

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ  
"Қ. И. СӘТБАЕВ атындағы ҚАЗАҚ ҰЛТТЫҚ ТЕХНИКАЛЫҚ ЗЕРТТЕУ  
УНИВЕРСИТЕТІ" КЕАҚ

Автоматика және ақпараттық технология институты  
Жоғары математика және моделдеу кафедрасы

**6B06103 « Математикалық және компьютерлік модельдеу»**

мамандығының студенті Жеңіснұр Нәзіранын дипломдық жұмысына берілген

**СЫН ПІКІР**

Мамандығы: 6B06103 « Математикалық және компьютерлік модельдеу»  
Дипломдық жұмыс тақырыбы: Қателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана отырып функцияларды жуықтау.

Орындалды:

а) графикалық бөлімі \_\_\_\_\_ парақ

ә) түсіндірме жазба \_\_\_\_\_ бет

б) математикалық есептерді шешу, модельдерді қорытындылау \_\_\_\_\_ бет

**ЖҰМЫСҚА ҚАТЫСТЫ ЕСКЕРТУЛЕР**

Кері таралу алгоритмі көп қабатты тікелей нейрондық желілерді оқыту әдістерінің бірі болып табылады.

Бұл дипломдық жұмыста студент моделдеу әдістерімен және технологияларымен өз бетінше танысып және нақты әдебиеттерді пайдалана отырып кері таралу әдісін пайдаланған.

Жұмыстың нәтижесінде мыналар зерттелген: нейрондық желілер туралы негізгі түсініктер, жасанды нейрондық желінің әртүрлі архитектурасы, кері таралу әдісімен нейрондық желілерді оқыту.

Дипломдық жұмыс – мақсатқа сай және жоғары техникалық дәрежеде орындалған. Жұмысқа қатысты ескертулер ретінде техникалық әдебиеттерді аударуда кеткен кейбір қателіктерді келтіруге болады.

**Жұмысқа берілген баға**

Дипломдық жұмысқа – 98% «өте жақсы» деген баға қоямын және Жеңіснұр Нәзира – 6B06103 «Математикалық және компьютерлік модельдеу» мамандығы бойынша бакалавр біліктілігіне лайық деп санаймын.

**Сын-пікір беруші**

ҚР ҒЖБМ ҒК Ақпараттық және есептеуіш

технологиялар институтының аға ғылыми

қызметкері, PhD

Алғазы К.Т.

« 30 \_\_\_\_\_ 2024 ж.



## ҒЫЛЫМИ ЖЕТЕКШІНІҢ

### ПІКІРІ

Дипломдық жұмыс  
(жұмыс түрінің атауы)

Жеңіснұр Назира  
(білім алушының Т.А.Ә.)

6В06103 - Математикалық және компьютерлік модельдеу  
(мамандық атауы мен шифрі)

Тақырыбы: Қателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана отырып функцияларды жуықтау

Дипломдық жұмыс авторы өзінің еңбекқорлығымен, тапсырылған жұмысты уақытымен орындауымен ерекшеленеді.

Дипломдық жұмыс 6В06103 «Математикалық және компьютерлік модельдеу» мамандығына қойылатын талапқа сай орындалған, технологиялық аппаратты математикалық моделдеу үшін қолдану теориясы кеңінен қолданылған.

Диплом жұмысы күрделі тақырыптың бір өзекті мәселесін толығымен қамтып, аша алған. Көркемдеу талаптарына сәйкес орындалған. Жұмыс авторы толығымен 98% жоғары бағалануға лайықты.



Ғылыми жетекші:

Жеңіснұр Назира (жұмыс түрінің атауы)  
кафедрасының аға оқытушысы

Лукпанова Л.Х.

«20» 05 2024 ж.



**Университеттің жүйе администраторы мен Академиялық мәселелер департаменті  
директорының ұқсастық есебіне талдау хаттамасы**

Жүйе администраторы мен Академиялық мәселелер департаментінің директоры көрсетілген еңбекке қатысты дайындалған Плагиаттың алдын алу және анықтау жүйесінің толық ұқсастық есебімен танысқанын мәлімдейді:

**Автор: Жеңіснұр Нәзира**

**Тақырыбы: Кателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана отырып функцияларды жуықтау**

**Жетекшісі: Ляззат Лукпанова**

**1-ұқсастық коэффициенті (30): 13.4**

**2-ұқсастық коэффициенті (5): 4.2**

**Дәйексөз (35): 0.4**

**Әріптерді ауыстыру: 2**

**Аралықтар: 0**

**Шағын кеңістіктер: 0**

**Ақ белгілер: 0**

**Ұқсастық есебін талдай отырып, Жүйе администраторы мен Академиялық мәселелер департаментінің директоры келесі шешімдерді мәлімдейді :**

Ғылыми еңбекте табылған ұқсастықтар плагиат болып есептелмейді. Осыған байланысты жұмыс өз бетінше жазылған болып санала отырып, қорғауға жіберіледі.

