{ds_j^{(q)}} \cdot \frac{\partial s_j^{(q)}}{\partial w_{ij}^{(q)}},$$

мұнда $y_j^{(q)}$ - шығыс, ал $s_j^{(q)}$ - ЖНЖ q -ші қабатындағы j -ші нейронының кірісінің салмақтанған қосындысы.

Белсендіру функциясының туындысы абциссаның барлық осінде анықталған болу керек болғандықтан, бірлік секіру функциясы және әртектілікпен белсендірудің басқа да функциялары қарастырылып отырған ЖНЖ үшін сай келмейді.

Оларда гиперболалық тангенс немесе классикалық сигмоид сияқты тегіс функциялар қолданылады.

Белсендіру функциясын біле отырып, (III.2.1.3) өрнегінің екінші

көбейткіші болып табылатын, оның $\frac{dy_j^{(q)}}{ds_{ij}^{(q)}}$ туындысын есептеуге болады. Мысалы, егер белсендіру функциясы:

1) классикалық сигмоид – логистикалық функция $f(s) = \frac{1}{1 + e^{-as}}$ болса, онда оның туындысы $f'(s) = a \cdot f(s) \cdot (1 - f(s))$ тең;

2) гиперболалық тангенс $f(s) = \frac{e^{ks} - e^{-ks}}{e^{ks} + e^{-ks}}$, онда оның туындысы $f'(s) = 1 - s^2$ тең.

(III.2.1.3) өрнегінің екінші көбейтіндісінің мәні логистикалық функция үшін мына формула арқылы есептеледі:

$$\frac{dy_j^{(q)}}{ds_j^{(q)}} = a \cdot y_j^{(q)} \cdot (1 - y_j^{(q)}) \quad (\text{III.2.1.4})$$

(III.2.1.3) өрнегінің үшінші көбейткіші $(q-1)$ қабатының I -ші нейронының шығысы ретінде болады, яғни:

$$\frac{\partial s_j^{(q)}}{\partial w_{ij}^{(q)}} = y_i^{(q-1)} \quad (\text{III.2.1.5})$$

Енді шығу қабатының нейрондарының синапстық байланыстарының салмақтары бойынша мақсаттық функцияның дербес туындысын оңай есептеуге болады.

(III.2.1.3) өрнегінің бірінші көбейткіші үшін қателіктер функциясын $y_j^{(Q)}$ бойынша дифференциалдай отырып алатынымыз:

$$\frac{\partial E}{\partial y_j^{(Q)}} = (y_j^{(Q)} - d_j^{(Q)}) \quad (\text{III.2.1.6})$$

Әрі қарай, (III.2.1.3) өрнегі негізінде (III.2.1.5) және (III.2.1.6) ескере отырып, келесі өрнекті алуға болады:

$$\frac{\partial E}{\partial w_{ij}^{(Q)}} = (d_j^{(Q)} - y_j^{(Q)}) \cdot \frac{dy_j^{(Q)}}{ds_j^{(Q)}} \cdot y_i^{(Q-1)} \quad (\text{III.2.1.7})$$

Ішкі қабаттардың нейрондардың синапстық байланыстар салмақтары үшін, біз (III.2.1.3) өрнегінен бірінші көбейткіш неге тең екендігін бірден жаза алмаймыз, дегенмен оны келесі түрде жазып көрсете аламыз:

$$\frac{\partial E}{\partial y_j^{(q)}} = \sum_k \frac{\partial E}{\partial y_k^{(q+1)}} \cdot \frac{dy_k^{(q+1)}}{ds_k^{(q+1)}} \cdot \frac{\partial s_k^{(q+1)}}{\partial y_j^{(q)}} = \sum_k \frac{\partial E}{\partial y_k^{(q+1)}} \cdot \frac{dy_k^{(q+1)}}{ds_k^{(q+1)}} \cdot w_{jk}^{(q+1)} \quad (\text{III.2.1.8})$$

Мұнда k бойынша қосу $(q + 1)$ қабатының нейрондарының арасында жүзеге асырылады. Жаңа айнымалы енгіземіз:

$$\delta_j^{(q)} = \frac{\partial E}{\partial y_j^{(q)}} \cdot \frac{dy_j^{(q)}}{ds_j^{(q)}} \quad (\text{III.2.1.9})$$

Сонда, шығу қабатындағы нейрондар үшін (III.2.1.6) және (III.2.1.9) ескере отырып, мынаны аламыз:

$$\delta_j^{(Q)} = (y_j^{(Q)} - d_j^{(Q)}) \cdot \frac{dy_j^{(Q)}}{ds_j^{(Q)}} \quad (\text{III.2.1.10})$$

(III.2.1.8) алғашқы екі көбейткіштің $\delta_k^{(q+1)}$ екенін көреміз. Сонда, (III.2.1.8) көмегімен q қабат нейрондарын $\delta_k^{(q)}$ арқылы ал, $(q+1)$ қабаттың нейрондарын $\delta_k^{(q+1)}$ шамасы арқылы өрнектеуге болады. $\delta_j^{(Q)}$ соңғы қабат үшін (III.2.1.10) арқылы есептеу жеңіл болғандықтан, барлық нейрондардың барлық қабаттары үшін $\delta_j^{(q)}$ мәнін рекурсивтік формула арқылы алуға болады:

$$\delta_j^{(q)} = \left[\sum_k \delta_k^{(q+1)} \cdot w_{jk}^{(q+1)} \right] \cdot \frac{dy_j^{(q)}}{ds_j^{(q)}} \quad (\text{III.2.1.11})$$

Енді (III.2.1.2) формуласын синапстық байланыстардың салмағын жетілдіру үшін (III.2.1.5) және (III.2.1.9) формулаларын ескере отырып, төмендегі түрде жазуға болады:

$$\Delta w_{ij}^{(q)} = -\eta \cdot \delta_j^{(q)} \cdot y_i^{(q-1)} \quad (\text{III.2.1.12})$$

Егер әртүрлі салмақ туындыларында қатты айырмашылық болса, онда бұл жағдайда (III.2.1.12) формуласы арқылы есептеу тиімсіз. Бұл s функциясының мәні кейбір нейрондар үшін модулі бойынша 1-ге жақын болғанда немесе кейбір салмақтардың модулі 1-ден көп үлкен болған жағдайға сәйкес келеді. Бұл жағдайда мақсатты функцияның бетіндегі ауысу кезіндегі шұғыл өзгерістерді тегістейтін кейбір инерциялықты салмақтарды түзету үдерісіне беру үшін (III.2.1.12) формуласына алдыңғы итерация салмақ өсімшесін түзетуге және локалдық минимумдарды еңсеруге мүмкіндік беретін кейбір λ инерциялық коэффициенті енгізіледі.

(III.2.1.12) формуласын жетілдірудің екі нұсқасы бар:

$$1) \Delta w_{ij}^{(q)}(t) = -\eta \cdot \delta_j^{(q)} \cdot y_i^{(q-1)} + \lambda \cdot \Delta w_{ij}^{(q)}(t-1), \quad (\text{III.2.1.13})$$

$$2) \Delta w_{ij}^{(q)}(t) = -\eta \cdot [(1-\lambda) \cdot \delta_j^{(q)} \cdot y_i^{(q-1)} + \lambda \cdot \Delta w_{ij}^{(q)}(t-1)],$$

(III.2.1.14)

мұнда λ – инерциялық коэффициент, t – ағымдағы итерация нөмері, η – оқыту жылдамдығының коэффициенті.

Сонымен, қателердің кері тарау алгоритмі арқылы ЖНЖ-ті оқытудың толық алгоритмі келесі қадамдар арқылы құрылады:

1-қадам. Барлық ЖНЖ синапстық байланыс салмақтарына кездейсоқ бастапқы мәндерді береміз. Осы орайда ЖНЖ кіру сигналдарының қандай да бір кездейсоқтық түрленуін жүзеге асырады және (III.2.1.1) қателіктер функциясының мәндері үлкен болуы мүмкін.

2-қадам. Оқыту жиынның ішінен кіріс векторларының бірін ЖНЖ кірісіне береміз және әрбір нейрондардың шығу мәнін есте сақтай отырып, ЖНЖ шығу мәнін есептейміз.

$$s_j^{(q)} = \sum_{i=1}^{I_j^{(q)}} w_{ij}^{(q)} \cdot y_i^{(q-1)}; \quad y_j^{(q)} = f(s_j^{(q)}); \quad y_j^{(0)} = x_r$$

мұнда $I_j^{(q)}$ – q қабатының j нейронының кіру саны; $j = 1, 2, \dots, J_q$; J_q – ығысуды беретін тұрақты шығыс жағдайы +1 нейронды ескерумен $(q-1)$ қабаттағы нейрондардың саны;

$q = Q, Q-1, \dots, 2, 1$; Q – ЖНЖ-дағы қабаттар саны;

$w_{ij}^{(q)} \cdot y_i^{(q-1)}$ – q қабаттағы j нейронның I -ші кірісі;

f – белсендірудің сигмоидалдық функциясы;

x_r – кіру векторының r -ші компоненті.

3-қадам. Алдымен Q -ші шығу қабаты үшін $\delta_j^{(Q)}$ мәнін (III.2.1.10) формуласы бойынша есептейміз және $\Delta w^{(Q)}$ синапстық байланыстар салмақтарының өзгеруін (III.2.1.12), (III.2.1.13) немесе (III.2.1.14) арқылы есептейміз. Сосын, барлық қалған q қабаттары үшін $\delta_j^{(q)}$ мәнін рекурренттік (III.2.1.11) формула көмегімен және $\Delta w^{(q)}$ салмақ өзгерісін (III.2.1.12), (III.2.1.13) немесе (III.2.1.14) көмегімен есептейміз.

4-қадам. Ағымдағы t – итерациясы үшін ЖНЖ синапстық байланыстардың салмақтарының барлығын келесі формула бойынша түзетеміз: $\Delta w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \Delta w_{ij}^{(q)}(t)$.

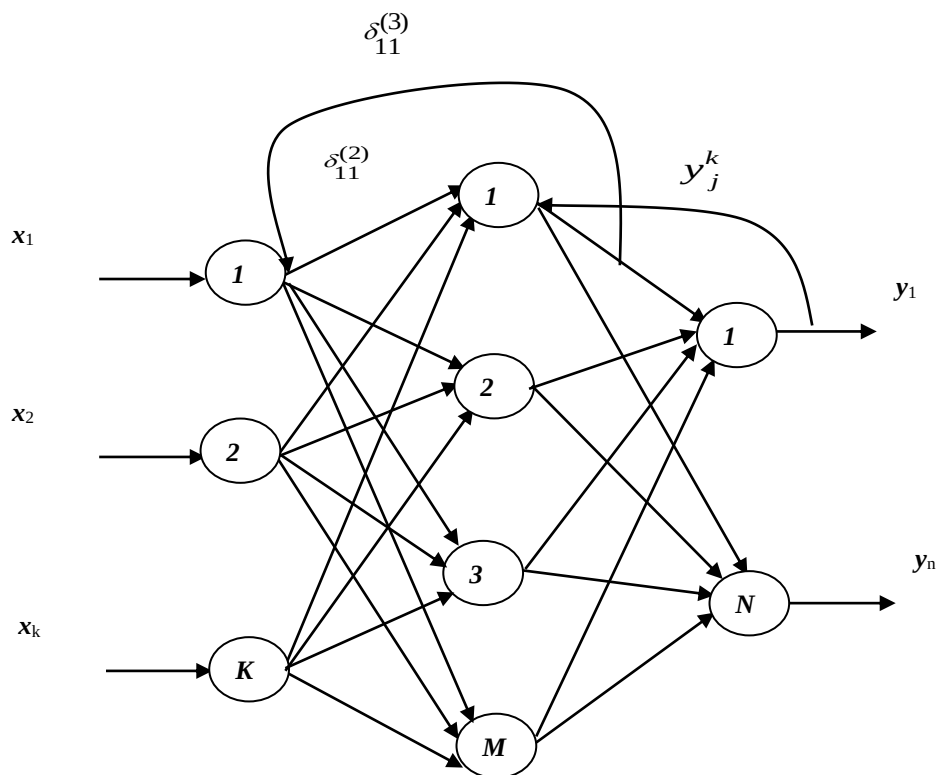
5-қадам. (III.2.1.1) мақсатты функциясының қателерінің мәнін есептейміз: егер, ол маңызды болса, онда екінші қадамға қайтып ораламыз, әйтпесе ЖНЖ-ны табысты оқытылды деп есептеп, жұмысты аяқтаймыз.

Басқаларды еске сақтау кезінде өзгелерін ұмытпау үшін, ЖНЖ 1-қадамда кезектесе кездейсоқ ретпен барлық оқытатын мысалдары ұсынылады.

$\Delta w_{ij}^{(q)} = -\eta \cdot \delta_j^{(q)} \cdot y_i^{(q-1)}$ өрнегінен $y_i^{(q-1)}$ шығу мәні нөлге ұмтылғанда, оқыту тиімділігі айтарлықтай төмендегендігін көруге болады. Екілік кіру векторлары кезінде орташа алғанда синапстық байланыстардың салмақтарының жартысы түзетілмейді, сондықтан нейрондар шығысының (0,1) мүмкін болатын мәндерінің аймағын логистикалық функцияларды қарапайым жетілдіру арқылы (0.5, 0.5) шегіне жылжыту керек.

Мысалы: $f(s) = \frac{1}{1+e^{-as}}$ –классикалық сигмоиды

$f(s) = -0.5 + \frac{1}{1+e^{-as}}$ түріне түрлендіреді. III.2.1-суретте қателерді кері тарату сұлбасы көрсетілген.



III.2.1-сурет. Қателердің кері таралу сұлбасы.

Қателердің кері тарату әдістерін көптеген табысты қолдануына қарамастан, ол ЖНЖ-ны өте ұзақ оқытады. Кейде ЖНЖ тіптен оқытылмауы да мүмкін. Ұзақ уақыт оқыту қадам мәндерін тиімсіз таңдау нәтижесі болуы мүмкін. Оқытудағы сәтсіздіктер әдетте келесі себептер бойынша туындайды: *ЖНЖ-ң дәрменсіздігі, локалдық минимумға және артықоқытуға түсу.*

III.2.1-тапсырмалар.

1. Қатені кері тарату әдісімен нейрондық желіні оқыту алгоритмін сипаттаңыз.
2. Шығу қабаты үшін дельта қатесін есептеу формуласын жазыңыз.
3. Шығу қабаты үшін синапстық байланыстардың салмағының өзгеру формуласын жазыңыз.
4. Қалған қабаттар үшін қателерді есептеу формуласын жазыңыз.
5. Барлық салмақтарды түзету формуласын жазыңыз.
6. Орташа квадраттық қатені есептеу формуласын жазыңыз.
7. Қателіктің кері тарату алгоритмі арқылы ЖНЖ-ны оқытудың толық алгоритмін сипаттаңыз.

III.2.1-сұрақтар.

1. Қателерді кері тарату әдісінің идеясы неде?
2. Қателерді кері тарату алгоритмінің жұмысының алдында не берілуі тиіс?
3. Қателердің критерийлері (мақсатты функция) ретінде не пайдаланылады?
4. Орташа квадраттық қателіктің функционалы қандай әдіспен азайтылады?
5. Қателерді кері тарату әдісімен нейрондық желіні оқыту үшін ең шұғыл түсу әдісі не үшін қажет?
6. Оқытудың әрбір қадамында градиенттік түсу әдісі нені қамтамасыз етеді?
7. Қателерді кері тарату әдісімен ЖНЖ оқытуда қандай жағдайларда сәтсіздік туындайды?

III.2.1-тестілер.

1. ЖНЖ оқыту әдістері оқытуға байланысты қандай үш іргелі қасиеттерді қарастырады?
А) сыйымдылықты, бейнелердің күрделілігін және есептеу күрделілігін;

В) анықтылықты, бейнелердің күрделілігін және сыйымдылық санын;

С) нәтижелілікті, бейнелердің жалпылығын және есептеу күрделілігін;

Д) салмақты, бейнелердің санын және есептеу күрделілігін;

Е) күрделілікті, бейнелердің анықтығын және есептеу күрделілігін.

2. Қателерді кері тарату әдісін қай жылы және кім әзірледі:

А) 1986 жылы Д. Румельхарт;

В) 1957 жылы Ф. Розенблатт;

С) 1950 жылы Н. Рочестр;

Д) 1949 жылы Н. Винер;

Е) 1943 жылы У. Мак-Каллок.

3. Шығу қабаты үшін дельта қатені есептейтін формула қайсысы?

А) $\delta_{ij}^{(Q)} = (y_{ij}^{(Q)} - d_{ij}^{(Q)}) \cdot \frac{dy_{ij}}{ds_{ij}};$

В) $w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \Delta w_{ij}^{(q)}(t);$

С) $\Delta w_{ij}^{(Q)} = -\eta \cdot \delta_{ij}^{(Q)} \cdot y_{ij}^{(Q-1)};$

Д) $\delta_{ij}^{(Q)} = \frac{dy_{ij}}{ds_{ij}};$

Е) $\delta_{ij}^{(Q)} = (y_{ij}^{(Q)} - d_{ij}^{(Q)}).$

4. Қатені азайту үшін ЖНЖ-нің Q шығару қабаты үшін синапстық байланыстар салмақтарының өзгеруін есептеуге арналған қандай формула?

А) $\Delta w_{ij}^{(Q)} = -\eta \cdot \delta_{ij}^{(Q)} \cdot y_{ij}^{(Q-1)};$

В) $w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \Delta w_{ij}^{(q)}(t);$

$$C) w_{ij}^{(Q)} = \delta_{ij}^{(Q)} \cdot y_{ij}^{(Q-1)};$$

$$D) \delta_{ij}^{(Q)} = \frac{dy_{ij}}{ds_{ij}};$$

$$E) \delta_{ij}^{(Q)} = (y_{ij}^{(Q)} - d_{ij}^{(Q)}).$$

5. Қателерді кері тарату әдісін басқаша қалай атайды?

- A) дельта-ережені жалпылау әдісі;
- B) дельта-ережені жекешелеу әдісі;
- C) дельта-ережені түрлендіру әдісі;
- D) дельта-ережені есептеу әдісі;
- E) дельта-ережені күрделілеу әдісі.

III.2.2. Больцман әдісімен нейрондық желілерді оқыту

Бұл тақырыпта Больцман әдісі бойынша ЖНЖ-ні оқыту сұрақтары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Больцман әдісі ақпараттық, теориялық және термодинамикалық принциптерден шығатын стохастикалық ережелерді қамтиды және қателік ретінде күйлерді корреляциялаудың екі режімдегі алшақтығы алынатын қателерді кері тарату әдісінің арнайы жағдайы ретінде қарастырылуы мүмкін.

Больцман әдісінің мақсаты көрінетін нейрондардың күйлері қалаулы ықтималдықты үлестіруді қанағаттандыратын салмақтарды баптау болады. Больцман әдісінде келесі алгоритм жүзеге асырылады:

1. Жасанды температураны беретін T айнымалысын анықтау және T -ға үлкен бастапқы мәнді беру.

2. Синапстық салмақтармен кірістердің жиынын ЖНЖ-не ұсыну және шығыстар мен мақсаттық функцияларды есептеу.

3. Синапстық байланыстар салмақтарының мәніне кездейсоқ өзгеріс енгізу және ЖНЖ-нің шығысы мен осы салмақтар мәндерінің жасалған өзгерістеріне сәйкес мақсаттық функциялардың өзгеруін санап шығу.

4. Егер мақсаттық функция азайса (жақсарса), онда осы салмақтар мәнінің өзгерісін сақтау, ал кері жағдайда, егер мақсаттық функция артса, онда осы өзгерістерді сақтау ықтималдығы Больцман үлестірімінің көмегімен есептеледі:

$$P(s) = e^{-\frac{s}{kT}},$$

мұнда $P(s)$ – s -тің мақсаттық функциядағы өзгеру ықтималдығы, k – есепке байланысты таңдап алынған Больцман тұрақтысына ұқсас тұрақты, T – жасанды температура.

Бірқалыпты таратылудан 0 мен 1 сандар аралығындағы кездейсоқ r саны таңдап алынады. Егер $P(s) > r$ болса, өзгеріс

сақталады, ал кері жағдайда салмақтың шамасы алдыңғы мәнге қайтарылады. Бұл жүйеге мақсаттық функцияны нашарлататын бағытқа кездейсоқ қадам жасауға мүмкіндік береді, сонымен қатар кез-келген кішкентай қадам мақсаттық функцияны арттыратын локалды минимумдардың шығуына мүмкіндік береді.

Больцман оқытуын аяқтау үшін мақсаттық функцияның рұқсат етілетін ең төменгі мәніне жетпейінше T температураны ақырындап кішірейте отырып, ЖНЖ-нің әр салмақтары үшін 3-ші және 4-ші қадамды қайталайды. Осы кезде басқа кіру векторы ұсынылады және оқыту үдерісі қайталанады. ЖНЖ оқыту жиындарының барлық векторларында, мақсаттық функция олардың барлығы үшін қол жетімді болғанға дейін мүмкін болатын қайталанулармен оқытылады.

3-ші қадамдағы салмақтың өзгерістері әр түрлі тәсілдерімен анықтала алады. Мысалы, жылулық жүйесіне байланысты w салмағының өзгеруі Гаусс үлестіріліміне сәйкес есептелінеді:

$$P(w) = e^{-\frac{w^2}{kT^2}},$$

мұнда $P(w)$ – w шамасына салмақтың өзгеру ықтималдығы, k – тұрақты, T – жасанды температура

Бізге салмақ өзгерісінің ықтималдығы емес Δw салмақ өзгеруінің шамасы керек болатындықтан, монте-Карло әдісі келесі түрде қолданылуы мүмкін:

1) $P(w)$ -ға сәйкес келетін, кумулятивті ықтималдықты табу. Бұл 0-ден w -ға дейінгі $P(w)$ -дан алынған интеграл. Бұл жағдайда $P(w)$ аналитикалық тұрғыдан интегралдана алмайтындықтан, оны сандық жолмен интегралдау керек, ал нәтижені кестемен ұсыну керек.

2) 0 мен 1 интервалындағы бірқалыпты үлестіруден кездейсоқ санды тандап алу керек. Осы шаманы $P(w)$ мәні ретінде қолдана отырып, кестеден синапстық байланыстар салмақтарының өзгеру шамасы үшін сәйкес мәнді тауып алу қажет.

Глобалдық минимумға жинақталатындай, температураның төмендеу жылдамдығы уақыт логарифміне кері пропорционал болатындығы көрсетілген Больцман машинасының қасиеттері көптеген жұмыстарда зерттеледі. Мұндай жүйедегі салқындату келесі формуламен көрсетіледі:

$$T(t) = \frac{T_0}{\log(1+t)}$$

мұнда $T(t)$ – уақыт функциясы болатын жасанды температура, T_0 – бастапқы жасанды температура, t –жасанды уақыт.

Бұл нәтиже салқындау (есептеу) жылдамдығы өте баяу екендігін болжайды. Сондықтан, Больцман машиналары оқыту үшін көп уақытты талап етеді.

III 2.2-тапсырмалар.

1. Больцман әдісінің жүзеге асырылуы алгоритмін сипаттаңыз.
2. Больцман үлестірімі үшін формуланы жазыңыз.
3. Гаусс үлестірімі үшін формуланы жазыңыз.
4. Үлкен уақыт ресурстарының Больцман машинасын талап ету себебін түсіндіріңіз.
5. Больцман оқытуының аяқталу себебін көрсетіңіз.
6. Оқытушы жиындарындағы барлық векторында ЖНЖ оқыту шарттарын көрсетіңіз.
7. Локалды минимумнан шығу шарттарын көрсетіңіз.

III 2.2-сұрақтар.

1. Больцман әдісі нені ұсынады?
2. Больцман тәсілін қалай қарастыруға болады?
3. Больцман тәсілінің мақсаты неде?
4. Больцман үлестірімін есептеу нені береді?
5. Гаусс үлестірімін қайда қолдануға болады?
6. Монте–Карло әдісін неге және қалай қолдануға болады?
7. Локалды минимумнан шығу мүмкіндігіне немен жетуге болады?

III.2.2-тестілер.

1. Больцман әдісінің мақсаты не?

A) көрінетін нейрондардың жағдайы қаланатын ықтималдықты үлестіруді қанағаттандыратын салмақтарды баптау;

B) анықтылықты, бейнелердің күрделілігін және сыйымдылық санын;

C) нәтижелілікті, бейнелердің жалпылығын және есептеу күрделілігін;

D) салмақты, бейнелердің санын және есептеу күрделілігін;

E) күрделілікті, бейнелердің анықтығын және есептеу күрделілігін.

2. Больцман үлестірімін есептейтін формула қайсысы?

A) $P(s) = e^{-\frac{s}{kT}}$;

B) $P(s) = e^{-\frac{s}{T}}$;

C) $P(s) = e^{-\frac{s}{k}}$;

D) $P(s) = e^{-\frac{1}{kT}}$;

E) $P(s) = k * e^{-\frac{s}{kT}}$.

3. Гаусс үлестірімін есептейтін формула қайсысы?

A) $P(w) = e^{-\frac{w^2}{kT^2}}$;

B) $P(s) = e^{-\frac{s}{kT}}$;

C) $P(w) = e^{-\frac{w^5}{T^2}}$;

$$D) P(s) = e^{-\frac{1}{kT}};$$

$$E) P(s) = k * e^{-\frac{s}{kT}}.$$

4. Салқындау температурасын есептейтін формула қайсысы?

$$A) T(t) = \frac{T_0}{\log(1+t)};$$

$$B) P(s) = e^{-\frac{s}{kT}};$$

$$C) T(t) = \frac{5T_0}{\log(1+t)};$$

$$D) T(t) = \frac{T_0}{\log(2+3t)};$$

$$E) P(s) = k * e^{-\frac{s}{kT}}.$$

5. 0 мен 1 аралығындағы бірқалыпты үлестіруден кездейсоқ r саны тандап алынған. Егер $P(s) > r$ болса, онда не болады?