Осы жұмыстағы ұқсастықтар плагиат болып есептелмейді, бірақ олардың шамадан тыс көптігі еңбектің құндылығына және автордың ғылыми жұмысты өзі жазғанына қатысты күмән тудырады. Осыған байланысты ұқсастықтарды шектеу мақсатында жұмыс қайта өңдеуге жіберілсін.

Еңбекте анықталған ұқсастықтар жосықсыз және плагиаттың белгілері болып саналады немесе мәтіндері қасақана бұрмаланып плагиат белгілері жасырылған. Осыған байланысты жұмыс қорғауға жіберілмейді.

**Негіздеме:**

**2024-05-24**

*Күні*



*Кафедра меңгерушісі*

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

«Қ.И.Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті»  
коммерциялық емес акционерлік қоғамы

Автоматика және ақпараттық технология институты  
Жоғары математика және модельдеу кафедрасы

Жеңіснұр Нәзира

Қателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана  
отырып функцияларды жуықтау

**ДИПЛОМДЫҚ ЖҰМЫС**

6B06103 - Математикалық және компьютерлік модельдеу

Алматы 2024

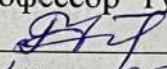


ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

«Қ.И.Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті» коммерциялық емес  
акционерлік қоғамы

Автоматика және ақпараттық технология институты  
Жоғары математика және модельдеу кафедрасы

ҚОРҒАУҒА ЖІБЕРІЛДІ  
Математика кафедрасының  
меңгерушісі физика-  
математика ғылымдарының  
кандидаты, қауымдастырылған  
профессор Тулешева Г.А.

  
«31» 05 2024ж

ДИПЛОМДЫҚ ЖҰМЫС

Тақырыбы: «Қателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана отырып  
функцияларды жуықтау»

6B06103 - Математикалық және компьютерлік модельдеу

Орындаған

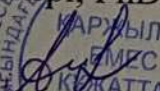
Жеңіснұр Н.

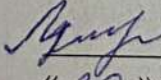
Сын-пікір беруші

Ғылыми жетекші

ҚР ҒЖБМ ҒК Ақпараттық және есептеуіш  
технологиялар институтының аға ғылыми  
қызметкері, PhD

Жоғары математика және  
моделдеу» кафедрасының  
аға оқытушысы

  
Алғазы К.Т.  
«30» мамыр 2024ж.

  
Лукпанова Л.Х.  
«30» мамыр 2024 ж.

Алматы 2024



ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

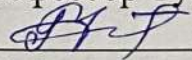
«Қ.И.Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті»  
коммерциялық емес акционерлік қоғамы

Автоматика және ақпараттық технология институты  
Жоғары математика және модельдеу кафедрасы

6B06103 – Математикалық және компьютерлік модельдеу

**БЕКІТЕМІН**

Математика кафедрасының  
меңгерушісі физика- математика  
ғылымдарының, қауымдастырылған  
профессор Тулешева Г.А.

  
« 31 » 05 2024ж

**Дипломдық жұмыс орындауға  
ТАПСЫРМА**

Білім алушы: Жеңіснұр Нәзира

Тақырыбы: Қателерді кері тарату әдісімен жасанды нейрондық желілерді пайдалана отырып функцияларды жуықтау

Университет ректорының 2023 жылғы «4» желтоқсандағы № 548-П/Ө бұйрығымен бекітілген.

Аяқталған жұмысты тапсыру мерзімі: 2024 жылғы «3-6» маусым.

Дипломдық жобаның бастапқы деректері:

Дипломдық жұмыста әзірленуге жататын мәселелердің тізбесі немесе дипломдық жұмыстың қысқаша мазмұны:

А Жұмысқа жалпы шолу

Б Жұмыстың құрылым бөлімі

В Алынған мәліметтерге негізделген қорытынды

Ұсынылатын негізгі әдебиеттер 7 кітап.