A) өзгеріс сақталады;

B) салмақтың шамасы алдыңғы мәнге қайтарылады;

C) қадам жасауға мүмкіндік береді;

D) өзгеріс сақталмайды;

E) оқыту аяқталады.

III.2.3. Коши әдісімен нейрондық желілерді оқыту

Бұл тақырыпта Коши әдісімен ЖНЖ оқыту қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Егер нейрондық желілерді III.2.2 бөлімде келтірілген оқыту әдісінде қадамның шамасын есептеу кезінде Больцман үлестірімінің орнына Коши үлестірімін пайдаланса, онда оқыту қадамын үлкейтіп, оқуды тездетуге болады. Шынында Коши үлестірімі шексіз (белгісіз) дисперсияға ие. Осындай жай өзгерістің көмегімен температураның максималды төмендеу жылдамдығы нейрондық желіні Больцман әдісі бойынша оқыту алгоритмінде болғандай логарифмге емес, сызықтық шамаға кері пропорционалды болады. Яғни, бұл тәуелділік келесідегідей болады:

$$T(t) = \frac{T_0}{1+t}$$

Коши үлестірімі келесі түрде болады:

$$P(x) = \frac{T(t)}{T(t)^2 + x^2},$$

мұнда $P(x)$ – x көлемінің ықтималдық қадамы.

Осындай алмастыру оқыту уақытын төмендетуге мүмкіндік береді және III.2.3.1- суреттегідей Коши үлестірімін Больцман үлестірімімен салыстырсақ, ол өте ұзын «жылу сыйымдылыққа» ие, осылайша үлкен қадам ықтималдығын үлкейтуге болады.



III.2.3.1-сурет Коши үлестірімі мен Больцман үлестірімдері.

(III.2.3.2) тендеуінде $P(x)$ үлестірім функциясы интегралдануы мүмкін, кейін x шамасы есептелінеді:

$$x_c = \rho T(t) \operatorname{tg}((P(x))), \quad (\text{III.2.7.3})$$

мұнда ρ – оқылу жылдамдығының коэффициенті, x_c – салмақ өзгерісі.

Енді Монте-карло әдісін қолдану өте жеңіл. Бұл жағдайда x -ті табу үшін біркелкі ашық $(-\pi/2, \pi/2)$ интервалында – тангенс функциясының шектеулігінде кездейсоқ сан алынады. Ол (III.2.3.3) формуласына $P(x)$ ретінде қойылады және ағымдағы температура көмегімен қадам шамасы есептелінеді.

Коши әдісі арқылы қол жеткізілген жақсартуға қарамастан, оқыту уақыты әлі де өте үлкен болуы мүмкін. Бұл үдерісті жылдамдату үшін термодинамикадағы жасанды жылу сыйымдылық әдісін қолдануға болады. Сондағы температураның төмендеу жылдамдығы оқыту үдерісінде есептелетін жасанды «жылу сыйымдылыққа» сәйкес өзгереді. Металды балқыту кезінде энергия деңгейінің дискретті өзгеруімен байланысты фазалық ауысулар болады. Әрбір фазалық ауысу кезінде жылу сыйымдылығы деп аталатын шаманың бірден өзгеруі орын алады.

ЖНЖ оқыту үдерісінде аналогиялық фазалардан өтеді. Фазалық шекке өткен кезде жасанды жылу сыйымдылық секіrmелі өзгеруі мүмкін. Бұл жылу сыйымдылық арнайы функциядағы температураның орташа жылдамдығының өзгеруі арқылы анықталады. Жәшіктегі шариктің мысалында бастапқы қатты сілкінуі мақсатты функцияның орташа ұлғаюын температураның кішкентай өзгерісіне байланыссыз жасайды, яғни жылу сыйымдылық тұрақтыға жуық. Соған ұқсас жүйе өте төмен температурада минимум нүктесінде қатады, сол үшін жылу сыйымдылық тағы да тұрақтыға жуық. Мақсатты функцияны жақсарту болмағандықтан, әрбір осы аймақтағы температураға үлкен өзгерістер жүретіндігі белгілі.

Сын температурада кішігірім өзгеріс мақсатты функцияның орташа мәнін үлкен өзгеріске алып келеді. Шарикке ұқсас алсақ,

«температурада» A -дан B -ға өту үшін шарик жеткілікті орташа энергияға ие болуы керек, бірақ B -дан A -ға жеткіліксіз ауысу кезінде арнайы функцияның орташа көлемі жылдам өзгеріске ұшыратады. Бұл сыни нүктелерде температураның алгоритмі өте жәй өзгеруі керек, пайда болған локалды минимумдағы A нүктесінде жүйе қатып қалмау керек. Сыни температура жасанды жылу сыйымдылықтың тез төмендеп кетуіне, яғни мақсатты функциямен қатар температураның өзгеруінің орташа жылдамдығының төмендеп кетуіне алып келуі мүмкін. Сын температураға жеткен кезде температураның өзгеру жылдамдығы глобалды минимумға сәйкесікті кепілдендіру үшін баяулауы керек. Қалған барлық температураларда тәуекелсіз температураның төмендеуінің аса жоғары жылдамдығы қолданыла беруі керек. Бұл оқыту уақытының төмендеуіне алып келеді.

III.2.3-тапсырмалар.

1. Температураны төмендету жылдамдығы үшін Коши әдісіндегі формуланы жазыңыз.

2. Оқыту уақытын тез төмендететін Коши үлестірімі үшін формуланы жазыңыз.

3. Жасанды жылу сыйымдылығы тез өзгере алатын жерді көрсетіңіз.

III.2.3-сұрақтар.

1. Коши әдісімен нейрондық желіні оқытудың басты идеясы неден тұрады?

2. Қандай құралдар арқылы Коши әдісімен нейрондық желіні оқытуды жақсартуға болады?

3. Глобалды минимумға сәйкестікті кепілдендіру үшін температура өзгерісінің жылдамдығы қашан баяулауы қажет?

III.2.3-тестілер.

1. Коши үлестірімі қандай формуламен есептеледі?

А) $P(x) = \frac{T(t)}{T(t)^2 + x^2};$

В) $P(x) = T(t)^2 + x^2;$

$$C) P(x) = \frac{T(t)}{T(t)^2 + x^2} + T(t)^2;$$

$$D) P(x) = \frac{1}{T(t)^2 + x^2};$$

$$E) P(x) = \frac{T(t)}{T(t)^2 + x^2} * x^2.$$

2. Коши үлестіріміндегі $P(x)$ нені көрсетеді?

- A) x шамасының ықтималдық қадамы;
- B) x шамасының оқыту жылдамдығын;
- C) салмақ өзгерісін;
- D) жасанды температураны;
- E) x шамасының өзгеруін.

3. Коши үлестірімі қандай мүмкіндікке ие?

- A) шексіз (белгісіз) дисперсияға;
- B) шекті (белгілі) дисперсияға;
- C) салмақ өзгерісіне;
- D) жасанды температураға;
- E) шексіз қадамға.

4. Коши үлестірімін Больцман үлестірімімен салыстырғанда қандай айырмашылығы бар?

- A) өте ұзын «жылу сыйымдылыққа» ие;
- B) өте қысқа «жылу сыйымдылыққа» ие;
- C) өте ұзын «жылу температурасына» ие;
- D) жасанды температураға ие;
- E) өте баяу «сыйымдылыққа» ие.

5. Жасанды жылу сыйымдылықтың тез төмендеп кетуіне, яғни мақсатты функциямен қатар температураның өзгеруінің орташа жылдамдығының төмендеп кетуіне алып келетін не?

- A) сыни температура;
- B) жылу сыйымдылық;
- C) жылу температурасы;
- D) жасанды температура;
- E) баяу сыйымдылық.

III.2.4. Хебб әдісімен нейрондық желілерді оқыту

Бұл тақырыпта өздігінен ұйымдасқан Хебб әдісімен ЖНЖ оқыту

сұрақтары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

ЖНЖ мұғалімсіз оқыту әдісінде синапстық байланыс салмақтарын құру үдерісі өздігінен жүреді. Мұнымен қатар, ЖНЖ өзінің құрылымын өзгерте отырып өздігінен ұйымдастырыла алады. Синапстық байланыстардың салмақтарын құру нейронда қолжетімді ақпараттар, яғни оның жағдайы және бар салмақтары негізінде ғана жүргізілуі мүмкін. Осы пікірлерге негізделіп жүйке жасушаларының өздігінен ұйымдастырылу принципіне ұқсас Хебб әдісі құрылды.

Хебб әдісі мұғалімсіз ЖНЖ-ні оқытудың ең көне ережесі болады, ол келесі нейрофизиологиялық бақылауларға негізделген: егер синапстың екі жағындағы нейрондар бір уақытта және ұдайы түрде белсендендірілсе, онда синапстың салмағы өседі және тек осы синапспен байланысқан белсендіру нейрондардан тәуелді болады. Сонымен қоса, әрбір нейрон аралық байланыстар салмағының өзгеруі осы синапсты құрайтын белсендіру нейрондардан ғана тәуелді болады, ол оқыту алгоритмінің жүзеге асырылуын қысқартады.

Хебб әдісімен ЖНЖ оқыту үдерісінде салмақты арттыру келесі түрде жазылған нейрондарды қоздыру, беру және қабылдау деңгейлерінің көбейтіндісімен анықталады:

$$w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)}, \quad (\text{III.2.4.1})$$

мұнда $y_i^{(q-1)}$ – $(q-1)$ қабатының I нейронның шығысы және j нейронның кірісі, $y_j^{(q)}$ – q қабатының j нейронның шығысы, $w_{ij}^{(q)}(t)$ и $w_{ij}^{(q)}(t-1)$ – q қабатының I нейронынан j нейронына сәйкесінше t және $t-1$ итерациясындағы синапс салмағы, η – оқыту жылдамдығының коэффициенті; $q = Q, Q-1, \dots, 2, 1$.

Хебб әдісін оқыту үдерісінде ЖНЖ-да қоздырылған нейрондардың арасындағы байланыстар күшейтіледі, оны (ІІІ.2.4.1) формуладан көруге болады.

Келесі қатынаспен анықталатын оқытудың дифференциалданған Хебб әдісі де бар:

$$w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - y_i^{(q-1)}(t-1)] \cdot [(y_j^{(q)}(t) - y_j^{(q)}(t))], \quad (\text{ІІІ.2.4.2})$$

мұнда $y_i^{(q-1)}(t)$ және $y_i^{(q-1)}(t-1)$ – сәйкесінше t және $t-1$ итерациясындағы $q-1$ қабатының I нейронның шығысы, $y_j^{(q)}(t)$ және $y_j^{(q)}(t-1)$ – сәйкесінше t және $t-1$ итерациясындағы q қабатының j нейронының шығысы, η – оқыту жылдамдығының коэффициенті.

Хебб оқытуының дифференциалды әдісінде (ІІІ.2.4.2) көрініп отырғандай, шығысы ұлғаю жағына қарай аса серпінді түрде өзгерген нейрондарды біріктіретін синапстар барлығынан гөрі күштірек оқытылады.

Жоғарыда келтірілген (ІІІ.2.4.1) және (ІІІ.2.4.2) формулаларын қолданып Хебб ережелерінің көмегімен ЖНЖ-ны оқытудың толық алгоритмі мына түрде болады:

1-қадам. Инициализациялау: ЖНЖ-дің барлық синапстық салмақтарына ең үлкен кездейсоқ бастапқы мәндері меншіктеледі.

2-қадам. Салмақталған қосындыны есептеу: ЖНЖ кірісіне кіріс үлгісі беріледі және салмақ барлық қабаттар бойынша тура үлестіру принципіне сәйкес таратылады және әрбір нейрон үшін келесі функция қолданылатын шығу мәні есептелінеді:

$$\sum_{i=1}^{I_j^{(q)}} w_{ij}^{(q)} \cdot x_i^{(q)} = s_j^{(q)}; \quad f(s_j^{(q)}) = y_j^{(q)},$$

мұнда $s_j^{(q)}$ – q қабатының j нейронының салмақталған қосындысы;

$I_j^{(q)}$ – q қабатының j нейронның кірістер саны; $j=1, 2, \dots, J_q$;

J_q – q қабаттағы нейрондар саны; $q=1, 2, \dots, Q$;

Q – ЖНЖ-гі қабаттар саны;

$w_{ij}^{(q)} \cdot x_i^{(q)}$ – q қабатындағы j нейронның I -ші кірісі;

f – белсендіру функциясы;

$y_j^{(q)}$ – q қабатының j нейронының шығысы;

3-қадам. *Салмақтардың реттелуі.* (III.2.4.1) немесе (III.2.4.2) көмегімен алынған ЖНЖ-ң шығыс мәндерінің негізінде $w_{ij}^{(q)}(n)$ синапстық байланыстар салмақтарын өзгерту жүргізіледі.

4-қадам. *Оқығандықты тексеру.* Егер ЖНЖ-нің шығыс мәндері берілген дәлдікпен тұрақталған болса, онда ЖНЖ сәтті оқытылды деп есептейміз және жұмысты аяқтаймыз. Әйтпесе, кезекпен барлық кірістегі бейнелерді ұсына отырып, ЖНЖ-ң шығыс мәндері тұрақты болғанға дейін 2-ші қадамнан бастап алгоритмді орындауды қайталаймыз.

Айта кететін мәселе, кіріс үлгілерінің әрбір сыныбына жауап қату түрі алдын-ала белгісіз болады және инициализациялау кезеңінде салмақтарды кездейсоқ таратуға шартталған шығу қабатының нейрондарының еркін байланысынан тұрады. Сонымен қатар, ЖНЖ-і бір сыныпқа апара отырып ұқсас үлгілерді жалпылауға қабілеті бар. Оқытылған ЖНЖ-ны тестілеу шығу қабатында сыныптың топологиясын анықтау мүмкіндігін береді. Оқытылған ЖНЖ-ге пікірлерді қолайлы ұсынуға келтіру үшін ЖНЖ-ны бір қабатпен толықтыруға болады, мысалы, бір қабатты персептронды оқыту алгоритмі бойынша ЖНЖ-нің шығыс реакциясын талап етілетін үлгіні бейнелеуге мәжбүрлеу керек.

III.2.4-тапсырмалар.

1. Хебб әдісімен ЖНЖ-ді оқыту сүйенетін бақылау түрлерін атап шығыңыз.
2. Хебб әдісімен ЖНЖ-ді оқытуда салмақтарды арттыру формулсын жазыңыз.
3. Хебб оқытуының дифференциалдық әдісінің формуласын жазыңыз.
4. Хебб ережесінің көмегімен ЖНЖ-ны оқытудағы толық алгоритмін сипаттаңыз.

5. Хебб әдісіндегі салмақталған қосындыны есептеу формуласын жазыңыз.

III.2.4-сұрақтар.

1. Хебб әдісі қандай принципке негізделіп құрылған?
2. Хебб әдісі бойынша ЖНЖ-ді оқыту үдерісінде не күшейтіледі?
3. Хебб оқытылуында дифференциалдық әдісте қандай синапстар күштірек оқытылады?
4. Хебб әдісінде оқытылғандықты тексеру қалай жүргізіледі?
5. Оқытылған ЖНЖ-ге пікірлерді қолайлы ұсынуға келтіру үшін нені қосуға болады?

III.2.4-тестілер.

1. Мұғалімсіз ЖНЖ-ні оқытудың ең көне ережесі болып табылатын әдістің қайсысы?
 - A) Хебб әдісі;
 - B) Коши әдісі;
 - C) Гаус үлестірімі;
 - D) Больцман үлестірімі;
 - E) Хопфилд әдісі.
2. Хебб ережелерінің көмегімен ЖНЖ-ны оқытудың толық алгоритміндегі 1-қадамда қандай үдеріс жүргізіледі?
 - A) Инициализациялау;
 - B) Салмақталған қосындыны есептеу;
 - C) Салмақтардың реттелуі;
 - D) Оқытылғандықты тексеру;
 - E) Нәтижені тексеру.
3. Хебб ережелерінің көмегімен ЖНЖ-ны оқытудың толық алгоритміндегі 2-қадамда қандай үдеріс жүргізіледі?
 - A) Салмақталған қосындыны есептеу;
 - B) Инициализациялау;

- С) Салмақтардың реттелуі;
- Д) Оқытылғандықты тексеру;
- Е) Нәтижені тексеру.

4. Хебб әдісіндегі алгоритмде әрбір нейрон үшін шығу мәні есептелетін формула қайсысы?

A) $\sum_{i=1}^{I_j^{(q)}} w_{ij}^{(q)} \cdot x_i^{(q)} = s_j^{(q)}; f(s_j^{(q)}) = y_j^{(q)};$

B) $\sum_{i=1}^{I_j^{(q)}} w_{ij}^{(q)} = s_j^{(q)}; f(s_j^{(q)}) = y^{(5)};$

C) $\sum_{i=1}^{I_j^{(q)}} x_i^{(q)} = s_j^{(q)}; f(s) = y;$

D) $w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1);$

E) $w_{ij}^{(q)}(t) = (t-1) + \eta \cdot y_i^{(q-1)} \cdot y_j^{(q)}.$

5. Хебб әдісінде қай кезде ЖНЖ сәтті оқытылды деп есептеліп, жұмыс аяқталады?

A) Егер ЖНЖ-нің шығыс мәндері берілген дәлдікпен тұрақталған болса;

B) Егер ЖНЖ-нің кіріс мәндері берілмесе;

C) Егер ЖНЖ-нің кіріс мәндері берілген дәлдікпен тұрақталмаса;

D) Егер ЖНЖ-нің барлық мәндері берілсе;

E) Егер ЖНЖ-нің мәндері берілген дәлдікпен аяқталса.

III.2.5. Кохонен әдісімен нейрондық желіні оқыту

Осы параграфта Кохоненның ЖНЖ оқыту әдісі қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Кохонен басқа “катені түзету әдісі” деп аталатын мұғалімсіз оқыту әдісін ұсынды және ол алдындағы итерацияның мәндеріне негізделген синапстық байланыс салмақтарының жетілдіруін қарастырады:

$$w_{ij}^{(q)}(t) = w_{ij}^{(q)}(t-1) + \eta \cdot [y_i^{(q-1)}(t) - w_{ij}^{(q)}(t-1)], \quad (\text{III.2.5.1})$$

мұнда $w_{ij}^{(q)}(t)$ және $w_{ij}^{(q)}(t-1)$ –қабаты t және $t-1$ итерациясында q қабатының I нейронынан j нейронына дейінгі синапс салмағы; η – оқудың жылдамдық коэффициенті; $y_i^{(q-1)}(t)$ – t ; q ; $q = 1, 2, \dots, Q$ итерациясында $(q-1)$ қабатының I нейронның шығысы; Q – ЖНЖ-дегі қабаттың саны.

Жоғарыда көрсетілген (III 2.5.1) формулада алдыңғы $y_i^{(q-1)}(t)$ қабаттың нейрондары мен синапстық салмағының шығыстарынан келетін нейронның шығу сигналдары арасындағы айырмашылықты минималдауға әкелетіндігін білдіреді. ЖНЖ оқыту үдерісі нейронаралық синапстық байланыс салмақтарының бастапқы мәндерінің коррекциясынан тұрады.

Оқытудың толық алгоритмі Хебб әдісіндегі сияқты құрылымға ие, бірақ 3 қадамда барлық қабаттардан нейрон таңдалады. Оның синапстық байланыстарының мәні кіріс үлгісіне максималды сай келеді және (III.2.5.1) формула бойынша салмақтардың бапталуы тек сол үшін жүргізіледі.

Аккредитация қабаттың қалған барлық нейрондарының жұмыс істеу қалпының тежелуімен және таңдалған нейронның қанығуын енгізумен болады. Мұндай нейронды таңдау, мысалы, шығыс мәндерінің векторымен синапстық байланыстар салмақтар векторының скалярлық көбейтіндісін есептеумен жүргізілуі мүмкін. Масималды көбейтінді жеңген нейронды береді.

$y_i^{(q-1)}(t)$ алдыңғы қабат шығысы және синапстық салмақтармен келіп түсетін нейрондардың кіру сигналдары арасындағы айырмашылықты азайту негізінде оқытудың басқа нұсқасы осы екеуінің арасындағы қашықтықты есептеуге алып келеді:

$$D_j = \sqrt{\sum_{i=0}^{n-1} (y_i^{(q-1)} - w_{ij})^2}, \quad (\text{III.2.5.2})$$

мұнда n – вектордың өлшемі, j – q қабатындағы нейронның индексі, I – $(q-1)$ қабатындағы нейрондар бойынша қосқыш индексі, w_{Ij} – нейрондарды біріктіретін синапстың салмағы; $(q-1)$ қабатындағы нейрондардың шығысы q қабаты үшін кіріс мәндері болады.

(III.2.5.2.) формуладағы түбірді алмауға болады, себебі әр түрлі D_j -дің салыстырмалы бағасы ғана маңызды.

Бұл жағдайда ең кіші ара қашықтықпен нейрон «жеңеді». Кей жағдайда жиі аккредитацияны алатын нейрондар қабаттарды барлық нейрондарының «құқықтарын теңдестіру» үшін қарастырылудан мәжбүрлі түрде шығарылады. Мұндай алгоритмнің қарапайым нұсқасы жаңа ғана жеңген нейронды тежеуден тұрады. Кохонен алгоритмі бойынша оқытуды қолданған кезде кіріс бейнелерін, сондай-ақ синапстық байланстардың салмақтық мәндерін инициализация және нормализация кезеңінде нормалдау практикасы пайдаланылады.

$$x_i = \frac{x_i}{\sqrt{\sum_{j=0}^{n-1} x_j^2}}, \quad (\text{III.2.5.3})$$

мұнда x_I – кіріс бейнесі векторының немесе синапстық байланыс салмақтарының векторының I - компоненті, n – оның өлшемділігі.

Бұл оқыту үдерісінің ұзақтығын қысқартады.

Кездейсоқ мәндердің синапстық байланыстар салмағының инициализациясы тығыз таратылған кіріс бейнелеріне сәйкес келетін әртүрлі сыныптардың бірігуіне немесе керісінше, бір сыныптағы жақын бейнелер жағдайында қосымша ішкі сыныптарға

бөлінуіне алып келеді. Мұндай жағдайды болдырмау үшін дөңес комбинация әдісі қолданылады. Оның мағынасы кіріс нормаланған бейнелер түрлендіруге түседі:

$$x_i = \eta(t) \cdot x_i + (1 - \eta(t)) \cdot \frac{1}{\sqrt{n}}, \quad (\text{III.2.5.4})$$

мұнда x_I – кіріс бейнесінің I - компонент, n – оның компонентінің жалпы саны, $\eta(t)$ – оқыту үдерісінде нөлден бірге дейін қзгеретін коэффициент. Нәтижесінде ЖНЖ-дің бастапқы кірісінде біркелкі бейнелер беріледі, ал уақыт ағынымен олардың барлығы бастапқыға бірігеді. Синапстық байланыс салмағы

$$w_0 = 1 / \sqrt{n}$$

шамасына тең инициализация қадамында орнатылады. Мұнда n – инициализацияланатын қабаттың нейрондары үшін вектор өлшемі;

Жоғарыда қарастырылған әдіс негізінде нейрондардың баейімделетін синапстық салмақтары бар ЖНЖ құрылады.

III.2.5-тапсырмалар.

1. Кохонен әдісі бойынша нейрондық желіні оқытудың басқа нұсқасының формуласын жазыңыз.

2. Синапстық байланыс салмақтарының бастапқы мәндерінің нормалдау формуласын жазыңыз.

3. Инициализациялау қадамында синапстық байланыс салмақтарының мәндерін орнатудың формуласын жазыңыз.

III.2.5-сұрақтар.

1. Кохонен әдісінің идеясы неден тұрады?

2. Кохонен әдісі бойынша нейрондық желіні оқыту неден тұрады?

3. Кохонен әдісі бойынша оқытудың толықтай алгоритмі қандай?

4. Кохонен әдісі бойынша нейрондық желіні оқытудың екінші нұсқасында қандай нейрон жеңімпаз болып саналады?

III.2.5-тестілер.

1. «Қатені түзету әдісі» деп аталатын мұғалімсіз оқыту әдісін ұсынған кім?

- A) Кохонен;
- B) Хебб;
- C) Гаусс;
- D) Больцман;
- E) Хопфилд.

2. Кохонен әдісінің Хебб әдісінен айырмашылығы неде?

- A) 3 қадамда барлық қабаттардан нейрон таңдалады;
- B) 3 қадамда барлық қабаттардан салмақ таңдалады;
- C) 2 қадамда барлық қабаттардан нейрон таңдалады;
- D) 1 қадамда барлық қабаттардан нейрон таңдалады;
- E) 4 қадамда барлық қабаттардан салмақ таңдалады.

3. $D_j = \sqrt{\sum_{i=0}^{n-1} (y_i^{(q-1)} - w_{ij})^2}$ формуласындағы w_{ij} нені көрсетеді?

- A) нейрондарды біріктіретін синапстың салмағы;
- B) нейрондардың шығысы;
- C) вектордың өлшемі;
- D) нейронның индексі;
- E) нейронның салмағы.

4. $x_i = \eta(t) \cdot x_i + (1 - \eta(t)) \cdot \frac{1}{\sqrt{n}}$ формуласындағы $\eta(t)$ не?

- A) оқыту үдерісінде нөлден бірге дейін өзгертін коэффициент;
- B) нейрондарды біріктіретін синапстың салмағы;
- C) вектордың өлшемі;
- D) нейронның индексі;
- E) нейрондардың шығысындағы коэффициент.

5. $w_0 = 1 / \sqrt{n}$ формуласын нені есептейді?

- A) синапстық байланыс салмағын;
- B) вектордың өлшемін;
- C) оқыту үдерісінде нөлден бірге дейін өзгертін коэффициентті;
- D) нейронның индексін;
- E) нейрондардың шығысындағы коэффициентті.

III.2.6. Нейрондық желіні жарыс әдісімен оқыту

Бұл тақырыпта ЖНЖ-ны жарыс әдісімен оқыту қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

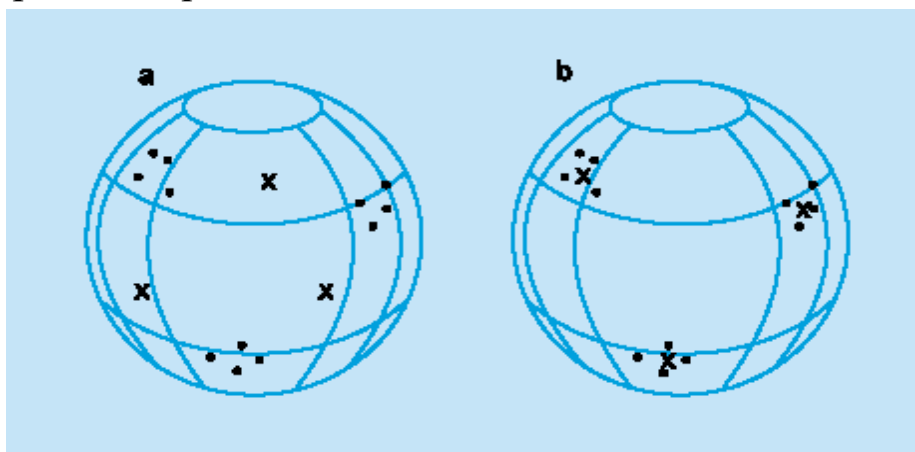
Жарыспен оқыту әдісінде, шығу нейрондарының жиыны бір уақытта белсендірілетін Хебб әдісінен айырмашылығы, шығу нейрондары белсенді болу үшін өзара жарысады. Әдетте бұл құбылыс "жеңімпаз бәрін алады" деген ережемен белгілі. Осындай оқытудың биологиялық нейрондық желілерде өз орны бар. Жарыс құралымен оқыту кіру деректерін кластерлеуге мүмкіндігін береді. Осындай мысалдар нейрондық желімен корреляцияға сәйкес топталады және бір элементпен бейнеленеді.

Бұл жарыса оқыту үдерісінде ұтып алған нейрон мен оның айналасында орналасқан нейрондардың синапстық байланыстарының салмақтарын түрлендіру жүзеге асырылады. Оқыту кезінде тек қана "жеңген" нейронның салмақтары ғана түрленеді. Бұл ереженің нәтижесіне кіру мысалына сәл жақындайтын ЖНЖ-де сақталған үлгіні (ұтқан нейронның салмақ векторын) өзгерту есебінен қол жеткізіледі. Ең үлкен шығу сигналы бар нейрон жеңімпаз болып жарияланады да, өзінің бәсекелестерін тежеу және көршілерін қоздыру мүмкіндігіне ие болады. Нейрон-жеңімпаздың шығу сигналы пайдаланылады және тек оған және оның көршілеріне өздерінің байланысу салмақтарын келесі формула бойынша түзетуге рұқсат етіледі:

$$\Delta w_{ij}(t+1) = w_{ij}(t) + \gamma[x_i - w_{ij}(t)]$$

III.2.6.1-суретте жарыс әдісімен оқытудың геометриялық бейнесі берілген. Кіру векторлары нормаланған және сфера бетінде нүктелермен бейнеленген. Векторлар салмағы үш нейрон үшін кездейсоқ мәндермен инициалданған. Олардың бастапқы мәндерін (оқытуға дейінгі) және соңғы мәндерін (оқытудан кейінгі) тиісінше III.2.6.1(а) және III.2.6.1(б)-суретте *X-пен* белгіленген. Мысалдардың үш тобының әрқайсысы, салмақтар векторы

табылған топтың ауырлық орталығына бапталған шығу нейрондарының бірімен табылған.



III.2.6.1-сурет Жарыс әдісімен оқыту мысалы:
(а) оқытуға дейін; (б) оқытудан кейін.

Егер оқыту жылдамдығының параметрі 0-ге тең емес болса, онда ЖНЖ-нің оқытылуы ешқашан тоқтамайтынын байқауымызға болады. Кейбір кіру үлгісі оқыту үдерісінің келесі итерациясында басқа шығу нейронын белсенді ете алады. Бұл оқыту жүйесінің тұрақтылығы туралы сұрақ қояды. Жүйе тұрақты болып саналады, егер оқытатын тандаманың ешбір мысалы оқыту үдеріс итерациясының ақырлы санынан кейін өзінің категорияға жататындығын өзгертпесе. Тұрақтылыққа жету тәсілдерінің бірі оқыту жылдамдығының параметрін 0-ге дейін біртіндеп азайтудан тұрады. Алайда, оқытуды бұл жасанды тежеу созымдылық деп аталатын және жаңа деректерге бейімділік қабілеттігіне байланысты басқа проблеманы тудырады. Жарыс әдісімен оқытудың бұл ерекшеліктері *Гроссбергтің тұрақтылық-созымдылық дилеммасы* деген атаумен әйгілі, ол келесі сұрақты қояды: "Жаңа құбылыстарға (иекмділікке) қалай үйрету керек және сонымен бірге бар білім өшпеу немесе бұзылмау үшін тұрақтылықты қалай сақтау керек?"

Карпенер мен Гроссберг әзірлеген ЖНЖ-лердің бейімді резонансты теориясында (ART1, ART2 және ARTMAP) бұл дилемманы шешуге қабілеті бар. Мұндай желіде жеткілікті шығыс элементтер саны бар, бірақ олар бұған қажеттілік туындағанға

дейін қолданылмайды. Элемент үлестірілген (үлестірілмеген) деп айтамыз, егер ол пайдаланылса (пайдаланылмаса). Оқытатын алгоритм қолда бар категория прототипіне түзету енгізеді, егер оған кіру векторы жеткілікті дәрежеде ұқсас болса ғана. Бұл жағдайда олар резонансқа ұшырайды. Ұқсастық дәрежесі k ($0 < k < 1$) ұқсастық параметрімен бақыланады, сондай-ақ ол категория санымен байланысты болады. Кіру векторы ЖНЖ прототипінің ешқайсысына жеткіліксіз ұқсас болса, жаңа категория құрылады және онымен прототиптің бастапқы мәні ретіндегі шығу векторымен үлестірілмеген элемент байланыстырылады. Егер үлестірілмеген элемент табылмаса, онда жаңа вектор ЖНЖ реакциясын шақырмайды.

Моделді көрсету үшін 0, 1 бинарлық кіріске есептелген ART1 желісін қарастырайық. ART1 сәулетінің қысқартылған сұлбасы III.2.6.2-суретте көрсетілген, ол толық байланыстары бар элементтердің екі қабатын қамтиды.



III.2.6.2-сурет. ART1 желісі.

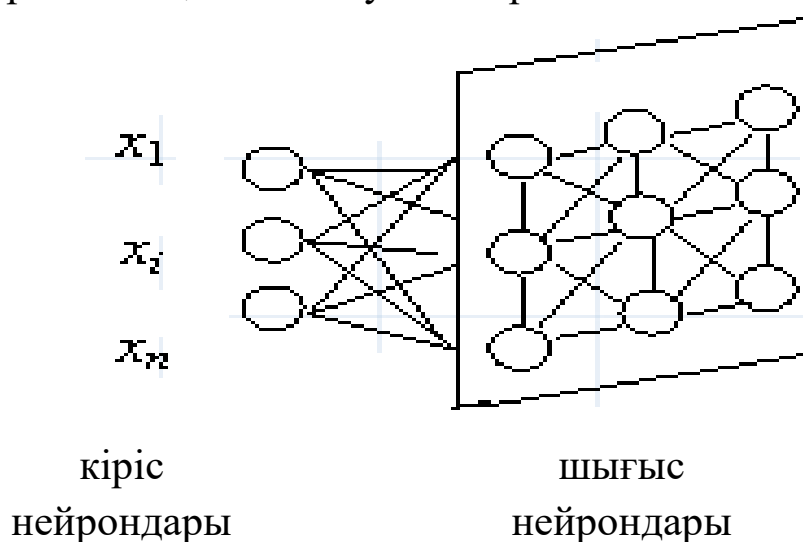
Жоғарыдан төменге бағытталған $\bar{w}_i$ салмақтық вектор кіру қабатының j элементіне сәйкес келеді, ал төменнен жоғарыға бағытталған $\bar{w}_i$ салмақтық вектор i шығу элементімен байланысты; $\bar{w}_i$ векторы $\bar{w}_i$ векторының нормалданған нұсқасы. $\bar{w}_j$ векторлары кластерлердің прототиптерін сақтайды.

Нормаландыру рөлі ұзындығы үлкен векторлардың ұзындығы кіші векторлардан басымды болдырмауда. R түсіру сигналы тек

ұқсастық берілген деңгейден төмен болғанда ғана тудырылады (генерацияланады).

ART1 моделі жаңа категориялар құруы мүмкін және желі өз сыйымдылығын бітірген кезде кіру мысалдарын шығарып тастай алады. Дегенмен желімен табылған категориялар саны ұқсастық параметріне сезімтал болады.

Жарысу әдісімен оқытылатын ЖНЖ-нің арнайы жағдайы өздігінен ұйымдасатын Кохонен картасы болады, ол топологияны сақтаудың сәтті қасиетіне ие және әрбір шығыс элементі үшін кеңістіктік аймақты анықтайды. Сонда жақын кіру мысалдары жақын шығыс элементтерін белсендіру етеді. III.2.6.3-суретте Кохонен картасының негізгі сәулеті көрсетілген.



III.2.6.3-сурет. Өздігінен ұйымдасқан Кохонен картасы.

Мәнісінде, Кохонен картасы екіөлшемді массив элементтерін береді, сондай-ақ әр элемент барлық n кіріс түйіндерімен байланысқан. Локалды аймақ шаршы, шеңбер, тіктөртбұрыш болуы мүмкін. Аймақтың бастапқы өлшемі көбінесе ЖНЖ өлшемінің $1/2$ -ден бастап $2/3$ -ке дейінгі аралықта тағайындалады және нақты заңға сәйкес (мысалы, экспоненциалды кему тәуелділігі бойынша) қысқартылады. Оқыту кезінде барлық жеңімпаз және оның көршілес элементтерімен байланысты барлық салмақтар түрлендіріледі.

III.2.6-тапсырмалар.

1. Жарысу әдісімен оқытудың геометриялық бейнесін құрыңыз.
2. Нейрондық желіні оқытуды жалғастыру шартын көрсетіңіз.
3. Оқыту жүйесінің тұрақтылық мен созымдылық шарттарын атаңыз.
4. Оқыту жүйесінің тұрақтылығына жету тәсілдердің бірін көрсетіңіз.
5. Гроссберг бойынша оқыту жүйесінің тұрақтылық-созымдылық дилеммасын айтыңыз.
6. Гроссберг бойынша оқыту жүйесінің тұрақтылық-созымдылық дилеммасын шешу қабілетіне ие теорияны атаңыз.
7. Кохоненнің өздігінен ұйымдастқан картасы түрінде жарысу әдісімен оқытылатын ЖНЖ-ң арнайы жағдайын сипаттаңыз.

III.2.6-сұрақтар.

1. Нейрондық желіні жарысу әдісімен оқытудың Хебб әдісінен айырмашылығы қандай?
2. Жарысу әдісінде қандай принцип әрекет етеді?
3. Жарысу әдісімен нейрондық желіні оқытуға не мүмкіндік беретін?
4. Жарысу әдісімен нейрондық желіні оқыту кезінде не жүзеге асырылады?
5. Жарысу әдісімен оқытуда "жеңімпаз" ережесіне ненің әсерінен қол жеткізіледі?
6. Жарысу әдісімен оқытудың геометриялық бейнесінде кіріс векторлары немен ұсынылған?
7. Қандай да бір кірістік үлгі оқыту үдерісіндегі келесі итерацияда басқа шығу нейрондарын белсендіру ете алады ма?

III.2.6-тестілер.

1. Жарысу әдісімен оқытудың Хебб әдісінен айырмашылығы неде?
 - А) шығу нейрондары өзара белсендіру болу үшін жарысады;
 - В) шығу нейрондары өзара тоқтау үшін жарысады;

- С) кіру нейрондары өзара белсендіру болу үшін жарысады;
- Д) шығу нейрондары өзара тоқтау үшін жарысады;
- Е) шығу нейрондары өзара шығу үшін жарысады.

2. Жарысу әдісімен оқыту кезінде қандай нейронның салмақтары ғана түрленеді?

- А) тек қана "жеңген";
- В) тек қана "ұтылған";
- С) тек қана "өшкен";
- Д) тек қана "шыққан";
- Е) тек қана "жойылған".

3. ЖНЖ-нің жылдамдық параметрі қандай болғанда, оқыту ешқашан тоқтамайды?

- А) 0-ге тең емес болса;
- В) 1-ге тең емес болса;
- С) 2-ге тең емес болса;
- Д) 0-ге тең болса;
- Е) кез келгенде мәнде.

4. "Жаңа құбылыстарға (иекмділікке) қалай үйрету керек және сол уақытта бар білім өшіп немесе бұзылып қалмау үшін тұрақтылықты қалай сақтау керек?" деген сұрақты қоятын тұрақтылық-иекмділік дилемасы кімдікі?

- А) Гроссберг;
- В) Хопфилд;
- С) Кохонен;
- Д) Хебб;
- Е) Гаусс.

5. Кохонен картасы қандай элементтерді береді?

- А) екіөлшемді массив;
- В) бірөлшемді массив;
- С) үш өлшемді массив;
- Д) төрт өлшемді массив;
- Е) бес өлшемді массив.

III.2.7. Тікелей әдіс пен кері әдісті біріктіру

Бұл тақырыпта тікелей әдіс пен кері әдісті біріктіру қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

ЖНЖ-ні оқыту кезінде қателерді кері тарату әдісі Коши үлестірімі пайдаланатын тікелей әдіске қарағанда артықшылыққа ие, себебі кейінгісінде кездейсоқ таңдаудың кесірінен көптеген қадамдар дұрыс емес бағытта орындалып, көп уақытты алады.

Оқытуды жеделдету үшін (қателік функциясының глобалды минимумын іздеу) оқытудың тікелей және кері қосу әдістерін біріктіруге болады.

Осы әдістерді біріктіру жақсы нәтиже беруі мүмкін. Кері тарату алгоритмімен есептелінетін салмақталған сомаға тең салмақтарды коррекциялау және Коши алгоритмімен берілетін кездейсоқ қадам әрбір әдіс жекелеп оқытатын жүйеге қарағанда глобалды минимумды тез табатын және жинақталатын бір жүйеге әкеледі. Қатені кері тарату әдісін қолданғанда да және Коши әдісін қолданғанда да орын алуы мүмкін ЖНЖ-нің дәрменсіздігінен айрылу үшін қарапайым эвристика пайдаланылады.

Екі әдісте туындайтын қателерді қарастырайық:

1. Қателіктің кері тарату әдісінің қиындықтары.

Кері тарату әдісінің қуатына қарамастан, оны қолдану кезінде бірнеше қиындықтар туындайды. Олардың бір бөлігі жаңа алгоритмді қолдану арқасында жеңілдетіледі.

Жинақтылық.

Жинақтылықты дәлелдеу дербес туындылардағы дифференциалды теңдеулер тілінде мүмкін болады, ол салмақ коррекциясы салмақтар кеңістігіндегі шексіз аз қадамдардың көмегімен орындалғанда ғана әділ болады. Алайда, шексіз аз қадамдар шексіз жинақтылық уақытына алып келеді. Әдіс практикалық қолданыста өз күшін жоғалтады, себебі, кері тарату қадамның ақырлы санында жинақталатындығы дәлелденбейді.

Эксперименттер көрсеткендей, әдетте ЖНЖ оқытылады, бірақ оқыту уақыты ұзақ болады және болжанбайды.

Дәрменсіздік.

Кейбір жағдайларда ЖНЖ оқыту кезінде салмақтарды жетілдіру ЖНЖ-нің нақты өзгеруіне әкелмейтін күйге түседі. Мұндай «дәрменсіздік» күрделі мәселе болады. Ол бір рет туындап, оқыту уақытын бірнеше ретке ұлғайтуы мүмкін.

ЖНЖ-ні оқыту үдерісіндегі дәрменсіздік салмақталған сомаға үлкен мәндерді беру үшін нейрондардың елеулі бөлігінде салмақтар мәндері өте үлкен болған жағдайда туындайды. Бұл нейрондардың барлығы немесе көбі сығушы белсендіру функциясы арументтерінің мәндері өте үлкен болатын және функцияның туындысы нөлге жақындайтын облыста (мысалы, сигмоидалды белсендіру функциясы үшін) жұмыс істеуіне әкелуі мүмкін.

Егер кері тарату алгоритмінде салмақтар мәнін есептеу кезінде формуладағы осы туындыны коэффициент (салмақтардың коррекция мәні туындының мәніне пропорционал) ретінде қолданса, онда оқыту үдерісі іс жүзінде тоқтауы мүмкін және ЖНЖ дәрменсіздігі туындауы мүмкін екендігі белгілі. Дәрменсіздікпен зақымдалған нейрондар үшін туындының нөлге жақындауы салмақтың түрленуін нөлге жақындатады.

Оқыту кезінде желі дәрменсіздене ме, жоқ па, оны алдын ала болжайтын теория жоқ. Қадамның кіші өлшемі дәрменсіздікке сирек әкелетіндігі эксперименталды түрде дәлелденген, алайда бір тапсырмаға кіші қадам басқасына үлкен болуы мүмкін. Дәрменсіздік құны жоғары болуы мүмкін.

Моделдеу кезінде машиналық уақыттың көптеген сағаттары дәрменсіздіктен шығу үшін кетуі мүмкін. Әдетте бұдан η қадам өлшемін кеміту арқасында қашуға болады, бірақ бұл оқыту уақытын арттырады және алгоритмнің жинақталуын баяулатады.

Дәрменсіздіктен алдын-ала қорғану үшін немесе одан кейінгі қайта қалыпқа келу үшін әртүрлі эвристикалар қолданылды, бірақ олар қазірше эксперименттік түрде ғана қарастырылуы мүмкін. Мұндай жағдайды жеңу үшін антиградиенттің тек бағытын

ескеретін және оның шамасын ескермейтін алгоритмнің жетілдірілуі бар. Сондай-ақ, оқыту үрдісі кезінде қадам өлшемін автоматты түрде коррекциялайтын қадамды таңдаудың бейімді алгоритмі бар.

Локалды минимум.

Қатені кері тарату әдісінде ЖНЖ синапстық байланыс салмақтарының мәнін коррекциялау үшін локалды минимумға бағытталған синапстық байланыс салмақтарының кеңістігінде қателердің беті бойынша градиентті түсу қолданылады. Бірақ локалды минимум бір жағдайларда тиімді шешім болып табылғанмен, басқа жағдайда тиімсіз болуы мүмкін.

Әдетте іс жүзінде күрделі ЖНЖ-нің қателік беті қатты тілімделген және жоғары өлшемді кеңістікте жазықтықтан, қатпарлардан, шыңдар мен құздардан тұрады. ЖНЖ локалды минимумға түсуі мүмкін және глобалды минимумды таппауы мүмкін. Локалды минимум нүктесінде барлық бағыттар жоғары қарай апарады, одан ЖНЖ шыға алмауы мүмкін.

ЖНЖ-ні оқытудың осы қиындығынан өту үшін жасырын қабаттарды және жасырын қабаттағы нейрондар санын көбейту есебінен салмақтар кеңістігінің сыйымдылығын кеңейтеді немесе бірнеше рет оқытады да, кейін жақсы нәтиже көрсеткен ЖНЖ-ні таңдайды. Сондай-ақ локалды минимумнан η оқыту жылдамдығын арттыру арқылы қысқа уақытқа шығуға болады. Кейде синапстық байланыс салмақтарының мәндерін өзгертуге шуды қосады. Бұл ЖНЖ-нің локалды минимумнан «секіруге» мүмкіндік береді.

Тіпті желі оқытылғаннан кейін де кері тарату көмегімен глобалды минимумның табылғандығы туралы айту мүмкін емес. Егер шешім қанағаттандырылықсыз болса, онда салмақтарға жаңадан кездейсоқ мән беруге тура келеді және желіні осы әрекет кезінде оқыту аяқталады немесе глобалды минимум қашан да бір табылады деген кепілдіксіз қайта оқуға тура келеді.

Артықоқыту

Оқытатын таңдамада жоғарғы дәлдік тесттік таңдамада нәтижелердің орнықсыздығына әкелуі мүмкін. ЖНЖ нақты

шарттарға (оқытатын таңдамаға) қаншалықты бейімделген болса, соншалықты аз жалпылауға және экстраполяциялауға қабілетті болады. Бұл жағдайда ЖНЖ функцияны емес, оқытатын таңдамада болатын шуылды моделдейді. Осы құбылыс *артықоқыту* деп аталады. Осы кемшілікпен күрестің бірден-бір құралы – ЖНЖ-ні артықоқытуды анықтау үшін пайдаланатын растайтын мысалдардың таңдамаларын қолдану. Растайтын таңдамамен жұмыс кезінде ЖНЖ мінездемесінің төмендеуі мүмкін болатын артықоқытуға меңзейді. Керісінше, егер растайтын деректер жиынынан мысалдарды бергенде қателік біртіндеп азайса, онда ЖНЖ әрі қарай оқытыла береді.

Осы әдістің кемшілігі оқытатын жиынында қолдануға болатын мысалдар санын азайту болып табылады. Себебі, оқытатын таңдама санының азаюы ЖНЖ жұмысының сапасын төмендетеді. Сонымен қатар, бастапқы деректерді оқытатын, тестілейтін және растайтын таңдамаға оңтайлы бөлу мәселесі туындайды. Тіпті кездейсоқ таңдамада деректер базасын әртүрге бөлу әртүрлі бағалар береді.

2. Коши алгоритмімен оқыту қиындықтары.

Больцман машинасымен салыстырғанда Коши машинасы беретін оқыту жылдамдығының жақсарғандығына қарамастан, жинақталу уақыты кері тарату алгоритімінің уақытынан 100 рет артық болуы мүмкін. Айта кетсек, желілік дәрменсіздік Кошидің оқыту алгоритмі үшін аса қауіпті, әсіресе ЖНЖ сызықты емес логистикалық функция үшін. Коши үлестірімінің ақырсыз дисперциясы салмақтардың шектелмген шамаға өзгеруіне алып келеді. Әрі қарай, салмақтың үлкен өзгерістері дәрменсіздіктің қауіпі пайда болатын желілік нейрондардың қатты қанығуына жиі әкелетін жағымсыз жағдайлар болғанда да кейде қабылданады.

III.2.7-тапсырмалар.

1. Қателерді кері тарату әдісіне қарағанда тікелей әдістегі кемшіліктердің болу себебін көрсетіңіз.

2. Нейрондық желінің тура әдісі мен кері оқыту әдісінің біріктіруінен туындайтын барлық оң нәтижелерді атаңыз.

3. Нейрондық желідегі кері қатені тарату әдісін оқытудағы қиындықтарды көрсетіңіз.

4. Нейрондық желідегі тура Коши әдісінің оқыту кезіндегі қиындықтарын көрсетіңіз

5. Желіні артықоқытудың туындау себебін түсіндіріңіз

III. 2.7-сұрақтар.

1. Нейрондық желідегі тура әдіс пен кері оқыту әдісінің бірігуінің туындау себебі неде?

2. Нейрондық желідегі тура әдіс пен кері оқыту әдісінің бірігуі арқылы қандай нәтижеге қол жеткізуге болады?

3. Нейрондық желіде оқытудың сәйкестігі дегеніміз не?

4. Нейрондық желідегі оқытудағы дәрменсіздігі деп нені айтамыз?

5. Желідегі дәрменсіздікті алдын ала анықтау теориясы бар ма?

6. Нейрондық желідегі кері қатені тарату әдісін оқытуда локалды минимумын анықтауда қандай әдіс қолданылады?

7. Желіге кері тарату әдісін оқытқаннан кейін глобалды минимум анықтала ма?

III.2.7-тестілер.

1. Оқытудың тікелей және кері әдістерін біріктіру не үшін қолданылады?

- А) оқытуды жеделдету үшін;
- В) оқытуды баялату үшін;
- С) глобалды максимумды табу үшін;
- Д) локалды максимумды табу үшін;
- Е) қатені іздеу үшін.

2. ЖНЖ оқыту кезінде салмақтардың түрленуі ЖНЖ-нің нақты өзгеруіне әкелмейтін күйді қалай атайды?

- А) дәрменсіздік;

- В) жинақтылық;
- С) локалды минимум;
- Д) локалды максимум;
- Е) қатені іздеу.

3. ЖНЖ оқытатын таңдамаларға қаншалықты бейімделген болса, соншалықты аз тұжырымдауға және экстраполяциялауға мүмкін болатын, сонымен қатар функцияны емес, оқытылатын таңдамада бар шуды моделдейтін құбылысты қалай атайды?

- А) артықоқыту;
- В) дәрменсіздік;
- С) локалды минимум;
- Д) локалды максимум;
- Е) қатені табу

4. Оқыту таңдама санының төмендеуі неге алып келеді?

- А) ЖНЖ жұмысының сапасын азайтады;
- В) ЖНЖ жұмысының санын азайтады;
- С) ЖНЖ жұмысының жылдамдығын азайтады;
- Д) локалды максимумды табады;
- Е) локалды минимумды табады.

5. Коши үлестірімінің шексіз дисперсиясы шектелмеген шамадағы нелердің өзгеруіне алып келеді?

- А) салмақтардың;
- В) жұмыс санының;
- С) жұмыс жылдамдығының;
- Д) кірістердің;
- Е) шығыстардың.

III.2.8. Қателерді кері тарату әдісі мен Коши әдісінің комбинациясы

Бұл тарауда қателерді кері тарату әдісі мен Коши әдісінің комбинациясы қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Қателерді кері тарату әдісі мен Коши әдісін қолданатын нейрондық желілерді оқытудың аралас алгоритмінде синапстық байланыстардың салмақтарын түзету екі компоненттен тұрады: (1) қателерді кері тарату алгоритмін пайдаланып есептелетін бағытталған компоненттер, және (2) Коши үлестірімімен анықталатын кездейсоқ компоненттер.

Бұл компоненттер әрбір салмақ үшін есептеледі және олардың қосындысы салмақтың өзгеретін шамасы болады. Коши алгоритіміндегідей, салмақтың өзгеруі есептелгеннен кейін мақсаттық функция есептеледі. Егер жақсару орын алса, онда өзгеріс сақталады. Кері жағдайда ол Больцман үлестірілімімен анықталатын ықтималдықпен сақталады.

Салмақтың түзетілуі жоғарыда көрсетілген теңдеулерді қолдану арқылы алгоритмдердің әрқайсысы үшін есептеледі:

$$w_{ij}^{(q)}(t+1) = w_{ij}^{(q)}(t) + \eta[\alpha \Delta w_{ij}^{(q)}(t) + (1-\alpha)\delta_j^{(q)} y_{ij}] + (1-\eta)x_c$$

мұнда $w_{ij}^{(q)}(t+1)$ – итерацияның $(t+1)$ -ші қадамындағы $(q-1)$ -ші қабаттың, I -ші нейронымен q -ші қабаттың j -ші нейронының синапс салмағы, $\Delta w_{ij}^{(q)}$ – синапстық байланыстың салмағының түзетілуі, η – Кошидің салыстырмалы шамаларын және салмақ қадамының компоненттеріндегі кері таратуды басқаратын оқыту жылдамдығының коэффициенті, $q=Q-1, Q-2, \dots, 2; 0<\eta<1$; $y_{j,k}^{(Q)}$ – k -ші оқытатын мысалдағы (үлгіді) j -ші нейрон үшін ЖНЖ-нің Q -ші шығу қабатының нақты шығысы.

Егер η нөлге теңелсе, онда жүйе толық Коши машинасына айналады. Ал егер η бірге теңелсе, онда жүйе кері тарату машинасына айналады.

Салмақтық функцияларды есептеулер арасында тек бір салмақты өзгерту тиімді болмайды. Бүтін қабаттың барлық салмақтарын өзгерту тиімдірек болып шықты, бірақ кейбір есептер үшін басқа стратегияны қолданған дұрыс.

Оқытудың аралас әдісімен желілік дәрменсіздіктен арылу.

Коши машинасы сияқты, егер салмақтың өзгеруі мақсаттық функцияны нашарлататын болса, онда Больцман үлестірімі көмегімен салмақтың жаңа мәнін сақтап қалу керек пе әлде алдыңғы мәнді қалпына келтіру керек пе екендігі шешіледі. Ендеше, салмақтар өсімдерінің жиындары нашарлауын сақтайды деген ақырғы ықтималдық бар болады. Коши үлестірімінің шексіз дисперсиясы (тангенстің өзгеру диапазоны – $-\infty$ -тен $+\infty$ -ке дейін) болғандықтан, көбіне желілік дәрменсіздікке алып келетін үлкен салмақтар өсімдерінің пайда болу ықтималдығы бар.

Салмақтық қадамның өзгеру диапазонының шектелуі туралы белгілі шешім осылай алынған алгоритмнің математикалық дұрыстығына сұрақ қойғызады. Шындығында, кейбір функцияны жүзеге асыру үшін үлкен салмақ керек болады, сонда екі үлкен салмақ шегеріліп аз айырым беретін жағдай эксперименттік тұрғыдан анықталған.

Басқа шешім қаныққан күйде болған нейрондардың салмақтарын рандомдаудан тұрады. Бұл шешімнің кемшілігі оқыту үдерісін бұзуы мүмкін, кейде оны шексіздікке дейін созады.

Дәрменсіздік проблемасын шешу үшін жаңа әдіс табылды, ол осыған дейін жеткен оқытуды бұзбайды. Қаныққан нейрондар олардың шығу сигналдарын өлшеу көмегімен анықталады. Бұл сигналдардың шамасы өздерінің шеткі оң немесе теріс мәндеріне жеткен кезде, осы нейронды қоректендіретін синапстық байланыстың салмақтарына сығымдаушы функция әсер етеді. Ол нейрондық шығу сигналын алу үшін қолданылатын әдіске ұқсас болады, тек оның өзгеріс диапазоны (+5,-5) интервал немесе басқа сәйкес жиын болатын жағдайдан басқа кезде. Сонда салмақтық коэффициенттердің түрленген мәні мынаған тең:

$$w_{mn} = \frac{-5+10}{1+\exp(-w_{mn}/5)}$$

Бұл функция өте үлкен салмақтың шамасын азайтады. Ал аз салмаққа азырақ әсерін тигізеді. Кейіннен, ол үлкен салмақтар арасындағы шамалы айырмашылықты сақтай отырып, симметрияны да сақтайды. Эксперименттік тұрғыда бұл функция нейрондарды қанығу күйінен ЖНЖ оқытуында қол жеткізілгенді бұзбай шығаратыны дәлелденген. Пайдаланатын функцияны оңталайландыру барысында көп күш жұмсалған жоқ, сонымен қатар, тұрақтының басқа мәндері аса тиімді болуы мүмкін.

Тәжірибелік нәтижелер.

Кері тарату әдісі мен Коши оқытуын пайдаланатын аралас алгоритм бірнеше үлкен желілерді оқыту үшін қолданылған. Мысалы, бұл әдіс арқылы қытай қолжазба иероглифын анықтайтын жүйе оқытылды. Дегенмен, оқыту ұзаққа созылуы мүмкін (шамамен, оқытуға машинаның 36 сағаты кеткен).

Басқа тәжірибеде бұл желі *эксклюзивті немесе* есебінде оқытылды да басқа алгоритмдермен салыстыру үшін тест ретінде қолданылды. ЖНЖ жинақтылығы үшін 76 оқыту жиынын ұсыну қажет болды. Салыстыру үшін, кері тарату әдісін қолдану барысында, орташа есеппен 245 ұсыныс және екінші ретті кері таратуды қолданғанда 4 986 итерация қажет болды.

Оқытудың ешқайсысы локалды минимумға алып келмеді. Сонымен қатар, 160 оқытудың ешқайсысында күтпеген потология анықталмады және желі әрқашан дұрыс оқытылды.

Таза Коши машинасымен жүргізген тәжірибелер нәтижесінде оқыту уақытының үлкен өзгерісіне алып келді. Мысалы, $\rho = 0,002$ тең болғанда, ЖНЖ орта есеппен 2 284 ұсыныс қажет болды.

Талқылау.

Кері тарату әдісі мен Коши оқытуын қолданатын аралас желі әрбір жеке алгоритмдерге қарағанда тезірек оқытылады және салыстырмалы түрде коэффициенттер мәндеріне қатысты сезімтал

емес. Коши алгоритмінде глобалды минимумға жетуге кепілдік беріледі, жүздеген тәжірибелер барысында оқыту желісі ешқашан локалды минимумның тұзағына түсіп көрмеді. Желілік дәрменсіздік мәселесі салмақты селективті қысу алгоритмі арқылы шешілді. Ол барлық берілген тесттік тапсырмаларда оқыту уақытын айтарлықтай ұлғайтпай жинақтылықты қамтамасыз етті.

Бірақ мұндай үміттендіретін нәтижелерге қарамастан, бұл әдіс әсіресе үлкен тапсырмаларда соңына дейін зерттелген жоқ. Айтарлықтай үлкен жұмыс оның артықшылықтары мен кемшіліктерін анықтауға жұмсалады.

III.2.8-тапсырмалар.

1. Нейрондық желіні оқытудың аралас алгоритміндегі әр компоненттің жұмысын сипаттап беріңіз.
2. Нейрондық желіні оқытудың аралас әдісі үшін синапстық байланыстар салмағын түзетуді есептеу формуласын жазыңыз.
3. Салмақтық коэффициенттердің түрленген мәндер формуласын жазыңыз.

III.2.8-сұрақтар.

1. Нейрондық желіні оқытудың аралас әдісінде салмақтарды түзету қалай жүзеге асырылады?
2. Нейрондық желіні оқытудың аралас әдісінде қандай жағдайда Коши машинасы және қандай жағдайда кері тарату машинасы қолданылады?
3. Нейрондық желіні оқытудың аралас әдісінде желілік дәрменсіздіктің алдын қалай алады?

III.2.8-тестілер.

1. Коши алгоритмінде, салмақтың өзгеруі есептелген соң, одан басқа не есептеледі?
 - А) мақсаттық функция;
 - В) межелік функция;
 - С) қадамдық функция;
 - Д) нейрон саны;
 - Е) салмақ саны.
2. Егер η нөлге теңелсе, онда жүйе қандай машинаға айналады?

- A) толық Коши;
- B) кері тарату;
- C) Тьюринг;
- D) Хебб;
- E) Кохонен.

3. Егер η бірге теңелсе, онда жүйе қандай машинаға айналады?

- A) кері тарату;
- B) толық Коши;
- C) Тьюринг;
- D) Хебб;
- E) Кохонен.

4. Коши үлестірімінің шексіз дисперсиясында тангенстің өзгеру диапазоны қандай?

- A) $-\infty$ -тен $+\infty$ -ке дейін;
- B) 1-ден $+\infty$ -ке дейін;
- C) $-\infty$ -тен +1-ге дейін;
- D) $-\infty$ -тен 0-ге дейін;
- E) 0-ден $+\infty$ -ке дейін.

5. Салмақтық коэффициенттердің түрленген мәні неге тең?

- A) $w_{mn} = \frac{-5+10}{1+\exp(-w_{mn}/5)}$;
- B) $w_{mn} = \frac{10}{1+\exp(-w_{mn}/5)}$;
- C) $w_{mn} = \frac{-5+10}{\exp(-w_{mn}/5)}$;
- D) $w_{mn} = \frac{-5+10}{10+\exp(-w_{mn}/10)}$;
- E) $w_{mn} = \frac{10}{10+\exp(-w_{mn}/25)}$;

III.2.9. ЖНЖ оқыту үдерісіндегі нейрондар санын баптау

Бұл тақырыпта ЖНЖ оқыту үдерісіндегі нейрондар санын баптау сұрақтары қарастырылады, тапсырмалар ұсынылады және сұрақтар беріледі. Оқу материалдарын дайындау барысында [5–38] деректер көздері пайдаланылды. Көп қабатты ЖНЖ оқытатын таңдамасымен берілген бейнені жүзеге асыру үшін, оның жасырын қабаттарында жеткілікті нейрондар саны болу керек.

Қазіргі уақытта берілген оқытатын таңдама бойынша ЖНЖ-не қажетті нейрон сандарын нақты анықтауға арналған формула жоқ. Бірақ, есепті шешу үшін ЖНЖ құруды қамтамасыз ететін және артықшылықты болдырмайтын мүмкіндікті беретін оқыту үдерісінде нейрондар санын баптау тәсілдері ұсынылған. Баптаудың бұл тәсілдерін екі топқа бөлуге болады: *қысқарту алгоритмдері (pruning algorithms)*, *конструктивтік алгоритмдер (constructive algorithms)*.

1. Қысқарту алгоритмдері.

Қысқартумен оқыту алгоритмінің жұмысы басында ЖНЖ-нің жасырын қабаттарында нейрондар саны алдын ала артық болады. Кейін, ЖНЖ-нен біртіндеп синапстар мен нейрондар жойыла бастайды. Қысқарту алгоритмдерін іске асырудың екі тәсілі бар:

1) Айыппұл функциялар әдісі.

Синапстық байланыс салмағының мәндері нөлден өзгеше болғандықтан, оқыту алгоритмінің мақсатты функциясына айыппұл енгізіледі, мысалы:

$$C = \sum_{i=1}^I \sum_{j=1}^J w_{ij}^2$$

мұнда w_{ij} – синапс салмағы, I –нейрон номері, j – кіріс номері; I – кіру сигналының өлшемі; J – жасырын қабаттағы нейрон саны.

2) Проекция әдісі.

Синапстың салмағы нөлденетін болады, егер оның мағынасы берілген диапазонға (аралыққа) түссе:

$$w_{ij} = \begin{cases} 0, & |w_{ij}| \leq \varepsilon \\ w_{ij}, & |w_{ij}| > \varepsilon \end{cases}$$

мұнда ε – қайсыбір тұрақты.

Қысқарту алгоритмдерінің кем дегенде, екі кемшілігі болады:

- артық болып табылатын жасырын қабаттардың нейрондар санын анықтайтын әдістеме жоқ, сол себепті алгоритм жұмысының алдында сол санды тауып алу қажет;
- алгоритмнің жұмыс істеу үдерісінде желі саны артық нейрондардан болады, сондықтан оқыту ақырын жүреді.

2. Конструктивтік алгоритмдер.

Конструктивтік алгоритмнің алдындағысы ретінде келесі қадамдардан тұратын коп қабатты нейрондық желілерді оқыту әдістемесін алуға болады:

1-қадам. Жасырын қабаттардағы нейрондардың бастапқы санын таңдау.

2-қадам. Синапс салмағына және ЖНЖ-нің ығысуына берілген аралықтан кездейсоқ мәндерді меншіктеуден тұратын, ЖНЖ инициализациялау.

3-қадам. Берілген таңдама бойынша ЖНЖ-ны оқыту.

4-қадам. Сәтті оқытылған жағдайда аяқтау. Егер ЖНЖ оқыту іске асырылмаса, онда нейрондардың саны арттырылады және 2-қадамнан 4-қадамға дейінгілер қайталанады.

Конструктивтік алгоритмдерде жасырын қабаттардағы нейрондар саны басында аз болады, кейін біртіндеп көбейеді. Бұл әдістен конструктивтік алгоритмдердің айырмашылығы, онда нейрондар саны артқанға дейін желінің алған дағдылары сақталады.

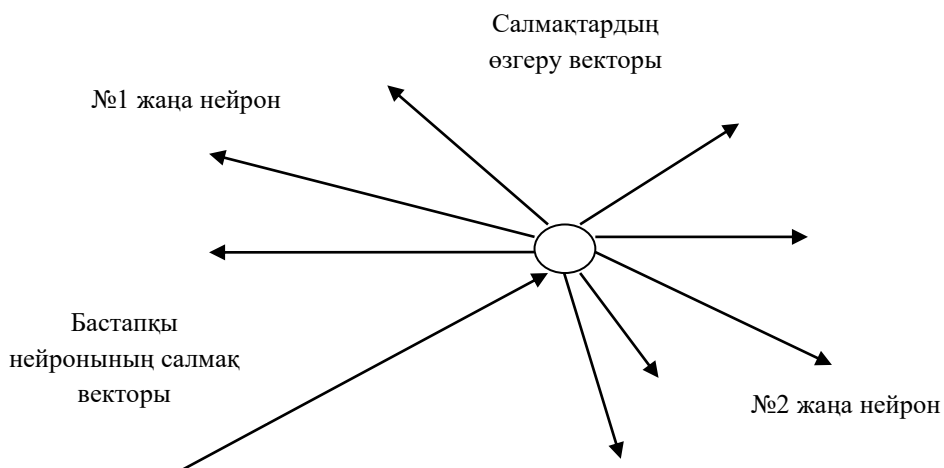
Конструктивтік алгоритмдерді желіге жаңа қосылған нейрондардың параметрлер мәндерін беру ережесімен ажыратады:

- параметр мәндері берілген аралықтағы кездейсоқ сандар болады;

– жаңа нейрондардың синапстық байланыс салмақтарының мәні ескі нейрондардың бірін *ыдырату* (*splitting*) жолымен анықталады.

Бірінші ереже маңызды есептеулерді қажет етпейді, бірақ оны қолдану жаңа нейрондарды әр қосқаннан кейін қате функциясы мәндерінің артуына алып келеді. Жаңа нейрондардың параметрлерінің мәндерін кездейсоқ беру нәтижесінде жасырын қабаттағы нейрондар санының артылып кетуі мүмкін. Нейрондарды ыдырату көрсетілген екі кемшіліктен айрылған. Алгоритмнің мәнісі осындай нейрондарды табудан және ыдыратудан тұрады. Ыдырау нәтижесінде ЖНЖ-да бір бастапқы нейронның орнына екі нейрон пайда болады. Бұл нейрондардың біріншісі бастапқы нейронның салмақтар векторының және басымды бағыттардың бірінің салмақтарының өзгеру векторының қосындысын көрсететін салмақтар векторынан тұрады. Басқа басымды бағыттардың салмақтарының өзгеру векторларын және бастапқы нейрон салмағының векторын қосу нәтижесінде екінші жаңа нейронның салмағы алынады.

III.2.9-суретте оқытудың қандай да бір қадамында жасырын қабаттың нейрон салмақтарының векторы және жеке оқытатын мысалдарға сәйкес келетін салмақтың өзгеру векторы көрсетілген.



III.2.9-сурет. Жеке оқылатын мысалдарына сәйкес келетін, жасырын

қабаттың нейрондар салмақтарының және өзгеруінің векторы және салмақтың өзгеру векторы.

Өзгеру векторының екі басымды бағыттары бар және кеңістікте сфералықтан маңызды айырмашылығы бар аймақты қалыптастырады.

Екі артық бағыты бар өзгеру векторларында нейрондарды ыдырату қажетті. Себебі, осындай нейрондардың болуы кері тарату әдісін оқыту кезінде осцилляцияға әкеліп соғады. Қателердің интегралды функциясы бар әдіспен оқыту кезінде мұндай нейрондардың болуы қателердің үлкен мәнінде локалды минимумға түсуіне әкеледі.

Ыдырату алгоритмі келесілерден тұрады:

- синапстық байланыс салмақтар өзгерістерінің векторларының ковариациялық матрицасын құру;
- алынған матрицаның меншікті векторларын және меншікті мәндерін Ойа (Oja) итерациялық алгоритмі көмегімен есептеу, осыларға сәйкес стохастикалық градиенттік көтеру мен Грамм-Шмидтің ортогоналдау орындалады;

ЖНЖ өлшемінің үлкеюі кезіндегі есептеу уақытының экспоненциалдық өсуі алгоритмнің кемшілігі болады. Көрсетілген кемшілікті болдырмау үшін бірқатар есептеулерді қажет етпейтін ыдырау алгоритмінің жеңілдетілген түрі ұсынылған. Берілген оқытатын таңдама бойынша осы алгоритммен құралған ЖНЖ-дегі нейрондардың жалпы саны, бастапқы алгоритм көмегімен құрылған ЖНЖ- дегіден бірнеше санға көбірек. Жеңілдетілген алгоритмде ыдырату үшін функционалы үлкен мәнді нейрон таңдалады:

$$F_j = \frac{\sum_{k=1}^N |\delta w_j^k|}{\left| \sum_{k=1}^N \delta w_j^k \right|},$$

мұнда δw_j^k – нейронның синапстық байланыс салмақтарының өзгеру векторы; k – оқытатын мысалының номері; j – нейрон номері; $j = 1, \dots, M$; M – нейрон саны; N – таңдау оқытуындағы мысал номері.

Осылайша, нейронды ыдырату үшін таңдау критеріі ретінде әртүрлі оқытатын мысалдарға сәйкес келетін, нейронның синапстық байланыстар салмақтарының өзгеру векторларының ұзындығының қосындысының, осы векторлар қосындысының ұзындығына қатынасы қолданылады. Ыдырату нәтижесінде бастапқы нейронның орнына желіге екі жаңа нейрон енгізіледі. Жаңа нейронның әрбір синапстық салмақтарының мәні –ескі нейронның сәйкес салмағы мен оған қосылған кейбір кішігірім шудың мәніне тең. Жаңа нейрон мен келесі қабат нейрондарының шығыс байланыстар салмақтарының шамасы келесі қабаттың сәйкес салмақтары бар қалыпты нейронның байланыстар салмағының шамасының жартысына тең. Қарапайым алгоритм, бастапқысы сияқты ыдыраудан кейін қателік функциясының өспейтініне кепілдік береді.

Ыдырату үшін нейрондарды таңдаудың сипатталған тәсілдерден басқа, сезімталдық талдау қолданылуы мүмкін. Бұл талдау барысында ЖНЖ параметрлері бойынша қателік функциясының екінші туындысына арналған Гессе матрицасы құрылады. Екінші туынды модулінің шамасы бойынша есепті шешу үшін берілген параметр мәнінің маңыздылығын бағалайды. Екінші туынды параметрінің кіші мәні нөлденеді. Сезімталдық талдау үлкен есептеу ресурстарын талап етеді.

III.2.9-тапсырмалар.

1. Айыппұлды функциялар әдісін сипаттаңыз.
2. Айыппұлды функция үшін формуланы жазыңыз.
3. Проекциялау әдісін сипаттаңыз.
4. Конструктивтік алгоритмдерді сипаттаңыз.
5. Конструктивтік алгоритмдердің айырмашылығын көрсетіңіз.

III.2.9-сұрақтар.

1. Қысқарту алгоритмін жүзеге асыруда қандай амалдар бар?
2. Ыдыратуға арналған функционал формуласы қандай?
3. Ыдырату алгоритмі нені қамтиды?
4. Ыдырату алгоритмі қандай кемшілікке ие?
5. Ыдырату үшін нейронды таңдаудың критеріі қандай?

III.2.9-тестілер.

1. Қысқарту алгоритмдерін іске асырудың қандай екі тәсілі бар?
 - A) айыппұл функциялар әдісі және проекция әдісі;
 - B) межелік функция әдісі және проекция әдісі;
 - C) қадамдық функция әдісі және проекция әдісі;
 - D) мақсаттық функция әдісі және проекция әдісі;
 - E) салмақтық функция әдісі және проекция әдісі.
2. Жаңа нейрондардың синапстық байланыс салмақтарының мәндері қалай анықталады?
 - A) ыдырату (splitting);
 - B) жинақтау;
 - C) санау;
 - D) проекциялау;
 - E) жою.
3. Ыдырату алгоритмінің кемшілігі не?
 - A) ЖНЖ өлшемінің үлкеюі кезіндегі есептеу уақытының экспоненциалдық өсуі;
 - B) ЖНЖ өлшемінің үлкеюі кезіндегі есептеу уақытының экспоненциалдық төмендеуі;
 - C) ЖНЖ өлшемінің үлкеюі кезіндегі жылдамдықтың экспоненциалдық өсуі;
 - D) ЖНЖ өлшемінің үлкеюі кезіндегі салмақтың экспоненциалдық кемуі;
 - E) ЖНЖ өлшемінің үлкеюі кезіндегі нейрондар санының өсуі.

4. $F_j = \frac{\sum_{k=1}^N |\delta w_j^k|}{\left| \sum_{k=1}^N \delta w_j^k \right|}$ формуласындағы δw_j^k нені көрсетеді?

А) нейронның синапстық байланыс салмақтарының өзгеру векторы;

В) нейрон номері;

С) нейрон саны;

Д) оқытатын мысалының номері;

Е) таңдау оқытуындағы мысал номері.

5. Ыдырату нәтижесінде не алынады?

А) бастапқы нейронның орнына желіге екі жаңа нейрон енгізіледі;

В) соңғы нейронның орнына желіге екі жаңа нейрон енгізіледі;

С) бастапқы нейронның орнына желіге үш жаңа нейрон енгізіледі;

Д) бастапқы нейронның орнына желіге екі ескі нейрон енгізіледі;

Е) бастапқы нейронның орнына желіге бес жаңа нейрон енгізіледі.

III.2.10. Нейрондық желі жұмысының сапасын бағалау

Бұл тақырыпта жасанды нейрондық желілер жұмысының сапасын бағалау қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Алдыңғы тарауларда ЖНЖ оқыту, оқытатын мысалдардың берілген жиынында ЖНЖ шығаратын нәтижелердің талап етілетін нәтижелерден ауытқуын көрсететін қандай да бір функцияның мәнін минималдауға негізделген екендігі көрсетілген. Өз кезегінде, минималдау мақсаты глобалдық минимумды іздеу болады. Бұған жету *оқытудың жинақтылығы* деп аталады. Глобалдық минимумды іздеу итерациялық үдеріс құралы – үйлеспеушілік бетін зерттейтін және онда глобалдық минимум нүктесін табуға талпынатын *оқытатын алгоритммен* жүзеге асырылады.

Қазіргі уақытта бір-бірінен оңтайландыру стратегиясымен және қателік критеріімен ерекшеленетін жүздеген оқыту алгоритмдері бар. Қателік критеріін сәтті таңдау үйлесімділіктің аса тегіс бетіне алып келеді және оқыту есебін жеңілдетеді. Ең қарапайым бағалау, әрбір k оқытатын мысалы үшін d_k қажет етілетін шығыс және y_k нақты алынған шығыс мәндері арасындағы айырым квадраттарының қосындысының K оқытатын мысалдарына орташаландырылған ретінде анықталатын орташа квадраттық қателік (MSE – *Mean Squared Error*) болады:

$$E(w) = \frac{1}{K} \cdot \sum_{k=1}^K (d_k - y_k)^2 \quad (\text{III.2.10.1})$$

MSE бағалау ЖНЖ шығу сигналдары барлық сигналдар үшін бірдей және берілген ε дәлдігімен әйгілі векторлармен сәйкес келуі керек, мұнда ε – *оқытудың сенімділік деңгейі*. Оқытудың деңгейлік сенімділігін есепке алу үшін MSE түрленген бағалауы қолднылады:

$$E(w) = \frac{1}{K} \cdot \sum_{k=1}^K \frac{(d_k - y_k)^2}{\varepsilon}, \quad (\text{III.2.10.2})$$

мұнда K – кіру векторының өлшемі, ε – интерпретация әдістеріне тәуелді әр түрлі диапазонға ие:

$0 < \varepsilon < 1$ – таңбалық интерпретация үшін;

$0 < \varepsilon \leq 2$ – «жеңімпаз барлығын алады» ережесі үшін;

$0 < \varepsilon \leq 2/(K-1)$ – реттік интерпретация үшін.

Оқытудың сенімділік деңгейі ЖНЖ тұрақтылығын қамтамасыз ету үшін енгізіледі. ЖНЖ тұрақтылық критеріі келесі түрде тұжырымдалған: ЖНЖ жұмысы тұрақты болып саналады, егер ЖНЖ-нің шығу сигналдарын ε -нен кіші шамаға өзгерткенде, интерпретация жауабы өзгермесе.

Бұл критерийді ЖНЖ-ны жеделдетіп оқытуды қамтамасыз ету үшін пайдалануға болады. Бағалауды (III.2.10.2) формуласы бойынша есептеу кезінде, интерпретациясы оның мәнін ε -нен аз шамаға өзгермейтіндей дұрыс жауаптар жиынын ғана қолданған жөн.

MSE бағалауын жалпылауға болады, егер $(d_k - y_k)^2$ айырмасының квадраттарын салмақтармен қосуды қолданатын болса:

$$E(w) = \frac{1}{K} \cdot \sum_{k=1}^K v_k (d_k - y_k)^2, \quad (\text{III.2.10.3})$$

мұнда v_k – оқытатын таңдамадағы k -ші мысалдың салмағы.

(III.2.10.3) бағалауын пайдалану оқытатын үлгіден сәйкес v_k салмағын тағайындай отырып, маңызды мысалдарды ерекшелеуге мүмкіндік береді.

Сондай-ақ, MSE бағалауы ақиқатқа жақын максимум критеріімен байланысқан Кульбак-Лейблер бағалауын да пайдаланады:

$$E(w) = \sum_{k=1}^M d_k \log \frac{d_k}{y_k} + (1 - d_k) \log \frac{1 - d_k}{1 - y_k},$$

мұнда M – ЖНЖ шығыстар саны.

ЖНЖ аппараттық жүзеге асыру кезінде (мысалы, IBM PC үшін ZISC Accelerator cards) және нейроимитаторларда жиі қолданылатын ЖНЖ жұмысының сапасын бағалау аса қарапайым болады:

$$E(w) = \sum_{k=1}^M |d_k - y_k|, \quad E(w) = \max_M |d_k - y_k|.$$

Уақыттық қатарларды талдау есептерін шешу үшін орташа салыстырмалы вариация бойынша бағалауды қолданған дұрыс болады:

$$arv(s) = \frac{\sum_{t \in S} (d_t - x_t)^2}{\sum_{t \in S} (d_t - \langle x_t \rangle)^2} = \frac{\sum_{t \in S} e_t^2}{N \sigma^2},$$

мұнда S – уақыттық қатар, e_t – t уақыт мезетінде d_t және x_t ақиқат мәндері арасындағы айырма, $\langle x_t \rangle$ – уақыттық қатардың орташа мәні үшін бағалау, N – уақыттық атардағы деректер саны.

III.2.10-тапсырмалар.

1. Формуланың мәнін түсіндіріңіз $E(w) = \frac{1}{K} \cdot \sum_{k=1}^K (d_k - y_k)^2$.

2. Формуланың мәнін түсіндіріңіз $E(w) = \frac{1}{K} \cdot \sum_{k=1}^K \frac{(d_k - y_k)^2}{\varepsilon}$.

3. Формуланың мәнін түсіндіріңіз $E(w) = \frac{1}{K} \cdot \sum_{k=1}^K v_k (d_k - y_k)^2$.

4. Формуланың мәнін түсіндіріңіз

$$E(w) = \sum_{k=1}^M d_k \log \frac{d_k}{y_k} + (1 - d_k) \log \frac{1 - d_k}{1 - y_k}.$$

5. Формуланың мәнін түсіндіріңіз $E(w) = \sum_{k=1}^M |d_k - y_k|$.

III.2.10-сұрақтар.

1. ЖНЖ оқыту неге негізделеді?
2. Оқыту жинақталығы дегеніміз не?

3. ЖНЖ тұрақтылық критеріі қалай тұжырымдалады?
4. Глобалдык минимумды іздеу қалай жүзеге асырылады?
5. Оқытудың сенімділік деңгейі не үшін енгізіледі?

III.2.10-тестілер.

1. Функцияны минималдау мақсаты не болады?
 - A) глобалдык минимумды іздеу;
 - B) глобалдык максимумды іздеу;
 - C) локалдык минимумды іздеу;
 - D) локалдык максимумды іздеу;
 - E) экстремумдарды іздеу.
2. Оқытудың жинақтылық нәтижесі не болады?
 - A) глобалдык минимумға жету;
 - B) глобалдык максимумға жету;
 - C) локалдык минимумды іздеу;
 - D) локалдык максимумды іздеу;
 - E) экстремумдарды іздеу.
3. MSE деген не?
 - A) орташа квадраттық қате;
 - B) оқыту жылдамдығы;
 - C) оқыту жинақтылығы;
 - D) оқытудың сенімділік деңгейі;
 - E) берілген дәлдік.
4. Кульбак-Лейблер бағалауының дұрыс жазылуын көрсетіңіз?

A) $E(w) = \sum_{k=1}^M d_k \log \frac{d_k}{y_k} + (1-d_k) \log \frac{1-d_k}{1-y_k} ;$

B) $E(w) = \sum_{k=1}^M \log \frac{d_k}{y_k} + \log \frac{1-d_k}{1-y_k} ;$

C) $E(w) = \sum_{k=1}^M d_k \log \frac{d_k}{y_k} + (1-d_k) ;$

D) $E(w) = \sum_{k=1}^M d_k \log \frac{d_k}{y_k} + \log \frac{1-d_k}{1-y_k} ;$

$$E) E(w) = \sum_{k=1}^M d_k + (1-d_k) \log \frac{1-d_k}{1-y_k} .$$

$$5. \quad arv(s) = \frac{\sum_{t \in S} (d_t - x_t)^2}{\sum_{t \in S} (d_t - \langle x_t \rangle)^2} = \frac{\sum_{t \in S} e_t^2}{N \sigma^2} \quad \text{орташа} \quad \text{салыстырмалы}$$

вариация бойынша бағалаудағы $\langle x_t \rangle$ мәні не?

- A) уақыттық қатардың орташа мәні үшін бағалау;
- B) уақыттық атардағы деректер саны;
- C) ақиқат мәндері арасындағы айырма;
- D) оқытудың сенімділік деңгейі;
- E) уақыттық қатар.

III.2.11 Оқыту әдістерін салыстыру

Бұл тақырыпта жасанды нейрондық желілер (ЖНЖ) салыстырулар қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Көптеген зертеушілер жоғарыда айтылған оқыту әдістерін жақсартуды және жалпылауды ұсынған. Бұл саладағы деректер көздерінің көптігінен біз оларды қамти да алмаймыз. Одан бөлек, соңғы баға беру үшін қазіргі таңда өте ерте. Бұл тәсілдердің кейбіреуі расымен іргелі болуы мүмкін, ал біразы уақыт өте келе жойылады.

3-ші қосымшада оқыту алгоритмдері ұсынылған және оларға байланысты оқыту парадигмалары, оқыту ережелері, ЖНЖ-нің архитектурасы, шешілетін есептер келтірілген. Әрбір оқыту алгоритмінің нақты архитектураға бағытталған және шектелген есептер сыныбына арналған өзінің парадигмасы мен оқыту ережесі бар.

4-ші қосымшада нейрондық желісінің сипаттамасы: қолдану аймағы, әртүрлі нейрондық желілердің артықшылықтары мен кемшіліктері берілген (тізім толық емес).

III.2.11-тапсырмалар.

1. Жасанды нейрондық желілерді оқыту алгоритмдері салыстырылатын критерийлерді көрсетіңіз.

2. Жасанды нейрондық желілер салыстырылатын сипаттамаларды көрсетіңіз.

3. Жасанды нейрондық желілер көбінесе қолданылатын пәндік аймақты көрсетіңіз.

1. II.2.11-сұрақтар.

2. Қандай есептерді шешкен кезде жасанды нейрондық желілер қолданылады?

3. Деректерді сығуға арналған есептерді шешу үшін оқытудың қандай парадигмасы тиімді?

4. Қандай жасанды нейрондық желілерде жетіспеушіліктеріне қарағанда артықшылықтары көп?

III.2.11-тестілер.

1. Жасанды нейрондық желілерде сызықты дискриминантты талдауда қандай оқыту ережесі қолднылады?

- A) Хебб;
- B) Больцман;
- C) Қателерді түзету;
- D) Жарыс;
- E) Хопфилд.

2. Жасанды нейрондық желілерде векторлы кванттауда қандай оқыту ережесі қолднылады?

- A) Жарыс;
- B) Больцман;
- C) Қателерді түзету;
- D) Хебб;
- E) Хопфилд.

3. ART1, ART2 оқытуда қандай желілер қолданылады?

- A) ART желілері;
- B) SOM Кохонен желісі;
- C) Хопфилд желісі;
- D) RBF желісі;
- E) Тікелей тарату желісі.

4. Жасанды нейрондық желілерде ассоциативті жадыны оқытуда желі архитектурасы қандай болады?

- A) Хопфилд желісі;
- B) SOM Кохонен желісі;
- C) ART желілері;
- D) RBF желісі;
- E) Тікелей тарату желісі.

5. Больцман машинасының негізгі кемшілігі неде?

- A) оқытудың баяу алгоритмі;
- B) қателерінің матрицасын қолдану керек;
- C) нақты аппроксимация құруға мүмкіндік бермейді;
- D) теориялық тұрғыдан әлсіз;
- E) көлемі қатаң шектелген.

III.3. Нақты нейрондық желілерді оқыту

III.3.1. Персептронды оқыту

Бұл тақырыпта персептронды оқыту сұрақтары қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Сызықтық бөліну критеріін қолдана отырып, бір қабатты персептрон қажет функцияны іске асыруға қабілетті ме екендігін шешуге болады. Нәтиже оң болғанына қарамастан, егер біз салмақтар мен меженің керекті мәнін анықтай алмасақ, оның беретін пайдасы шамалы болады. Егер компьютерде немесе басқа да электрондық құрылғыда осы мәндерді есептеуге арналған процедура жүзеге асырылатын болса, онда персептрон белгілі бір мағынада өздігінен бапталатын болады. Сол себептен салмақты баптау процедурасын оқыту алгоритмі деп атайды және желі «оқытылады» дейді.

Алғашқы рет Розенблатт персептронды мұғаліммен оқыту алгоритмін ойлап тапты және ол персептрон өзі жүзеге асыра алатынның барлығына оқытыла алатындығын дәлелдеді.

Розенблат ұсынған персептронды оқыту алгоритімінде оқытатын таңдама, әрбіреуі үшін өзінің талап етілетін шығысы көрсетілетін, кіру векторларының жиынынан тұрады. Жалпы жағдайда персептронды келесі алгоритм арқылы оқытуға болады:

1. $w_1, w_2, \dots, w_n$ синапстық салмақтарға қандай да бір бастапқы мәндерді беру. Мысалы, нөлге тең.
2. $X = x_1, x_2, \dots, x_n$ кіру мәнін беру және y шығысты есептеу. Егер y дұрыс болса, онда ол 4-ші қадамға, әтпесе 3-ші қадамға өту.
3. Дельта-ережені (төменде берілген) пайдалана отырып, синапстық байланыстар салмақтарының жаңа мәндерін есептеу.
4. Желі оқытатын таңдамадалар векторлардан күтілетін нәтижені шығарғанға дейін немесе ауытқу қандай да бір межеден төмен болғанша, персептронды оқытудың осы алгоритмінің 2-4 қадамдарын қайталау.

Ендеше, егер қандай да бір кіру үлгісінде персептрон сигналы ақиқат болса, онда ештеңені түзетіп керегі жоқ, егер жалған болса салмақтарды түзету жүргізіледі.

Салмақтарды түзетудің келесі ережелері бар:

1. Егер y шығысы қате және $y=0$ болса, онда 1 берілген кірістердің салмақтарын арттыру қажет.

2. Егер y шығысы қате және $y=1$ болса, онда 1 берілген кірістердің салмақтарын азайту қажет.

Бұл ережелердің түсіндірмелері.

Айталық, кіріске қандай да бір оқытатын екілік векторы X берілсін. Бұл векторға $y=1$ шығысы сәйкес келеді және бұл шығыс қате. Сонда, бірлік кірістерге қосылған салмақтар кемітілу қажет, себебі олар қате нәтижені беруге ұмтылады. Осыған ұқсас, егер қандай да бір басқа X оқытатын векторына қате шығыс $y=0$ сәйкес болса, онда 1 берілген кірістердің салмақтары кемітілу қажет.

Енді салмақтарды түзету ережелерінің математикалық моделі болып табылатын, *дельта- ереже* келтірейік.

Қажет етілетін шығыс d және нақты шығыс y арасындағы айырмаға тең δ шамасын енгізейік:

$$\delta = d - y$$

Сонда, түзетуден кейін персептрон салмақтары мынаған тең:

$$w_i(t+1) = w_i(t) - \eta \delta(t) x_i,$$

мұнда x_i – синапстық w_i салмағына сәйкес келетін кіріс шамасы, t – персептронды оқытудың ағымдағы итерация нөмірі, η – салмақтар өзгеруінің орташа мәнін басқаруға мүмкіндік беретін оқытудың жылдамдық коэффициенті.

Көбейтіндіге x_i шамасын қосу кірісінде нөлге сәйкес келетін салмақтардың өзгеруін болдыртпауға мүмкіндік береді.

1-қадам. w_0 ығысуының және $w_I(0)$, ($I=1,2,\dots,N$) синапстық байланыстар салмақтарының қандай да бір кішкентай кездейсоқ сандармен инициалдау.

2-қадам. Персептронға жаңа үлгі ұсыну: $(x_1, \dots, x_N)$ кіру векторын және талап етілетін d шығу сигналын.

3-қадам. Персептронның шығу сигналын есептеу:

$$y(t) = f\left(\sum_{i=1}^N w_i(t) \cdot x_i(t) + w_0\right)$$

3-қадам. Нақты шығыстың дұрыстығын тексеру:

а) егер нақты шығыс дұрыс болса, онда 4-ші қадамға өту, әйтпесе,

б) егер нақты шығыс дұрыс болмаса және нөлге тең болса, онда барлық кірістерді өздеріне сәйкес салмақтарға қосу, әйтпесе,

с) егер нақты шығыс дұрыс болмаса және бірге тең болса, онда әрбір кірісті оған сәйкес салмақтан азайту.

2-ші қадамға өту.

4-қадам. Егер әлі де ұсынылмаған үлгі бар болса, онда 2-ші қадамға өту, әйтпесе персептрон оқытылған болып есептеледі және жұмыс аяқталады.

Осылайша персептрон, оның кірісіне бір-бірлеп үлгілер жиынын бере отырып және барлық үлгілер үшін қажетті шығысқа жеткенге дейін салмақтарды баптай отырып оқытылады.

Дельта-ережесі деп аталатын персептронды оқыту алгоритмін маңызды жалпылау бұл әдісті үздіксіз кірістер мен шығыстарға ауыстырады. Оның қалай алынғанын түсіну үшін персептронды оқыту алгоритмінің 3-ші қадамы δ -ның енгізілуі арқылы жалпы түрде тұжырымдалуы мүмкін. Мұнда δ – қажет шығыс d және нақты шығыс y айырмасына тең, яғни $\delta = (d - y)$. Сонымен бірге:

$\delta=0$ жағдайы шығыс дұрыс болғанда және персептронда ештеңе өзгермесе 3 а) қадамға сәйкес келеді;

$\delta>0$ жағдайы шығыс дұрыс емес болғанда және нөлге тең болса, персептронда белсендіру ұлғайғанда 3 б) қадамға сәйкес келеді;

$\delta<0$ жағдайы шығыс дұрыс емес болғанда және бірге тең болса, персептронда белсендіру азайғанда 3 с) қадамға сәйкес келеді.

Бұл жағдайлардың кез келгенінде персептронды оқыту алгоритмі сақталады, егер δ әрбір x_i кірісінің мәніне көбейтілсе

және бұл көбейтінді синапстық байланыстың сәйкес салмағына қосылса. Жалпылау мақсатымен «оқыту жылдамдығы» $\eta (0 < \eta < 1)$ коэффициенті енгізіледі, ол салмақтар өзгерісінің орташа мәнін басқаруға көмектесетін δx_i -ге көбейтіледі. Бұның алгебралық түрде жазылуы мына түрде болады:

$$w_i(t+1) = w_i(t) + \eta \cdot (d - y(t)) \cdot x_i(t)$$

мұнда $w_i(t)$ —түзетуге дейінгі i -ші салмақтың мәні; $w_i(t+1)$ — түзетуден кейінгі i -ші салмақтың мәні.

Дельта-ереже персептронның синапстық байланысының салмағын әрбір полярлық шығыстың талап етілетін және нақты мәндеріне сәйкес, үздіксіз және бинарлы кіріс және шығыс үшін түрлендіреді. Бұл қасиеттер жаңа қолданымдар жиынын ашуға септігін тигізді.

Розенблатт персептронды оқыту үшін қолданылатын үлгілер (векторлар) екі сызықтық-бөлінетін сыныптардан алынатын болса, онда персептронның алгоритмі жинақты болатындығын және шешімдер жазықтығы осы екі класты бөліп тұратын гипержазықтық түрінде қалыптасатындығын дәлелдеді. Персептронның шығу (есептегіш) қабатының өлшемділігін арттыра отырып және оған бірнеше нейронды қоса отырып, сыныптау есептерін сыныптардың көп санымен шешуге болады.

Персептронды оқыту мұғалімді-жұптар жиынын талап етеді., яғни $\langle X^k$ кіру сигналдарының векторы, d^k шығу сигналының күтілетін мәні $\rangle$

Оқыту (w_{ij} салмақтарды іздеу) мақсаттық функцияны минималдау есебіне алып келеді

$$E(W_i) = \frac{1}{2} \cdot \sum_{k=1}^p (y_i^k - d_i^k)^2$$

Өкінішке орай, персептрон үшін $E(W_i)$ минимумды табу үшін $f(S_i)$ функциясының үздікті орай тек қана нөлдік қатарды оптимизациялау әдісі қолданылмақ. Тәжірибеде персептронды оқыту үшін келесі қарапайым алгоритмнен тұратын *персептрон ережесі* қолданылады.

1. Нейронның w_{Ij} ($j=0,1,2,...,N$) салмақтарының бастапқы мәндері таңдап алынады.

2. Әр оқытатын $\langle X^k, d^k \rangle$ жұбы үшін шығу салмақтарының мәнін айқындаудың циклдер қатары (оның нөмірлерін t арқылы белгілейміз) келесі формула бойынша орындалады:

$$w_{Ij}(t+1) = w_{Ij}(t) + \delta w_{Ij}(t),$$

мұнда:

$$\delta w_{Ij}(t) = 0, \text{ егер } y_I(t) = \delta^k_I;$$

$$\delta w_{Ij}(t) = x^k_j, \text{ егер } y_I(t) = 0, \alpha \delta^k_I = 1;$$

$$\delta w_{Ij}(t) = -x^k_j, \text{ егер } y_I(t) = 1, \alpha \delta^k_I = 0;$$

α –оқыту жылдамдығы немесе қадамы.

Ағымдағы оқытытын жұбты өңдеу барысы мына жағдайда аяқталады:

а) барлығы $\delta w_{Ij}(t) = 0$ болатын циклінде;

б) циклдердің межелік санына жеткеннен соң.

Персептрон ережесі сызықтық персептронды оқыту үшін де, сондай-ақ, межелік белсендіру функциясы бар персептронды оқыту үшін де қолданыла алатын, *дельта-ереже (delta rule)* деп аталатын әмбебап ереженің дербес жағдайынан тұратындығын атап кеткен жөн. Дельта ережесін межелік белсендіру функциясы персептрон үшін 60-ыншы жылдары Ф. Розенблат, ал оны сызықтық желі үшін Б. Уидроу (B.Widrow) және Т. Хофф (T.Hoff) ұсынған.

$$\delta w_{ij}(t) = x_j^k (d_i^k - y_i(t)).$$

Бір қабатты персептронды оқытудың дельта ережесі нейрондық желінің орташа квадраттық қатесін минималауды ұйғарады, ол $p \in N$ кіру үлгілері үшін келесі қатынаспен анықталады:

$$E_s = \sum_{k=1}^p E(k) = \frac{1}{2} \sum_{k=1}^p \sum_{j=1}^m (y_j^k - d_j^k)^2$$

мұнда $E(k)$ – k -шы үлгі үшін нейрондық желінің орташа квадраттық қатесі; y_j^k және d_j^k – k -шы үлгі үшін нейрондық желінің сәйкес шығыстық және эталондық мәндері.

Орташа квадраттық қатені минималдау үшін *градиенттік түсу* әдісі қолданылады. Бір қабатты перцептронды оқытудың екі негізгі тәсілі бар: *тізбектік оқыту (online learning)* және *топтық оқыту (batch learning)*. Тізбектік оқыту кезінде синапстық байланыстардың жетілдіруі нейрондық желіге оқытатын таңдамадан әр үлгіні бергеннен кейін орындалады. Бұл жағдайда градиенттік түсу әдісінде бір кіру үлгісі үшін нейрондық желінің квадраттық қателігі қолданылады:

$$E = \frac{1}{2} \sum_{j=1}^m (y_j - d_j)^2$$

Олай болса градиенттік түсу әдісіне сәйкес синапстық байланыстың салмақтарын және нейрондық желінің межесін уақыт өтуімен келесі өрнек бойынша өзгерту керек:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha \frac{dE}{dw_{ij}(t)},$$

$$w_{0j}(t+1) = w_{0j}(t) - \alpha \frac{dE}{dw_{0j}(t)},$$

α – оқытудың жылдамдығы немесе қадамы.

Осы жерден тізбектік оқыту үшін дельта-ережені алуға болады:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha (y_j - e_j) x_i,$$

$$w_{0j}(t+1) = w_{0j}(t) - \alpha (y_j - e_j) x_i,$$

Тізбектік оқыту кезіндегі бейімді қадамның шамасы мынаған тең:

$$\alpha(t) = \frac{1}{1 + \sum_{i=1}^n x_i^2},$$

мұнда n – нейрондық желінің кіріс саны; $i = 1, 2, \dots, n$.

Топтық оқытуды қолданғанда синапстық байланыстардың жетілдіруінің кірісіне p үлгіні бергеннен кейін орындалады. Бұл жағдайда дельта-ережені келесі түрде ұсынуға болады:

$$w_{ij}(p) = w_{ij}(0) - \alpha(t) \sum_{k=1}^p (y_j^k - e_j^k) x_i^k,$$

$$w_{0j}(p) = w_{0j}(0) - \alpha(t) \sum_{k=1}^p (y_j^k - e_j^k).$$

Топтық үйрету кезіндегі бейімді қадамның шамасы:

$$\alpha_j^k(t) = \sum_l (y_j^l - e_j^l) (1 + \sum_i x_i^l x_i^k), \quad \alpha(t) = \sum_k \sum_l (y_j^k - e_j^k) a_j^k / \sum_k \sum_l (a_j^k)^2.$$

Персептрондар теориясы тура таратылған жасанды нейрондық желілердің көптеген типтерін зерттеу үшін классикалық негізі болды.

Алайда персептрондар кейбір қарапайым есептердің шешімін таба алмайды. Дәл математикалық тәсілдерді пайдаланып персептрондардың нені орындай алатындығын және неге оқытуға болатынын қатаң түрде дәлелдеуге болады. Мұнымен қоса, *бейнелейтіндігін және оқытылатындығын ажырата білген* маңызды. Бейнелейтіндік ұғымы персептронның белгілі бір функцияны моделдеу қабілетіне қатысты. Оқытылатындық осы функцияны іске асыру үшін синапстық байланыстардың салмақтарын баптаудың жүйелік процедурасының бар болуын талап етеді.

III.3.1-тапсырмалар.

1. Персептронды оқытудың қадамдарын атап шығыңыз.
2. Бірқабатты персептронды оқытудың екі негізгі тәсілін атаңыз.
3. Мақсатты функцияның минималдауын шешу үшін арналған формуласын келтіріңіз.
4. Топтық оқыту кезіндегі бейімдік қадамның шамасын анықтау формуласын жазыңыз.

III.3.1-сұрақтар.

1. Дельта-ереже деп аталатын персептронды оқыту алгоритмін жалпылау маңыздылығын түсіндіріңіз?

2. Персептронның шығыс (есептегіш) қабатының өлшемділігінің артуы және оған бірнеше нейронды қосу арқылы неге қол жеткізіледі?

3. Персептронның ережесінің алгоритмі қандай қадамдардан тұрады?

III.3.1-тестілер.

1. Алғашқы рет персептронды мұғаліммен оқыту алгоритмін ойлап тапқан кім?

- A) Розенблатт;
- B) Больцман;
- C) Хебб;
- D) Кохонен;
- E) Хопфилд.

2. Салмақтарды түзету ережелерінің математикалық моделі не?

- A) дельта-ереже;
- B) перспетрон-ережесі;
- C) қателіктер ережесі;
- D) дәлдіктер ережесі;
- E) омега-ереже;

3. Әр оқытылатын $\langle X^k, d^k_I \rangle$ жұбы үшін шығыс салмақтарының мәнін айқындау қандай формула бойынша орындалады?

- A) $w_{Ij}(t+1) = w_{Ij}(t) + \delta w_{Ij}(t)$;
- B) $w_{Ij}(t+3) = 3w_{Ij}(t) + \delta w_{Ij}(t)$;
- C) $w_{Ij}(5t+1) = 5w_{Ij}(t) + \delta w_{Ij}(t)$;
- D) $w_{Ij}(t+1) = w_{Ij}(t) + 2\delta w_{Ij}(t)$;
- E) $w_{Ij}(t+2) = w_{Ij}(t) + \delta w_{Ij}(t)$.

4. Орташа квадраттық қатені минималдау үшін қандай әдіс қолданылады?

- A) градиенттік түсу;
- B) градиенттік ауытқу;
- C) градиенттік оқу;
- D) көлденең түсу;
- E) көлбеу түсу.

5. Топтық үйрету кезіндегі бейімді қадамның шамасын есептеу үшін қайсы формула қолданылады?

A) $\alpha(t) = \sum_k \sum_j (y_j^k - e_j^k) a_j^k / \sum_k \sum_j (a_j^k)^2$;

B) $\alpha(t) = \sum_k \sum_j (y_j^k - e_j^k) / \sum_k \sum_j (a_j^k)$;

C) $\alpha(t) = \sum_k \sum_j (y_j^k - e_j^k) a_j^k + \sum_k \sum_j (a_j^k)^2$;

D) $\alpha(t) = \sum_k \sum_j (y_j^k - e_j^k) a_j^k - \sum_k \sum_j (a_j^k)^2$;

E) $\alpha(t) = \sum_k \sum_j (y_j^k - e_j^k) a_j^k * \sum_k \sum_j (a_j^k)$;

III.3.2. Сигмоидалық нейронды оқыту

Бұл тақырыпта сигмоидалық нейрондарды оқыту қарастырылған, мысалдар келтірілген, тапсырмалар берілген және тест сұрақтары қойылған. Оқу материалдарын дайындау кезінде келесі дерек көздері пайдаланылды [5–38].

Сигмоидалық нейронды оқыту үдерісінде мақсатты оқыту функциясын $E(W_i)$ минималдай алатын, градиенттік әдістер қолданылып «мұғаліммен оқыту» парадигмасы пайдаланылады. Мұнда персептроннан айырмашылығы, мақсатты функцияның минимумын іздеген кезде 1-ретті іздеуді оңтайландырудың әдісі қолданылады. Бұл әдісте w_{ij} салмақ коэффициенттерінің мақсатқа бағытталған өзгерісі теріс градиент бағытымен орындалады:

$$E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r \left(\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k \right)^2 \rightarrow \min_w, \quad (\text{III.3.2})$$

мұнда $y_i^k = f(s_i^k) = f\left(\sum_{i=0}^m w_{ij}^k x_i^k\right)$ – сигмоидалық функция;

$x = [x_0, x_1, \dots, x_n]^T$ – кіру векторы, $x_0=1$, егер поляризациясы бар болса, $x_0=0$, егер поляризация болмаса; x_i – i -ші нейронның шығысындағы оған сәйкес келетін мән.

Үздіксіз белсендіру функциясын қолдану оқыту үдерісінде градиенттік әдістерді пайдалануға мүмкіндік береді. Бәрінен бұрын жылдам түсу әдісін іске асырған оңай, оған сәйкес $W = w_1, w_2, \dots, w_n$ салмақ векторын айқындау мақсатты функцияның теріс градиенті бағытында жүргізіледі. Егер бұл функция (III.3.2) өрнегімен анықталған болса, онда градиенттің j -ші құраушысы келесідегідей болады:

$$\nabla_j E(w_i) = \frac{dE(w_i)}{dw_{ij}} = \sum_{k=1}^r (y_i^k - d_i^k) \frac{dy_i^k}{dw_{ij}} = \sum_{k=1}^r (y_i^k - d_i^k) \frac{df(s_i^k)}{ds_i^k} x_j^k,$$

мұнда w_{ij} салмағын айқындау жылдам түсу әдісі бойынша мақсатты функцияның теріс градиенті бағытында жүргізіледі.

$\delta_i^k = (y_i^k - d_i^k) \frac{df(s_i^k)}{ds_i^k}$ деп белгілесек мынау болады

$$\frac{dE(w_{ij})}{dw_{ij}} = \sum_{k=1}^r \delta_i^k \cdot x_j^k.$$

Осылайша, дискреттік тәсілмен – $\langle x^k, d_i^k \rangle$ әрбір эталондық жұп үшін кіру салмағын айқындау циклдарының серияларымен де сигмоидалдық нейронды оқыту мүмкін болады (персептрон ережесін қараңыз). Мұнда әрбір циклден кейінгі салмақ түзетулері келесі формуламен орындалады:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k,$$

мұнда η – мәні $(0, 1)$ аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті.

Салмақ өзгерісін келесі айырымдық теңдеуді шеше отырып та іске асыруға болады:

$$\frac{dw_{ij}(t)}{dt} = -\mu \delta_i^k x_j^k,$$

мұнда $\mu \in (0, 1)$ – оқыту коэффициенті.

1-ретті іздеуді оңтайландырудың барлық әдістері – глобалды экстремумге жеткізетініне кепілдік бермейтін локалды іздеу әдісі екенін ескерген жөн. Бұл жетіспеушіліктен арылу үшін моментпен оқыту ұсынылған, ондағы салмақ түзетулері келесі түрде орындалады:

$$w_{ij}(t+1) = w_{ij}(t) - \eta \cdot \delta_i^k(t) \cdot x_j^k + \alpha \cdot (w_{ij}(t) - w_{ij}(t)).$$

Формуладағы соңғы қосынды момент деп аталады және алдыңғы циклдегі салмақ өзгерісінің нақтылығын сипаттайды. (α мына $(0,1)$ диапазонында таңдалады) Моментті құрайтын локалды минимум (градиентті құраушы $\delta_i^k(t) \cdot x_j^k$ 0-ге ұмтылған) нүктесіне жақындаған кезде, іздеуді локалды минимум аймағынан күшті перспективалық аймаққа шығарады деген үміт қалыптасқан.

Нейрондарды оқытудың градиентті әдісі тек локалды экстремумға жетуге кепілдік беретіндігін айта кеткен жөн, ол полимодалды $E(w)$ үшін глобалды минимумнан жеткілікті түрде алыс болуы мүмкін. Оның шекарасынан шығу қарапайым жылдам түсу алгоритмін қолдану кезінде мүмкін емес, сондықтан тәжірибеде оны шашумен немесе моментпен қолданады.

Бұл жерде салмақты айқындау үдерісі тек градиентті функцияның ақпараты арқылы ғана анықталмайды, сонымен қатар, салмақ өзгерісінің нақты тренді бойынша да алынады:

$$\Delta w_{ij}(t+1) = -\eta \delta_i x_j + \alpha \Delta w_{ij}(t)$$

мұнда $\alpha \in (0, 1)$ – момент коэффициенті, немесе оңтайландырудың стохастикалық әдісін қолдану.

Салмақты таңдауға моменттің әсері α мәнінің өсуімен арта түседі, сондықтан α -ны оқыту үдерісінде момент басым болмайтындай етіп таңдау керек. Өйткені ол оқыту алгоритмінің тиянақсыз (жинақсыз) болуына әкеліп соғады.

III.3.2-тапсырма.

1. Униполярлы белсендіру функциясының формуласын жазыңыз.
2. Биполярлы белсендіру функциясының формуласын жазыңыз.
3. Униполярлы белсендіру функциясының туындысын есептеу формуласын жазыңыз.
4. Биполярлы белсендіру функциясының туындысын есептеу формуласын жазыңыз.

III.3.2-сұрақтар.

1. Униполярлы белсендіру функциясының графигі қандай түрге ие?
2. Униполярлы белсендіру функциясының туындысының графигі қандай болады?

3. Биполярлы белсендіру функциясының графигі қандай болады?

4. Биполярлы белсендіру функциясының туындысының графигі қандай болады?

III.3.2-тестілер.

1. Мақсатты функцияның минимумын іздеген кезде қандай әдіс қолданылады?

- A) 1-реттегі оңтайландыруды іздеу әдісі;
- B) 2-реттегі оңтайландыруды іздеу әдісі;
- C) 3-реттегі оңтайландыруды іздеу әдісі;
- D) 1-реттегі минимумды іздеу әдісі;
- E) 1-реттегі максимумды іздеу әдісі.

2. $E(W_i) = \frac{1}{2} \sum_{k=1}^r (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{k=1}^r (\sum_{i=0}^m w_{ij}^k x_i^k - d_i^k)^2 \rightarrow \min_w$ мақсатты

функцияны есептеу формуласындағы $y_i^k = f(s_i^k) = f(\sum_{i=0}^m w_{ij}^k x_i^k)$ өрнегі

не болады?

- A) сигмоидалық функция;
- B) тангенстік функция;
- C) кері функция;
- D) синусоидалық функция;
- E) қателіктер функциясы.

3. $w_{ij}(t+1) = w_{ij}(t) - \eta \delta_i^k(t) x_j^k$ формуласындағы η нені береді?

A) мәні (0, 1) аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті;

B) мәні (-1, 1) аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті;

C) мәні (-1, 0) аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті;

D) мәні $(-2, -2)$ аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті;

E) мәні $(-10, 10)$ аралығынан эмпирикалық түрде таңдап алынатын оқыту коэффициенті.

4. Момент арқылы оқыту ұсынылатын салмақ түзетулері қандай түрде орындалады?

A) $w_{ij}(t+1)=w_{ij}(t)-\eta \cdot \delta_i^k(t) \cdot x_j^k+\alpha \cdot (w_{ij}(t)-w_{ij}(t));$

B) $w_{ij}(t+1)=w_{ij}(t)-\eta \cdot \delta_i^k(t) \cdot x_j^k+\alpha \cdot (w_{ij}(t)(t));$

C) $w_{ij}(t+1)=w_{ij}(t)-\eta \cdot \delta_i^k(t) \cdot x_j^k-w_{ij}(t);$

D) $w_{ij}(t+1)=-\eta \cdot \delta_i^k(t) \cdot x_j^k+\alpha \cdot (w_{ij}(t)-w_{ij}(t));$

E) $w_{ij}(t+1)=w_{ij}(t)-\delta_i^k(t) \cdot x_j^k+(w_{ij}(t)-w_{ij}(t)).$

5. $w_{ij}(t+1)=w_{ij}(t)-\eta \cdot \delta_i^k(t) \cdot x_j^k+\alpha \cdot (w_{ij}(t)-w_{ij}(t))$

формуласындағы соңғы бөлікті қалай атайды?

A) момент;

B) үдеу;

C) коэффициент;

D) қате;

E) салмақ.

СӨЗ СОҢЫ

Қазіргі кезде жеке нейрондардың моделдері мен жасанды нейрондар желісі моделдерінің көптеген түрлері жасалды. Бұл жұмыста жеке жасанды нейрондарға қатысты жиналған білімдерді жалпылауға әрекет жасалды және осы мақсатқа сүйене отырып, нейрондардың қолданыста бар моделдері үшін әр түрлі: желі топологиясында нейрондардың орналасуы бойынша, нейрондардың аппараттық жүзеге асырылуына бойынша, нейрон моделінің математикалық сипатталуын әзірлеу дәрежесі бойынша және нейрон тәртібінің сапалық сипатталуын бойынша сыныптаулар келтірілді. Соңғы белгі бойынша нейрондар формалды (оған да өз кезегінде басқа бірнеше белгілері бойынша сыныптаулар келтіріледі), физиологиялық және феноменологиялық.

Қолданыстағы моделдерді шолу нәтижесінде табиғи нейронды өте қарапайым түрде сипаттайтын жасанды нейрондардың жеткілікті саны бар екендігі айқын болады. Олар сыныптау мәселелерін шешуде және т.б. үлгілерді тану саласында қолданыс тапты. Сонымен қатар, нейронды сипаттағанда, нейронның іс-әрекетін сандық сипаттау мақсатын қоятын көптеген моделдер бар. Дегенмен, табиғи нейронның барлық жаңа қасиеттерін бейнелеуге тырысу арқылы жасанды нейронды күрделендіру айтарлықтай прогреске және нейрондық есептердің нәтижелерін жақсартуға алып келетіндігі әлі күнге дейін толық анықталған жоқ. Бірақ бұл үдеріс үнемі дамып, жалғасып отыратындығы белгілі.

ДЕРЕКТЕР КӨЗДЕРІ

1. Кантор Г. Труды по теории множеств. – М.: Наука, 1985. – 430 с. – (Классики науки).
2. Куратовский К., Мостовский А.. Теория множеств / Перевод с английского М. И. Кратко под редакцией А. Д. Тайманова. – М.: Мир, 1970. – 416 с.
3. Выгодский М.Я. Справочник по элементарной математике. 2009 г.; М.: АСТ, 512 с.
4. Зайцев В. В., Рыжков В. В., Сканави М. И. Элементарная математика. Повторительный курс. – Издание третье, стереотипное. – М.: Наука, 1976. – 591 с.
5. Arbib M., ed. The Handbook of Brain Theory and Neural Networks. MIT Press, 1995.
6. Bishop C. M. Neural Networks for Pattern Recognition. Oxford University Press Inc., 2003.
7. Grossberg S. 1969. Some networks that can learn, remember and reproduce any number of complicated space-time patterns. Journal of Mathematics and Mechanics, 19:53-91.
8. Haykin S. Neural Networks: A Comprehensive Foundation. NY: Macmillan, 1994.
9. Hecht-Nielsen R. 1988. Applications of Counterpropagation networks. Neural Networks 1: 131-39.
10. Khanna T. Foundations of neural networks, Addison-Wesley Publishing Co., 1990.
11. Kohonen T. 1988. Self-organization and associative memory. 2d ed. New-York, Springer-Verlag.
12. Kohonen T., "Self-Organizing Maps"(2-nd edition), Springer, 1997.
13. Kröse B., Smagt P. An Introduction to Neural Networks. University of Amsterdam, 1996.
14. Lipman R. An Introduction to computing with neural nets // IEEE Acoustic, Speech and Signal Proc. Magazine, 1987, № 2, L.4-22.

15. Smith L. An Introduction to Neural Networks. Unpublished draft, University of Stirling, 2001. <http://www.cs.stir.ac.uk/~lss/NNIntro/InvSlides.html>.
16. Zurada J.M. Introduction To Artificial Neural Systems. Boston: PWS Publishing Company, 1992.
17. Головкин В. А. Нейронные сети: обучение, организация и применение. Учеб. пособие для вузов / Общая ред. А.И. Галушкина. – М.: ИПРЖР, 2001. – 256 с.
18. Голубев Ю. Ф. Нейросетевые методы в мехатронике. – М.: Изд-во Моск. унта, 2007. – 157 с. – ISBN 978-5-211-05434-9.
19. Горбань А. Н., Дунин-Барковский В. Л. и др. Нейроинформатика. – Новосибирск: Наука, 1998.
20. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: учебное пособие для вузов. – 2-е изд., перераб. И доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – 400 с.
21. Короткий С. Нейронные сети: обучение без учителя. <http://www.orc.ru/~stasson/n3.zIp>
22. Короткий С. Нейронные сети: основные положения. <http://www.orc.ru/~stasson/nl.zIp>
23. Кошимский В.И., Смирнов Д.А, Нейронные сети и их применение в системах управления и связи. – М.: Горячая линия – Телеком, 2002.
24. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – 2-е изд., стереотип. – М.: Горячая линия Телеком, 2002.
25. Минский М. , Пейперт С. Перцептроны. Москва: Мир, 1971.
26. Мкртчян С.О.. Нейроны и нейронные сети. Москва: Энергия, 1971.
27. Нейронные сети. Алгоритм обратного распространения [электронный ресурс]: <http://ai.obrazec.ru/neur-2.html>
28. Розенблатт Ф. Принципы нейродинамики: Перцептрон и теория механизмов мозга: Пер. с англ. – М.: Мир. – 1965. – 480 с.
29. Стариков А. Введение в RBF сети, <http://www.basegroup.ru/neural/rbf.htm>

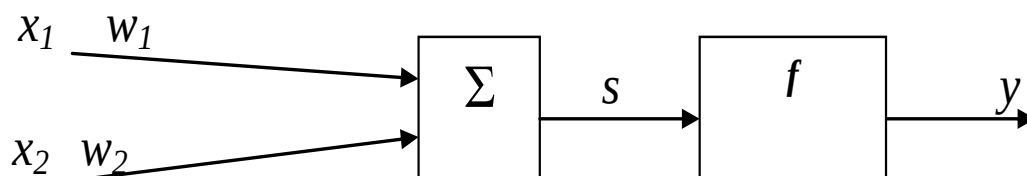
30. Суровцев И.С., Клюкин В.И., Пивоварова Р.П. Нейронные сети.- Воронеж: ВГУ, 1994. 225 с.
31. Тарков М.С. Нейрокомпьютерные системы—Интернетуниверситет информационных технологий—ИНТУИТ.ру, БИНОМ. Лаборатория знаний, 2006. -144с.
32. Терехов С.А. Лекции по теории и приложениям искусственных нейронных сетей. Лаборатория Искусственных Нейронных Сетей НТО-2, ВНИИТФ. Снежинск, 1998.
33. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. М.: Мир, 1992. 240 с.
34. Хайкин С. Нейронные сети: полный курс, 2-е издание, исправленное.: Пер. с англ. – М.: ООО «И.Д. Вильямс», 2006. – 1104.
35. How many kinds of Kohonen networks exist? Internet FAQ Archives. Online Education
36. Hecht-Nielsen, R. (1990), Neurocomputing, Reading, MA: Addison-Wesley, ISBN 0-201-09355-3.
37. Kohonen, T. (1989/1997/2001), Self-Organizing Maps, Berlin – New York: Springer-Verlag. First edition 1989, second edition 1997, third extended edition 2001, ISBN 0-387-51387-6, ISBN 3-540-67921-9
38. Kohonen, T. (1988), Learning Vector Quantization, Neural Networks, 1 (suppl 1), 303.

1-ҚОСЫМША. ПЕРСЕПТРОННЫҢ БЕЙНЕЛІЛІГІ

Бұл қосымшада «Эксклюзивті НЕМЕСЕ» функциясының жүзеге асырылу проблемаларын бір қабатты персептронның екі әдісі арқылы шешуге болатындығы қарастырылған: бірінші әдісте оның мүмкін еместігі дәлелденсе, ал екіншісінде керісінше жүзеге асу мүмкіндігі дәлелденеді.

М. Минский мен С. Пейперт теориялық түрде персептронның көптеген қарапайым есептерді шеше алмайтынын, соның ішінде 0 немес 1 мәнін қабылдайтын және аргументтердің біреуі ғана 1-ге тең болғанда 1 мәнін қабылдайтын, және қарсы жағдайда 0 мәнін қабылдайтын екі аргументі бар «Эксклюзивті НЕМЕСЕ – eXclusive OR (XOR)» бинарлы функциясын жүзеге асыра алмайтынын дәлелдеген соң, бір қабатты персептрондарға деген қызығушылық жоғалып кетті.

XOR функциясының жүзеге асуының мүмкін еместігін П.1.1-суретте ұсынылған екі x_1 және x_2 кіру сигналдарымен бір нейрондық персептронның көмегімен көрсетуге болады.



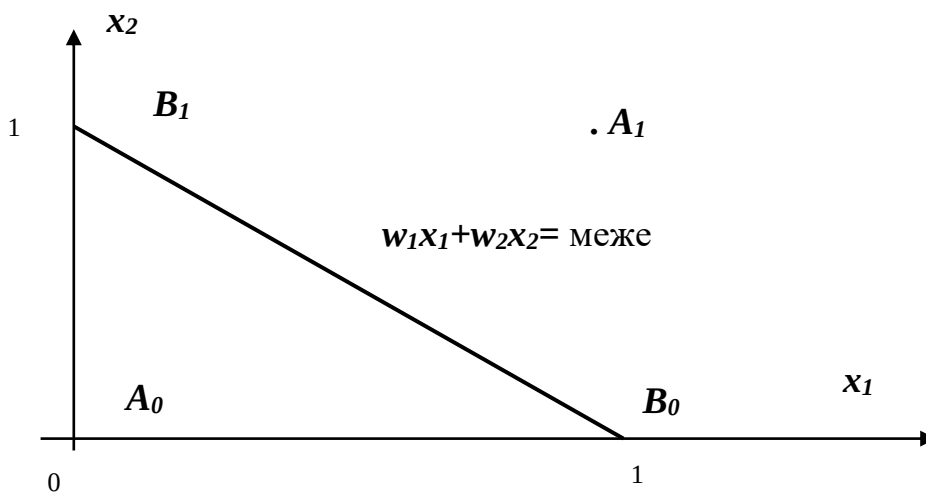
П.1.1-сурет Екі кірісі бар персептрон .

Бұл жерде f белсендіру функциясы y шығысы егер s мәні 0,5-тен үлкен немесе тең болса, 1 мәнін қабылдайтындай, ал s мәні 0,5-тен кіші болса, 0 мәнін қабылдайтындай қарапайым меже болады.

$$y = f(s) = \begin{cases} 1, & \text{егер } s \geq 0,5 \\ 0, & \text{егер } s < 0,5 \end{cases}$$

П.1.2-суретте x_1 және x_2 кіру сигналдарының барлық мүмкін болатын комбинациялары (x_1, x_2) жазықтығындағы төрт нүктеден тұрады: $A_0(0,0)$, $B_0(1,0)$, $B_1(0,1)$, $A_1(1,1)$, сонымен қатар нөлдік

шығыстарды беретін комбинациялар A_0 және A_1 деп белгіленсе, ал бірлік шығыстар сәйкесінше B_0 және B_1 таңбаланады.



П.1.2-сурет. XOR бейнелілік проблемасы

П.1-кестеде XOR функциясы үшін семантикалық моделді (кіріс және шығыс арасындағы талап етілетін байланыс) көрсетеді.

П.1-кесте. XOR функциясының моделі.

Нүктелер	x_1	x_2	Талап етілетін шығыс
A_0	0	0	0
B_0	1	0	1
B_1	0	1	1
A_1	1	1	0

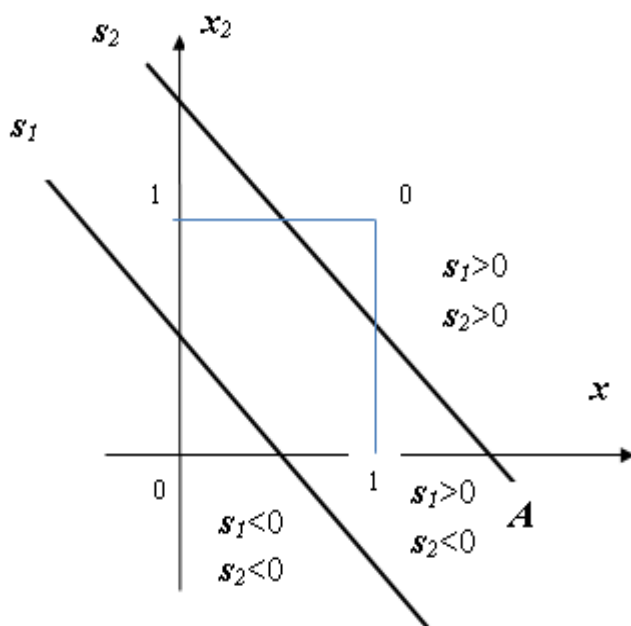
Нейрон келесі есептеулерді орындайды $w_1x_1 + w_2x_2 = s$. Бұл жерде w_1 және w_2 салмақтарының ешқандай комбинациялары П.1-кестесіндегі кіріс пен шығыс арасындағы қатынасты бере алмайды. Осы шектеуді түсіну үшін s -ті 0,5 мәнінде қалдырайық. Бұл жағдайда біз $w_1x_1 + w_2x_2 = 0,5$ теңдеуін аламыз. Бұл теңдеу x_1 және x_2 бойынша сызықты болғандықтан, яғни осы теңдеуді қанағаттандыратын x_1 және x_2 барлық мәндері (x_1, x_2) жазықтығындағы түзу бойында жатады. Бұл түзудегі x_1 және x_2 үшін кез келген мәндер s үшін 0,5 межелік мәнін береді. Кіріс

мәндері түзудің бір жағында межелік мәнінен үлкен s мәнін береді, ендеше $y = 1$. Түзудің екінші жағында $y = 0$ ете отырып, межелік мәнінен кіші s мәнін береді. w_1, w_2 мәндерінің және меженің өзгеруі түзудің көлбеуі мен орналасуын өзгертеді.

П.1-кестесінде берілген XOR функциясын персептрон жүзеге асыру үшін, A_0 мен A_1 нүктелері түзудің бір жағында болатындай, ал B_0 мен B_1 нүктелері түзудің екінші жағында болатындай орналастыру керек. Оның мүмкін еместігін байқау оңай. Салмақ пен меже мәндеріне қандай мәнді берсек те, персептрон XOR логикалық функциясын бейнелеу үшін талап етілетін кіріс және шығыс арасындағы қатынасты жүзеге асыра алмайды.

Алайда, бұл мәселе тек межелік немесе монотонды өспелі үздіксіз сигмоидалық белсендіру функциясына ғана қатысты. Бір максимумды сигналды немесе унимодалды белсендіру (Гаусстың) функциясын қолданғанда бір қабатты персептрон XOR проблемасын шеше алады.

XOR функциясының жүзеге асу мүмкіндігін сигналды белсендіру функциясы бар персептрон арқылы көрсетуге болады. Бұл жағдайда П.1.3-суреттегідей бірліктер мен нөлдік аймақтарын екі түзу арқылы бөліп көрсету керек.



П.1.3-сурет. XOR проблемасының шешімі.

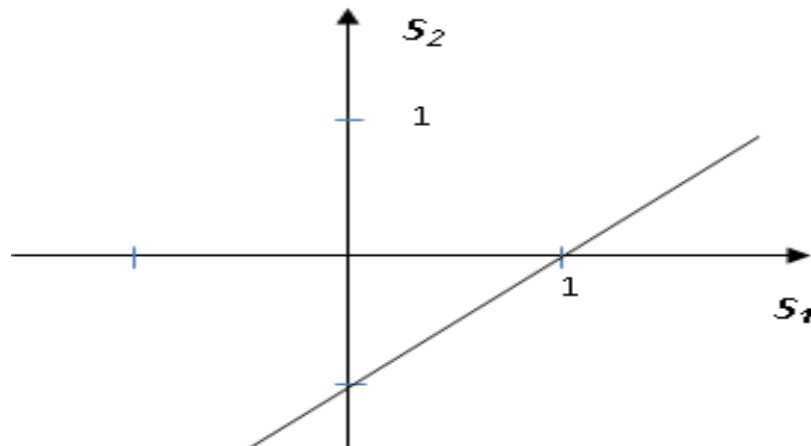
Бірлік сыныбын анықтайтын A аймағы $s_1 > 0$ және $s_2 < 0$ шарттарымен сипатталады.

Түзу теңдеуі келесідей түрде болады:

$$s_1 = x_2 + x_1 - 0.5,$$

$$s_2 = x_2 + x_1 - 1.5.$$

Бұл жерден П.1.4-суретте көрсетілген $s_2 = s_1 - 1$ тәуелділікті алуға болады.

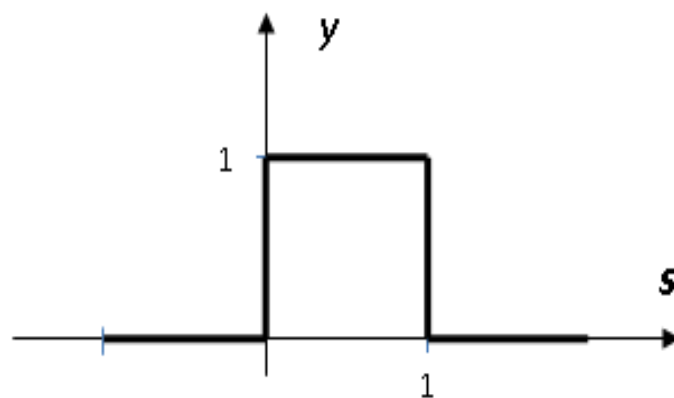


П.1.4 –сурет s_1 -дің s_2 -ден тәуелділігі

Егер $s_1 > 0$ және $s_2 < 0 \Rightarrow 0 < s_1 < 1$ болса, онда осыдан келесі белсендіру функциясын алуға болады:

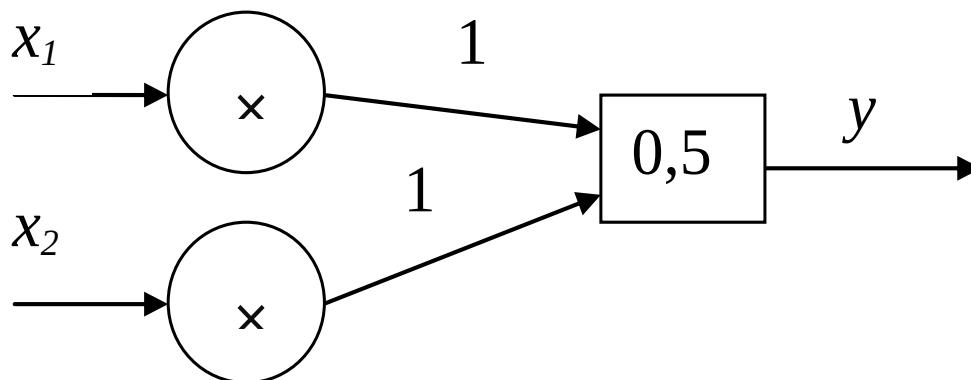
$$y = f(s) = f(s_1) = \begin{cases} 1, & \text{егер } 0 < s < 1, \\ 0, & \text{кері жағдайда} \end{cases}$$

П.1.5-суретте осы сигналды функцияның графигі көрсетілген.



Сурет П.1.5. Сигналды белсендіру функциясы.

Нәтижесінде *XOR* есебін шешу үшін бір қабатты персептрон сәйкесінше 1 мен 0,5 тең салмақ пен межеге ие болады. *XOR* функциясын жүзеге асыру үшін бір қабатты персептронның құрылымдық сұлбасы П.1.6-суретте көрсетілген.



П.1.6-сурет Функцияны жүзеге асыратын бір қабатты персептрон

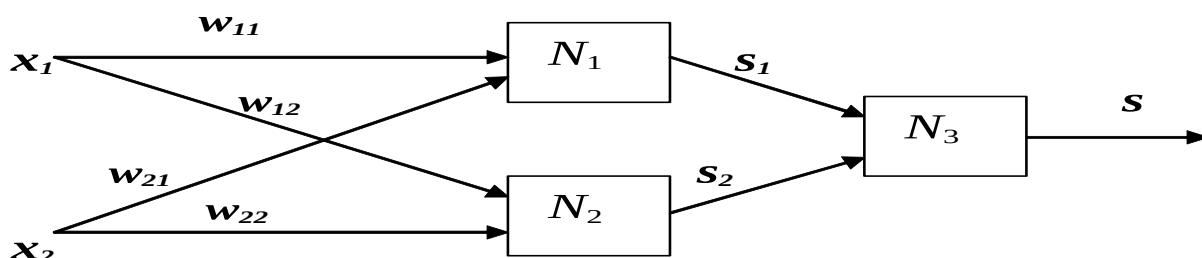
Осылайша, сигналды белсендіру функциясын қолданғанда бір қабатты персептрон *XOR* есебін шеше алады, себебі ол бейнелердің кіріс кеңістіктерін екі түзу арқылы сыныптарға бөледі.

Радиалды базистік белсендіру функциясымен бір қабатты персептронды қолданғанда осы есепті шешу үшін персептронның шығу мәнін, егер ол нақты мәннен үлкен болса, 1-ге тең деп қабылдау керек.

2-ҚОСЫМША. ПЕРСЕПТРОН МЫСАЛДАРЫ

Берілген қосымшада екі қабатты және үш қабатты персептрон мысалдары қарастырылған.

Екі қабатты персептрон екі кірісі бар, бірінші қабаттағы екі нейронға түсірілген, ол екінші қабаттағы жалғыз нейронмен қосылған. Бұл персептронның құрылымдық сұлбасы П.2.1-суретте көрсетілген.

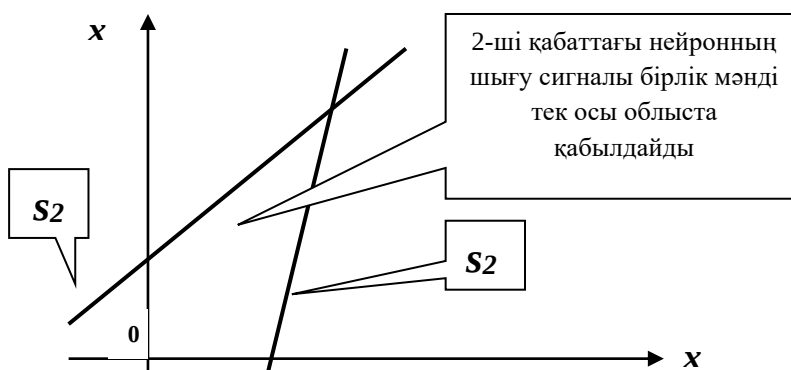


Шығу нейроны үшін шегі 0,75-ке тең, ал оның екі салмағы да 0,5-ке тең болсын. Сонда шекті арттыру үшін шығысында бірлік пайда болу үшін, шығысындағы бірінші деңгейлі екі нейронда бірге тең болу керек.

$$s = 0,5 * s_1 + 0,5 * s_2 = 0,5 * 1 + 0,5 * 1 = 1$$

П.2.1-сурет. AND функциясын іске асыру үшін екі қабатты персептрон.

екі жартылай жазықтыққа бөледі, біреуі жоғарғы сызықтан төмендегі кірістер үшін, басқасы—төменгі сызықтан жоғарыдағы кірістер үшін бірлік шығысты қамтамасыз етеді. Екінші қабаттағы нейронының шығу сигналы тек V-түріндегі аймақ ішінде ғана бірлік мәнге тең.



П.2.2-сурет. Жазықтықты жарты жазықтықтарға бөлу

Осыған ұқсас, бірінші қабатта әрі қарай жазықтықты бөлетін және үшбұрышты формадағы аймақты құрумен үш нейрон қолданылуы мүмкін. Кіру қабатына жеткілікті нейрондар санынан қосумен кез келген қажетті формадағы дөңес көпбұрыш құрылуы мүмкін. Олар берілген сызықтармен аймақтардың үстінен *AND* операциясының көмегімен құрылғандықтан, барлық мұндай көпжақтар дөңес, ендеше тек дөңес аймақтар ғана пайда болады. Дөңес аймақты құрамайтын нүктелер, екі қабатты желідегі жазықтықтың басқа нүктелерінен бөліне алмайды.

Екінші қабаттағы нейрон *AND* функциясымен шектелмейді. Ол салмағы мен шегі қолайлы таңдап алынған жағдайда басқа да көптеген функцияларды жүзеге асыра алады. Мысалы, бірінші қабаттың нейронының кез келген бірлік шығысы, *OR* логикасын іске асыра отырып, екінші қабаттың нейронының шығысында бірдің пайда болуына алып келетіндей етіп жасауға болады. Екі айнымалыдан 16 екілік функциялар бар. Егер қолайлы түрде салмақ пен шек таңдап алынса, онда олардың ішінен 14-ін (*XOR* және *NOTXOR*-дан басқаларын) орындауға болады.

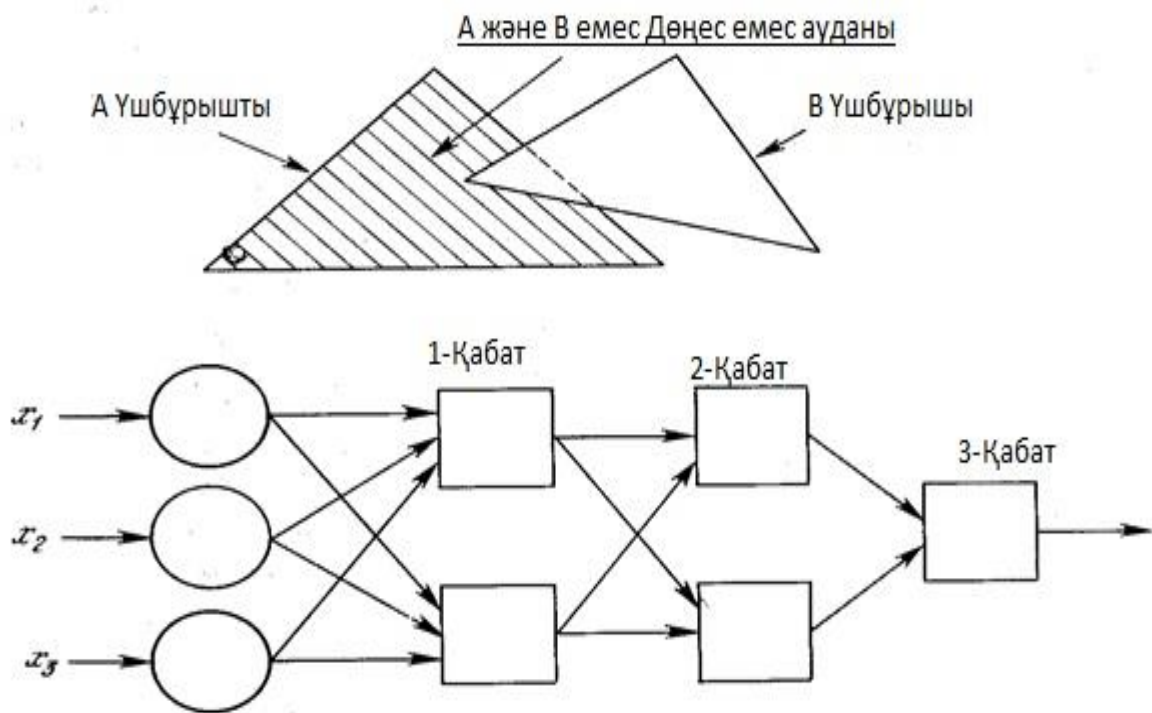
Екі қабатты персептронның математикалық моделі келесі өрнекпен көрсетіледі:

$$y_j = f_2 \left(w_{20j} + \sum_{j=1}^J w_{2j} f_1 \left(w_{10i} + \sum_{i=1}^I w_{1i} x_i \right) \right),$$

мұнда f_1 , f_2 – бірінші және екінші қабат белсендендіру функциялары; I – бірінші қабаттағы нейрондар саны; J – екінші қабаттағы нейрондар саны; w_{0i}^1 – бірінші қабаттағы I -ші нейронның бастапқы жағдайы; w_{0i}^2 – екінші қабаттағы j -ші нейронның бастапқы жағдайы; w_{1i} және w_{2j} – бірінші қабаттағы I -ші нейронның және екінші қабаттағы j -ші нейронның коэффициенттері, x_i – кіру векторының I -ші координатасы, y_j – шығу векторының j -ші координатасы.

Үш қабатты персептрон аса жалпы нейрондық желі болады. Оны жіктейтін мүмкіндіктер тек жасанды нейрондар және

салмақтар санымен ғана шектелген. Дөңестікке шектелулер жоқ. Енді, үшінші қабаттағы нейрондар кіріс ретінде дөңес көпбұрыштар жиынын қабылдайды, және олардың логикалық комбинациясы дөңес емес болуы мүмкін. П.2.3-суретте, A және B екі үшбұрышы, « A және B » функциялары арқылы қиыстырылып, дөңес емес аймақты беретіндігі көрсетілген. Нейрондарды және салмақтарды қосу кезінде көпбұрыштар саны шексіз артуы мүмкін. Бұл кез келген формадағы аймақты кез келген дәлдікпен жуықтауға мүмкіндік береді. Сонымен қатар, екінші қабаттың барлық шығыс аймақтары қиылысуы міндетті емес. Осылай болғандықтан, кіру векторы оның біріне тиісті болған кезде, шығысында әр сайын бірді бере отырып, әр түрлі дөңес және дөңес емес аймақтарды біріктіру мүмкін болады.



П.2.3-сурет 3 қабатты желімен берілген шешімдердің ойыс аймағы

Шығу сигналының мәніне қарай: егер кіру сигналы A сыныбына тиесілі болса $+1$; егер кіру сигналы B сыныбына тиесілі болса -1 шешімі қабылданады.

Үш қабатты персептронның математикалық моделі келесі өрнекпен көрсетіледі:

$$y_j = f_3(w_{30k} + \sum_{k=1}^K w_{3k} f_2(w_{20j} + \sum_{j=1}^J w_{2j} f_1(w_{10i} + \sum_{i=1}^I w_{1i} x_i))),$$

мұнда w_{10i} , w_{1i} , w_{20j} , w_{2j} , w_{30k} , w_{3k} – бірінші қабаттың I -ші нейронының, екінші қабаттың j -ші нейронының, үшінші қабаттың k -ші нейронының сәйкесінше бастапқы жағдайлары мен салмақтық коэффициенттері; I – бірінші қабаттағы нейрондар саны; J – екінші қабаттағы нейрондар саны; K – үшінші қабаттағы нейрондар саны;

Персептронды оқыту. Бірінші сыныпқа тиесілі s_1 нүктелер жиыны (шығу сигналы – 1) және екінші сыныпқа тиесілі s_2 нүктелер жиыны (шығу сигналы – 0) берілген. Белгілер кеңістігін сызықты түрде екі бөлікке бөлетін гипержазықтық теңдеуін құру керек, W векторын табу және w_0 – ығысуын табу керек.

Берілген:

s_1 : (-2, 4); (0, 3); (0, 0);

s_2 : (-2, 0); (-1, -5); (3, -6); (2, -5).

Табу керек. Белгілер кеңістігін екі бөлікке бөлетін, гипержазықтықтың теңдеуін.

Шешімі:

2. Сызықтық бөлу ережесін масса орталығын бөлу көмегімен қарастырамыз.

Салмақ векторының бастапқы мәнін анықтаймыз.

$$W = \left(\sum_{i=1}^n X_1^i - \sum_{j=1}^m X_2^j \right) / (n+m),$$

онда X_1^i – бірінші сынып жиынының векторы, ал X_2^j – бірінші сынып жиынының векторы.

$$W = \left(\sum_{i=1}^3 X_1^i - \sum_{j=1}^4 X_2^j \right) / (3+4) =$$

$$= \frac{(2,4)+(0,3)+(0,0)-(-2,0)-(-1,-5)-(3,-6)-(2,-5)}{7}$$

$$= \frac{(2,7)-(2,-16)}{7} = \frac{(0,23)}{7} = (0, 3,285)$$

Салмақ векторының бастапқы мәні $w_0=0$; $w_1=0$; $w_2=3,285$.

3. Персептронды оқытуды жеке мысалдар бойынша қарастырайық, ол үшін оқытатын таңдаманы аламыз және кезекпен шығу сигналының мәнін есептейміз.

$$X_1^1 = (2,4), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 0 + 0 + 4 \cdot 3,285 = 13,14 ,$$

$$y = f(s) = \begin{cases} 1, & \text{егер } s > 0 \\ 0, & \text{егер } s \leq 0 \end{cases}, \Rightarrow y = 1;$$

$$X_1^2 = (0,3), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 0 + 0 + 3 \cdot 3,285 = 9,857$$

$$y = f(s) = 1;$$

$$X_1^3 = (0,0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 0 + 0 + 0 \cdot 3,285 = 0 ,$$

$$y = f(s) = 0;$$

Салмақтарды түзетуді мына формула бойынша орындаймыз
 $w_i(t+1) = w_i(t) + \gamma x_i$, $\gamma = 0.8$ шығатыны:

$$w_0(t+1) = w_0(t) + \gamma x_0 = 0 + 0,8 \cdot 1 = 0.8 ;$$

$$w_1(t+1) = w_1(t) + \gamma x_1 = 0 + 0,8 \cdot 0 = 0 ;$$

$$w_2(t+1) = w_2(t) + \gamma x_2 = 3,285 + 0,8 \cdot 0 = 3,285 .$$

Енді $W = (0,8; 0; 3,285)$

$$X_1^1 = (2,4), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 2 \cdot 0 + 4 \cdot 3,285 = 13,942$$

$$y = f(s) = 1;$$

$$X_1^2 = (0,3), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 0 \cdot 0 + 3 \cdot 3,285 = 10,65 ,$$

$$y = f(s) = 1;$$

$$X_1^3 = (0,0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 0 + 0 \cdot 3,285 = 0,08$$

$$y = f(s) = 1;$$

$$X_2^1 = (-2, 0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + (-2) \cdot 0 + 0 \cdot 3,285 = 0,8,$$

$$y = f(s) = 1;$$

Деректердің екінші тоб үшін жауап 0 мәні сәйкес келу керек.

Салмақтарды түзетуді мына формула бойынша орындаймыз

$$w_i(t+1) = w_i(t) - \gamma x_i, \gamma = 0.8 \text{ шығатыны:}$$

$$w_0(t+1) = w_0(t) - \gamma x_0 = 0,8 - 0,8 \cdot 1 = 0;$$

$$w_1(t+1) = w_1(t) - \gamma x_1 = 0 - 0,8 \cdot (-2) = 1,6;$$

$$w_2(t+1) = w_2(t) - \gamma x_2 = 3,285 - 0,8 \cdot 0 = 3,285.$$

$$\text{Енді } W = (0; 1,6; 3,285)$$

$$X_1^1 = (2, 4), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0 + 2 \cdot 1,6 + 4 \cdot 3,285 = 16,34,$$

$$y = f(s) = 1;$$

$$X_1^2 = (0, 3), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0 + 0 \cdot 1,6 + 3 \cdot 3,285 = 9,855,$$

$$y = f(s) = 1;$$

$$X_1^3 = (0, 0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 3,285 = 0,$$

$$y = f(s) = 0;$$

Салмақтарды түзетуді мына формула бойынша орындаймыз

$$w_i(t+1) = w_i(t) + \gamma x_i, \gamma = 0.8 \text{ шығатыны:}$$

$$w_0(t+1) = w_0(t) + \gamma x_0 = 0 + 0,8 \cdot 1 = 0,8;$$

$$w_1(t+1) = w_1(t) + \gamma x_1 = 1,6 + 0,8 \cdot 0 = 1,6;$$

$$w_2(t+1) = w_2(t) + \gamma x_2 = 3,285 + 0,8 \cdot 0 = 3,285.$$

$$\text{Енді } W = (0,8; 1,6; 3,285)$$

$$X_1^1 = (2, 4), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 2 \cdot 1,6 + 4 \cdot 3,285 = 17,14,$$

$$y = f(s) = 1;$$

$$X_1^2 = (0, 3), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 0 \cdot 1,6 + 3 \cdot 3,285 = 10,655,$$

$$y = f(s) = 1;$$

$$X_1^3 = (0,0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 0 \cdot 1,6 + 0 \cdot 3,285 = 0,8$$

$$y = f(s) = 1;$$

$$X_2^1 = (-2,0), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + (-2) \cdot 0 + 0 \cdot 3,285 = -2,4,$$

$$y = f(s) = 0;$$

$$X_2^2 = (-1,-5), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + (-1) \cdot 1,6 + (-5) \cdot 3,285 = -17,225,$$

$$y = f(s) = 0;$$

$$X_2^3 = (3,-6), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 3 \cdot 1,6 + (-6) \cdot 3,285 = -14,11,$$

$$X_2^4 = (3,-6), s = 1 \cdot w_0 + x_1 w_1 + x_2 w_2 = 1 \cdot 0,8 + 2 \cdot 1,6 + (-5) \cdot 3,285 = -12,425,$$

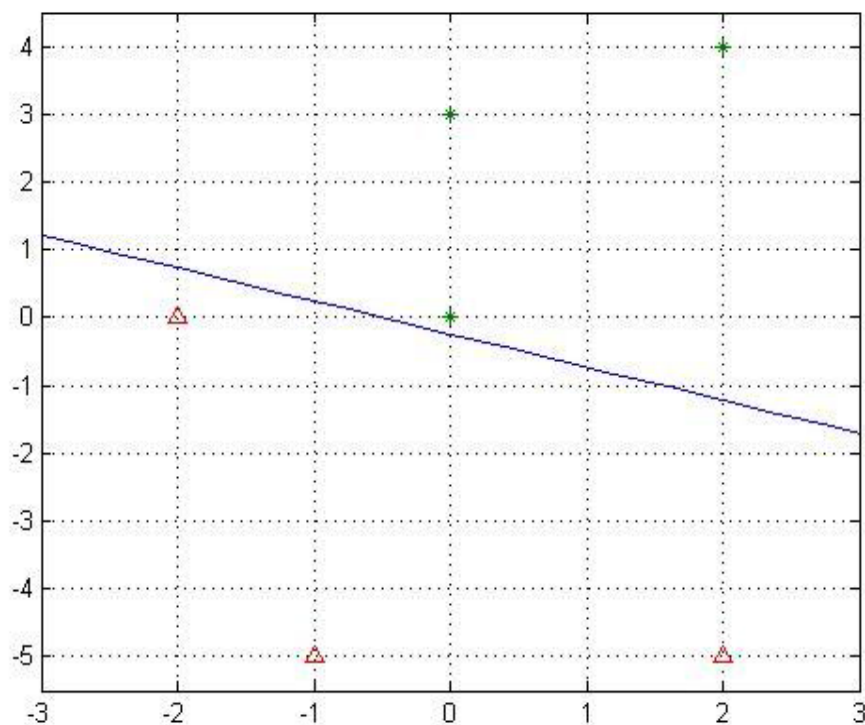
$$f(s) = 0.$$

Бірқабатты персептрон оқытылды. Жауабы: $W = (0,8; 1,6; 3,285)$

Бөліп тұрған гипержазықтық теңдеуі:

$$0,8 + 1,6 \cdot x_1 + 3,285 \cdot x_2 = 0.$$

Төменгі суретте гипержазықтықты бөлуші көрсетілген.



3-ҚОСЫМША. ЖАСАНДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕР СИПАТТАМАЛАРЫ

Оқыту алгоритмі	Оқыту парадигмалары	Оқыту ережелері	Желі архитектурасы	Шешілетін есептер
Персептронды оқыту алгоритмдері. Кері тарату	Мұғаліммен	Қателерді түзету	Бір қабатты және көп қабатты персептрон	Үлгілерді сыныптау, Функциялар аппроксимациясы, Болжау, бсқару
Больцман алгоритмін оқыту		Больцман	Рекуррентті	Үлгілерді сыныптау
Сызықты дискриминантты талдау		Хебб	Көп қабатты, тікелей тарату	Деректерді талдау, Үлгілерді сыныптау
Векторлы кванттау		Жарыс	Жарыс	Сынып ішінде категориялау. Деректерді сығу
ARTMap			ART желісі	Үлгілерді сыныптау

Саммон проекциясы	Мұғалімсіз	Қателерді түзету	Көп қабатты, тікелей тарату	Сынып ішінде категориялау, Деректерді талдау
Негізгі компоненттерді талдау		Хебб	Тікелей тарату және жарыс	Деректерді талдау, Деректерді сығу
Ассоциативті жадыны оқыту			Хопфилд желісі	Ассоциативті жад
Векторлық кванттау		Жарыс	Жарыс	Категориялау, Деректерді сығу
SOM Кохонен			SOM Кохонен	Категориялау, Деректерді талдау
ART1, ART2			ART желілері	Категориялау
RBF алгоритмін оқыту	Аралас	Қателерді түзету және жарыс	RBF желісі	Үлгілерді сыныптау, Функцияларды аппроксимациялау, Болжау, басқару.

4-ҚОСЫМША. ЖАСАНДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕРДІҢ ТҮРЛЕРІ

Жасанды нейрондық желілердің түрлері	Шешілетін есептер	Кемшіліктері	Артықшылықтары
Жасанды резонанс–1, ART-1 Network (Adaptive Resonance Theory Network–1)	Бейнелерді тану, кластерлік талдау	Желінің жұмысы кезінде нейрондар санының шексіз өсуі. Шу болған кезде бақыланбай эталондар санының өсуімен байланысты айтарлықтай мәселелер туындайды.	Мұғалімсіз оқып үйрену
Кохонен желісі (Kohonen's Neural Network)	Кластерлі талдау, бейнелерді тану, сыныптау.	Кластер саны алдын ала белгілі болса, желіні кластерлік талдау жасау үшін қолдануға болады.	ART желісіне қарағанда, Кохонен желісі кедергілер болған жағдайда да жұмыс істеуге қабілетті, себебі сынып саны тұрақты болады, салмағы баяу түрлендіріледі, және салмақты орнату оқыту аяқталғаннан кейін бітеді. (ABT желісінде баптау үздіксіз жалғасады)

Екі бағытты ассоциативті жады (ЕАЖ). BI-Directional Associative Memory (BAM)	Ассоциативті жад, бейнелерді тану.	ЕАЖ көлемі қатаң шектелген, сонымен қатар ЕАЖ жұмыс істеу кезінде күтпеген жағдайлар орнауы мүмкін. Жалған жауаптар шығуы мүмкін.	Автоассоциативті жадпен салыстырғанда (мысалы, Хопфилд желісімен), екі бағытты ассоциативті жады жалпы жағдайда әртүрлі көлемді X және Y векторлары арасында ассоциация құруға мүмкіндік береді. Осы мүмкіндіктер есебінен гетероассоциативті жады автоассоциативті жадқа қарағанда, қосымшаның ауқымды сыныптарында қолданылады. Синапстық байланыс салмағын құру үдерісі қарапайым және жылдам. Жұмыс істеу барысында желі жылдам тарайды.
Больцман машинасы	Бейнелерді тану, сыныптау.	Оқытудың баяу алгоритмі.	Алгоритм желі үшін локалды минимумнан шығу мүмкіндігін береді.
Кері тарату	Бейнелерді тану,	Көп критерийлі есеп кері	Кері тарату – көп қабатты

(Neural Network with Back Propagation Training Algorithm)	сыныптау, болжам жасау, сөйлеу синтезі, бақылау, бейімді басқару, экспертті жүйелерді құру.	таратудың оңтайландыру әдісінде бір критерилілер жиыны ретінде қарастырылады – әрбір итерцияда желі параметрлерінің мәні өзгеріп отырады, жұмысты тек таңдалған оқыту мысалымен жақсартады. Мұндай қадам әрине оқыту жылдамдығын тежейді. Кері таратудың классикалық әдісі сызықтық сәйкестік алгоритміне жатады. Сәйкестіктің жылдамдығын арттыру үшін екінші туынды функция қателерінің матрицасын қолдану керек.	нейрондық желілерді оқытудың бірінші тиімді алгоритмі. Оқытудың әйгілі алгоритмдер қатарында, оның көмегімен көптеген тәжірибелік есептер шешілген және шешілуде.
Қарама-қарсы тарату желісі (Counter Propagation Network)	Бейнелерді тану және қайта қалпына келтіру (ассоциативті жад), деректерді қысу (жоғалтулармен), статикалық талдау.	Желі нақты аппроксимация құруға мүмкіндік бермейді. Бұл желі кері тарату желілеріне айтарлықтай артықшылық беріп қояды. Желінің жетілдіру жұмысы теориялық тұрғыдан	Қарама-қарсы тарату желісі қарапайым. Ол кіріс сигналдар жиынынан статикалық сипаттамаларды алу мүмкіндігін береді. Желі жылдам оқытылады. Оның оқыту

		әлсіз болып келеді.	жылдамдығы кері тарату желісінен 100 есе аз болып келеді. Бұл желі аппроксимация бастамасы жылдамдықты талап ететін қосымшалар үшін пайдалы. Желі практикалық есептерді шешетін функцияны құруға және керісінше мүмкіндік береді.
Delta Bar Delta желісі	Бейнелерді тану, сыныптау.	DBD стандартты алгоритмі моменттерге негізделген эвристиканы қолданбайды. Тіпті, жай ғана сызықтық коэффициенттерді арттырудың өзі оқыту жылдамдығының едәуір өсуіне алып келеді, ол салмақтар кеңістігінде секірістердің пайда болуына әкеп соғады. Коэффициенттерді геометриялық кемітудің өзі кейде жеткіліксіз жылдам.	DBD парадигмасы алгоритмдерді кері тарату есебінен конвергенциялау үдерісін арттыру үшін параметрлердің және оқыту үдерісінде салмақтарының өзгеруі туралы қосымша ақпарат береді.
Тікелей	Хэмминг желісімен	Желі қабаттарының саны – кіру	MAXNET желісінен өзгеше

байланыстар арқылы максимумды іздеу (Feed-Forward MAXNET)	бірге, нейрожелілік жүйе құрамында бейнелерді тану.	сигналдарының көлемінің артуына байланысты болады.	циклдық функционалды желісіне қарағанда, бұл желіде шешімді алу үшін есептеу көлемі алдын ала белгілі болады.
Гаусс сипаттауышы (Neural Gaussian Classifier)	Бейнелерді тану, сыныптау.	Примитивті бөлетін беттер (гипержазықтықтар) бейнелерді танудың тек ең қарапайым есептерін шешуге мүмкіндік береді. Кіру сигналдарын априорлы орналастыру әртүрлі сыныптарға сәйкес екендігі белгілі.	Моделдің программалы немесе аппаратты жүзеге асуы өте қарапайым. Ығысу мен синапстық байланыстар салмағының құрылуы қарапайым және жылдам алгоритм бойынша орындалады.
Генетикалық алгоритм (Genetic Algorithm)	Бейнелерді тану, сыныптау, болжау.	Түсінуге және программалық жүзеге асуы күрделі.	Генетикалық алгоритм бейімделу рельефтерінің глобалды минимумін іздеуде өте тиімді, себебі осы аймақта нейрондық желілер параметр мәндерінің көп бөлігі зерттеледі. Оқыту жылдамдығы айтарлықтай жоғары, дегенмен градиентті

			алгоритмнің жинақталу жылдамдығына қарағанда баяуырақ болып келеді. Генетикалық алгоритм нейрондық желілер параметрлерінің дискретті мәндерін пайдалануға мүмкіндік беретіндігі – нейрондық желілердің аппаратты жүзеге асуын жеңілдетеді және жалпы оқыту уақытын қысқартады.
Хэмминг желісі (Hamming Net)	Бейнелерді тану, сыныптау, ассоциативті жад, кедергі кезінде сигнал берудің сенімділігі .	Желі дұрыс тануға қабілетті(сыныптау) тек шулы кіру сигналдарын ғана . Тек екілік кіру сигналдарын пайдалану мүмкіндігі айтарлықтай қолдану аясын шектейді.	Желі өте қарапайым және тез жұмыс істейді. Шығу сигналы (шешімі) нейрондардың бір ғана қабаты арқылы сигналдарды өту нәтижесінде қалыптасады. Салыстыру үшін, көп қабатты желі сигналы бірнеше рет желі нейрондары арқылы өтеді, және шешім алу үшін қажетті итерация саны алдын ала белгілі болмайды. Хопфилд желісіне

			қарағанда, Хэмминг желісінің сыйымдылығы кіріс өлшеміне байланысты емес, ол нейрондардың санына дәл тең. Хопфилд желісіне кіру сигналының өлшемі 100, 10 үлгілерін сақтауға болады, осылайша ол 10000 синапсовқа ие болады. Хэмминг желісінде осындай сыйымдылықпен 1000 ғана синапсы болады.
Хопфилд желісі (Hopfield Network)	Ассоциативті жад, мазмұны бойынша адрестеу; бейнелерді тану; оңтайландыру мәселелері (комбинаторлық оңтайландыруды қоса алғанда)	Желі сыйымдылыққа шағын. Сонымен қатар, толтырылған бейнелермен қатар желіде олардың жағымсыз жақтары да сақталады. Өлшемі мен кіру сигналдарының түрі өлшем және шығу сигналының түріне сәйкес келеді. Бұл айтарлықтай бейнелерді тануда мәселелерді желіде пайдалануды шектейді. Сигналының өлшемін арттыру	Бұрмаланған сигналдарды қалпына келтіруге мүмкіндік береді.

		арқылы синапстардың саны квадратты ұлғайтылады.	
Кіру жұлдызы (Instar)	Желілерде бейнелерді тану үшін қолданылуы мүмкін.	Әр жұлдыз тым қарапайым функцияны іске асырады. Мұндай жұлдыздардан кез келген берілген бейнені суреттейтін нейрондық желі құру мүмкін емес. Бұл кіру жұлдыздарының практикалық қолдануын шектейді.	Кіру жұлдызы биологиялық нейрондық желілердің компоненттерінің кейбір функцияларын моделдейді және мидың жеке бөліктерінің жеткілікті түрде жақсы моделі бола алады. Практикалық мәселелерді шешуде кіру жұлдыздары қарапайым және тез оқытылатынын желілерді құру үшін қолданылуы мүмкін.
Шығу жұлдызы (Outstar)	Бейнелерді тану үшін нейрондық желілер компоненті ретінде пайдаланылуы мүмкін.	Әр жұлдыз тым қарапайым функцияны іске асырады. Осы жұлдыздардан тұратын нейрондық желілердің есептеу мүмкіндіктері шектелген.	Практикалық есептерді шешу кезінде шығу жұлдыздары жылдам қарапайым оқыту желілерін құру үшін қолданылуы мүмкін.
Күйдіруді имитациялайтын нейрондық желі (Neural Networks)	Күйдіруді имитациялайтын алгоритм көмегімен әр түрлі мөлшердегі	Үлкен мөлшерлі нейрондық желілерді оқыту үдерісінде жинақтылық деңгейінің төмен болуы.	Алгоритмге қойылған «жылу тербелістері», локалдық минимумдарды болдырма мүмкіндігін береді. Күйдіруді

wIth SImlated AnnealIng TraInIng AlgorIthm)	векторлардың бейнесін құруға болады . Оған бейнелерді тану, бейімді басқару, көп параметрлі идентификация, болжау және диагностикала есептері қойылады.		имитациялайтын алгоритм бейімді рельефті нейрондық желіде глобалдық оптимумды іздеу үшін қолданылуы мүмкін екендігі көрсетілген.
Максимумды іздеу желісі (MAXNET)	Нейрожелілік жүйеде бейнелерді тану құрамымен Хэмминг желісімен біріге.	MAXNET нейрондық желісінің жұмыс істеу итерациясының саны алдын ала белгісіз. Ол кіру сигналын қайсысы ең жоғары мәнге ие екендігін анықтайды. Алайда, жұмыс істеу кезінде желі максимал сигналдың мәнін сақтамайды. Кіру сигналының өлшемі артқан кезде синапстардың саны квадратты ұлғайтылады.	Желі жұмысының қарапайымдылығы.
Радиалдық негізделген желі (RadIal	Бейнелерді тану, сыныптау.	Жасырын қабаттағы нейрондардың функциясын құру үшін алдын ала	Оқыту кезеңінің болмауы.

BasIs FunctIon Network)		эталондар, эвристикалар саны белгілі болуы керек. Осы түрдегі желілер салыстырмалы жинақы және оңай оқытылады. Радиалдық базалы желі: бір жасырын қабат, тек жасырын қабат нейрондары ғана сызықты емес активті функцияға ие, кіріс және жасырын қабаттарының синапстық байланыстарының салмақтары бірге тең сияқты ерекшеліктері бар	
Бірқабатты персептрон (Single Layer Perceptron)	Бейнелерді тану, сыныптау.	Қарапайым бөлуші жазықтықтықтар танудың аса қиын емес мәселерін шешуге ғана мүмкіндік береді.	Программалық немесе аппараттық моделдерді іске асыру өте жеңіл. Оқыту алгоритмі оңай және жеңіл.

МАЗМҰНЫ

БІРІНШІ БАСЫЛЫМҒА АЛҒЫ СӨЗ	4
ЕКІШІ БАСЫЛЫМҒА АЛҒЫ СӨЗ	7
I. НЕЙРОНДЫҚ ЖЕЛІЛЕРДІҢ МАТЕМАТИКАЛЫҚ НЕГІЗДЕРІ	8
I. 1. Жиындар теориясының элементтері	8
<i>I. 1.1. Жиындар және оларға амалдар</i>	8
<i>I.1.2. Қатынастар және оларды бейнелеу тәсілдері</i>	15
I.2. Функциялар теориясының элементтері	24
<i>I.2.1. Функцияның анықтамасы, мәні және туындысы</i>	24
<i>I.2.2. Сандық функциялар және олардың қасиеттері</i>	31
<i>I.2.3. Нақты функциялар.</i>	36
II. НЕЙРОНДЫҚ ЖЕЛІЛЕР ТЕОРИЯСЫНЫҢ ЭЛЕМЕНТТЕРІ	42
II.1 Негізгі түсініктер	42
<i>II.1.1 Табиғи нейрондар</i>	42
<i>II.1.2. Табиғи нейрондық желілер</i>	48
<i>II.1.3 Жасанды нейрондар және нейрондық желілер тарихы</i>	52
<i>II.1.4. Жасанды нейронның формалды моделі</i>	60
<i>II.1.5. Жасанды нейрондарды сыныптау</i>	68
<i>II.1.6. Нейрондық желілерді сыныптау</i>	73
II.2. Белсендіру функциясы	80
<i>II.2.1. Белсендіру функциясы туралы жалпы мағлұмат</i>	80
<i>II.2.2. Сызықтық белсендіру функциясы</i>	86
<i>II.2.3. Межелік белсендіру функциясы</i>	94
<i>II.2.4. Модулдық белсендіру функциясы</i>	99
<i>II.2.5 Сигмоидальқ белсендіру функциялары.</i>	103
<i>II.2.6. Ара тәріздес белсендіру функциясы</i>	109
<i>II.2.8. Дәрежелі белсендіру функциясы</i>	117
<i>II.2.9. Квадраттық түбір белсендіру функциясы</i>	121
<i>II.2.10. Экспоненциалды белсендіру функциясы</i>	124
<i>II.2.11. Кіріс мәндерін азайтатын белсендіру функциясы</i>	128
<i>II.2.12. Құлдилау айнымалысы бар белсендіру функциясы</i>	132
<i>II.2.13. Радиалды-базистік белсендіру функциясы</i>	137
<i>II.2.14. Бәсекелесуші белсендіру функциялары</i>	142
II.3. Нейрондардың моделдері	146
<i>II.3.1. Персептрондар</i>	146
<i>II.3.2. Сигмоидальқ нейрондар</i>	156
<i>II.3.3. Паде-нейрондар</i>	164
<i>II.3.4. Квадраттық қосқышы бар нейрондар</i>	168
<i>II.3.5. Бейімделген сызықтық нейрондар</i>	173
<i>II.3.6. Сигма-Пи нейрондар</i>	177
<i>II.3.7. WTA типті нейрондар</i>	180

II.3.9. Гроссберг жұлдыздары	187
II.3.10. Есептеуіші сәйкесті нейрондар	193
II.3.11. Кубтық нейрондар	197
II.3.12 Радиалды нейрондар	201
II.3.13. Стохастикалық нейрондар	206
II.4. Нейрондық желілердің моделі	212
II.4.1. Бір қабатты нейрондық желілер	212
II.4.2 Бір қабатты персептрондар	218
II.4.3 Көп қабатты нейрондық желілер	224
II.4.4 Көп қабатты персептрондар	235
II.4.5. Кохоненнің нейрондық желілері	242
II.4.6 Аралас нейрондық желілер	252
III. НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ОҚЫТУ	259
III.1. Нейрондық желілерді оқытудың парадигмалары	259
III.1.1. Нейрондық желілерді оқытудың негізгі түсініктері	259
III.1.2. Нейрондық желілердің мұғаліммен оқыту парадигмасы	267
III.1.3. Нейрондық желілерді мұғалімсіз оқытудың парадигмасы.	273
III.1.4. Нейрондық желілерді аралас оқыту парадигмасы	281
III.2. Нейрондық желілерді оқыту әдістері	287
III.2. Қателерді кері тарату әдісімен нейрондық желілерді оқыту	287
III.2.3. Коши әдісімен нейрондық желілерді оқыту	304
III.2.4. Хебб әдісімен нейрондық желілерді оқыту	308
III.2.5. Кохонен әдісімен нейрондық желіні оқыту	313
III.2.6. Нейрондық желіні жарыс әдісімен оқыту	317
III.2.7. Тікелей әдіс пен кері әдісті біріктіру	323
III.2.8. Қателерді кері тарату әдісі мен Коши әдісінің комбинациясы	329
III.2.10. Нейрондық желі жұмысының сапасын бағалау	341
III.2.11 Оқыту әдістерін салыстыру	346
III.3. Нақты нейрондық желілерді оқыту	348
III.3.1. Персептронды оқыту	348
III.3.2. Сигмоидалық нейронды оқыту	357
СӨЗ СОҢЫ	362
ДЕРЕКТЕР КӨЗДЕРІ	363
1-ҚОСЫМША. ПЕРСЕПТРОННЫҢ БЕЙНЕЛІЛІГІ	366
2-ҚОСЫМША. ПЕРСЕПТРОН МЫСАЛДАРЫ	371
3-ҚОСЫМША. ЖАСАНДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕР СИПАТТАМАЛАРЫ	378
4-ҚОСЫМША. ЖАСАНДЫ НЕЙРОНДЫҚ ЖЕЛІЛЕРДІҢ ТҮРЛЕРІ	380

НЕЙРОНДЫҚ ЖЕЛІЛЕР (Оқулық)

Оқулық мазмұнына баспа бөлімі жауап бермейді

Таралымы 100 дана
Пішіні 60x84 1/16
Қаріп түрі “Times New Roman”
Тапсырыс №. 15061

Л.Н.Гумилев атындағы Еуразия ұлттық университетінің
типографиясында басылып шығарылды.
010008, Астана қаласы, Қажымұқан, 13/1



Аты-жөні, туылған күні: Алтынбек Әмірұлы (Амирович) ШӘРІПБАЙ (ШАРИПБАЕВ), 16 маусым 1952 жылы.

Білімі. 1974 жылы С.М. Киров атындағы ҚазМУ "Қолданбалы математика" мамандығы бойынша бітірді, ал 1981 жылы М.В. Ломоносов атындағы ММУ аспирантурасын "Есептеу машиналары, жүйелері мен желілерінің математикалық және программалық қамтымы" мамандығы бойынша аяқтады. 1989 жылы М. Тореца атындағы ШТ ММИ «Француз тілі» курсын, ал 1990 жылы Поль Валерий университетінде (Франция) ғылыми тағылымдамадан өтті.

Қазақ, орыс, ағылшын және француз тілдерін біледі.

Ғылыми атағы, дәрежесі. «Информатика, есептеу техникасы және басқару» мамандықтар тобы бойынша физика-математика ғылымдарының кандидаты, техника ғылымдарының докторы және профессор, Халықаралық ақпараттандыру академиясының академигі, ҚР педагогика ғылымдары академиясының академигі, ҚР ғылым, техника және білім саласындағы мемлекеттік сыйлығының жылғы лауреаты.

Жұмыс орны. 1974-1975 жылдары С.М.Киров атындағы ҚазМУ – математик-программалаушы; 1975-1981жылдары М.В.Ломоносов атындағы ММУ – стажер-зерттеуші, аспирант; 1981-1992 жылдары ҚазМУ-де ассистент, доцент, кафедра меңгерушісі; 1992-1997 жылдары ҚР ҰҒА Информатика мен басқару проблемалар институты – зертхана меңгерушісі; 1997-1998 жылдары Қазақ мемлекеттік басқару академиясы – кафедра меңгерушісі; 1998-2000 жылдары Ақпараттық технологиялар орталығы – директор; 2000 жылдан бастап қазіргі уақытқа дейін Л.Н. Гумилев атындағы ЕҰУ – кафедра меңгерушісі, профессор, «Жасанды интеллект» ғылыми-зерттеу институтының директоры.

Ғылыми-практикалық әрекет: ғылыми қызығушылығы – программалау, жасанды интеллект, ақпараттық қауіпсіздік, цифрлық технологиялар, электрондық оқыту. Программалау саласында ол алмастырымдылық және логикалық программалау тілдерінің семантикаларын формалданды, программалық және аппараттық құралдарды верификациялауды автоматтандыру әдістері құрылды. Осы зерттеулер негізінде «Символдық процессор тіліндегі программаларды ішінара верификациялау» кандидаттық диссертациясы және «Есептеу машиналары мен жүйелерінің программалық және аппараттық құралдарын верификациялау» докторлық диссертациясы қорғалды. Алынған ғылыми нәтижелері 1976-1991жылдары КСРО стратегиялық ғылыми орталықтарында енгізілді: «Ғарыштық жүйелерді имитациялау тілінің трансляторы» Ұшу-сынау институтында (Жуковск қ.); «Цифрлық сұлбаларды верификациялау жүйесі», Ғылыми-өндірістік орталығында (Зеленоград қ.); «Көппроцессорлы есептеу кешені үшін параллелді программалау жүйесі» Электрондық техникасының халықаралық ғылыми орталығында (Москва қ.).

Жасанды интеллект (компьютерлік лингвистика) саласында қазақ тілінің математикалық теориясы және қазақ тілінің жазбаша және ауызша сөйлемдерін талдау мен синтездеу әдістері жасалды. Бұл қазақ мәтіндерін грамматикалық талдау мен синтезді және сөйленгенді жазбашаға, ал жазылғанды сөйлеуге айналдыруды автоматтандырып, қазақ тілінде жасанды интеллект жасауға мүмкіндік берді. Нәтижесінде компьютер қазақ тіліндегі сөйлемдердің мағынасын түсініп, адам қойған сұраққа жауап бере алатын, ал өзі қойған сұраққа адам берген жауаптың дұрыстығын бағалай алатын болды (<https://abctv.kz/ru/tv/zhasandy-intellekt>).

Ақпараттық қауіпсіздік және ақпаратты қорғау саласында бейклассикалық (ықтимал, бұлдыр) логика негізінде тәуекелділікті бағалау мен аудит жасау және балама математикалық (автоматтық, кванттық және нейрожелілік) моделдер негізінде сенімді криптожүйелер жасауға мүмкіндік беретін әдістер даярланды.

Цифрлық технологиялар саласында (Қазақ әлпбиін кодтау, Автоматтандырылған жүйелерді жасауға техникалық тапсырмалар, Программалық қамтымды сынақтан өткізуді жоспарлау және басқа) мемлекеттік стандарттар даярланды және тапсырушы мекемелерде ендірілген әртүрлі атоматтандырылған жүйелер (Құнды қағаздарды автоматтандырылған тіркеу жүйесі, Құнды қағаздарды автоматтандырылған депозиттер жүйесі, Коммуналдық төлемдерді есептеудің автоматтандырылған жүйесі, Депутаттарды сайлаудың автоматтандырылған жүйесі, Есеп комитетінің автоматтандырылған ақпараттық базасы және т.б.) даярланды.

Электрондық оқыту саласында «Электрондық оқу басылымдарының генераторы», «Мемлекеттік қызметкерлерді қазақ тіліне қашықтан оқыту жүйесі», «Қазақ тілінде сұрақ-жауап интеллектуалды жүйесі», «Интеллектуалды электрондық университет», «Қазақ тілінде білімді беру, бақылау және бағалаудың интеллектуалды жүйесі» және 40-тан астам электрондық оқу басылымы даярланды.

Ғылыми-әдістемелік әрекет: 500-дей ғылыми еңбектер, 11 оқулық (Информатика қазақ және орыс тілдерінде, Тілдер мен автоматтар теориясы қазақ және орыс тілдерінде 2-басылым, Нейрондық желілер қазақ және орыс тілдерінде 2-басылым, Ақпараттық қауіпсіздік және ақпаратты қорғау 3-бөлім) және 13 оқу құралдарын (Языки символьной обработки, Технология программирования, Математика для компьютерных наук, Mathematics for Computer Science, Реляциялық деректер базасын құру мен әкімшіліктеу және т.б.) шығарды, 3 монография (Доказательство правильности программных и аппаратных средств компьютеров, Құнды қағаздар анықталығы, Қазақ жазуын латын әлібиіне көшіру проблемалары) және 5 терминологиялық (Қазақша-орысша, орысша-қазақша информатика және есептеу техника бойынша терминологиялық сөздік 2-басылым, Қазақша-орысша, орысша-қазақша ақпараттық және коммуникациялық технологиялары бойынша терминологиялық сөздік, Қазақ жазуын латын әлібиіне көшіруде қажет орфографиялық сөздік) және 4 түсіндірме сөздіктер (Ғылыми-салалық информатика және есептеу техника бойынша түсіндірме сөздік, Ағылшын, орыс және қазақ тілдерінде информатика бойынша түсіндірме сөздік, Ағылшын, орыс және қазақ тілдерінде автоматтанрылған жүйелер бойынша терминдер мен анықтамалар, Ағылшын, орыс және қазақ тілдерінде программалық құжаттамалардың бірыңғай жүйесі бойынша терминдер мен анықтамалар) шығарды, сонымен қатар 30-дан астам авторлық куәліктер алынған, 9 МЖМББС даярланды.

Ғылыми кадрлерді даярлау: Информатика, есептеу техникасы және басқару мамандықтар тобы бойынша 5 ғылым докторын, 8 ғылым кандидатын және Информатика бойынша 8 PhD дайындады.

Халықаралық ынтымақтастық. Шетелдердің ғылыми орталықтарының шақыруы бойынша дәріс оқу мен ғылыми баяндама жасау үшін ол келесі елдерде болды: Болгария –1989 ж.; Франция –1990 ж., 1996 ж., 2015 ж.; ҚХР –1993 ж., 2010 ж., 2014 ж.; Германия –1998 ж., 2015 ж.; Түркия –1999 ж., 2008 ж., 2013 ж., 2015 ж.; Тайланд – 2003 ж., 2015 ж. ; Израиль – 2006 ж.; Ұлыбритания – 2006 ж.; Мальта – 2007 ж., Жапония –2012 ж.; Корея – 2013 ж.; Италия – 2013 ж., 2018 ж.; Испания – 2014 ж., 2016 ж., 2018 ж., ОАЕ – 2016 ж., Австрия – 2016 ж.; Малайзия – 2017 ж., Португалия – 2017 ж.;

Марапаттаулар мен гранттар: 2001 ж. – ҚР ғылым, техника және білім беру саласындағы Мемлекеттік сыйлығының лауреаты; 2003 ж. – ҚР ғылымының дамуына қосқан үлесі үшін медалі; 2005, 2010, 2021 ж. ҚР БҒМ «ЖОО үздік оқытушысы» гранттары; 2014 ж. – Кеден одағын дамытқаны үшін медалі; 2015 ж. – «ҚР Конституциясына 20 жыл» Юбилейлік медалі;