Дипломдық жұмысты дайындау  
КЕСТЕСІ

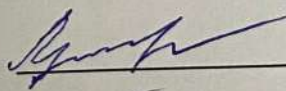
Бөлімдердің атаулары, әзірленетін мәселелердің тізбесі	Ғылыми жетекшіге ұсыну мерзімдері	Ескерту
Тақырыпқа байланысты арнайы әдибеттерді қарастыру	02.02.2024	орындалды
Дипломдық жұмыстың жоспарын құру	09.04.2023	орындалды
Негізгі бөлімді қарастыру	15.04.2024	орындалды
Дипломдық жұмысты қорытындылау	19.05.2024	орындалды

Қолтаңбалар

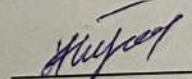
Аяқталған дипломдық жобаға консультанттар мен нормоконтролер оларға қатысты жұмыс бөлімдерін көрсете отырып

Бөлімдердің атаулары	Кеңесшілер Т.А.Ә. (мұғ. дәрежесі, атағы)	Қол қойылған Күні	Қолы
Негізгі бөлім	Лукпанова Л.Х. техника ғылымдарының магистрі, аға оқытушы	30.05.2024	
Норма бақылаушы	Шатманов Ж.Ж. физ.-мат. ғылымдарының кандидаты, қауымдастырыл ған профессор	29.05.2024	

Ғылыми жетекшісі

 Л.Х. Лукпанова

Білім алушы тапсырманы орындауға алды

 Н.Жеңіснұр

Күні

«30» мамыр 2024ж

## **АНДАТПА**

Дипломдық жұмыста кері таралу әдісін қолданып, жасанды нейрондық желіні қолдану арқылы функцияны жуықтаймыз.

Дипломдық жұмыс барысында келесі нәтижелерге қол жеткіземіз:

- Нейрондық желілер туралы негізгі түсінік.
- Жасанды нейрондық желісінің әртүрлі архитектуралары.
- Кері таралу әдісімен нейрондық желілерді оқыту.
- Нейрондық желі арқылы функцияларды жуықтау.

## **АННОТАЦИЯ**

В дипломной работе мы аппроксимируем функцию, используя метод обратного распространения ошибки и искусственную нейронную сеть.

В ходе дипломной работы мы добьемся следующих результатов:

- Базовое понимание нейронных сетей.
- Различные архитектуры искусственных нейронных сетей.
- Обучение нейронных сетей методом обратного распространения ошибки.
- Аппроксимация функций нейронной сетью.

## **ABSTRACT**

In our thesis we approximate the function using backpropagation and an artificial neural network.

During the thesis we will achieve the following results:

- Basic understanding of neural networks.
- Various artificial neural network architectures.
- Training of neural networks using backpropagation method.
- Approximation of functions by a neural network.



## МАЗМҰНЫ

	КІРІСПЕ	7
1	Нейрондық желілер	8
1.1	Жасанды нейрондық желінің анықтамасы	8
1.2	Нейрондық желінің архитектурасы. Көпқабатты перцептрон (MLP)	11
2	Нейрондық желілерді оқыту	14
2.1	Кері таралу алгоритмі	14
2.2	Кері даярлау және жалпылау	21
3	Функцияның жуықтауы	25
3.1	Нейрондық желілер арқылы функцияларды даярлау	25
3.2	Matlab жүйесінде көпқабатты желіні енгізу	30
3.3	Жұмыстың орындалу тізімі	36
	Қорытынды	37
	Пайдаланылған әдебиеттер тізі	38

## КІРІСПЕ

Нейрондық желілерді зерттеу 20 ғасырдың 40-шы жылдарында басталды. Ең алғашқы нейрондық желілерді 1943 жылы Маккалох пен Питтс зерттеген. Кейінірек олар зақым келген, көрінбейтін кесіндерді тану үшін желілік парадигмаларды зерттеді.

Нейрондық желілерге қойылатын негізгі міндеттердің бірі ол, қандайда бір үлгіні тану. Олар қандайда бір үлгіні классификациялаудан тұрады, яғни желіні қандайда бір классқа тағайындаудан тұрады. Бастапқы кезде нейрондық желіге анықтамалық, базалық кескіндер беріледі. Кейінірек оған қосымша өзге кескіндерді жалғайды. Осылайша нейрондық желі кескіндерге сәйкестік іздеп белгілі алгоритмді пайдаланым салыстыруға тырысады. Жалпы айтқанда нейрондық желілер кескінді кластерлеу әдісін орындайды деп айта аламыз. Кластерлік талдауды бағалы қағаздар нарығын зерттеушілер пайдаланатындықтан, нейрондық желілерді акциялардың бағасын болжау үшін де пайдалануға болады және бұл өзекті мәселе болып табылады.

Нейрондық желілерді сот сараптамасында немесе құқық қорғау органдарының қылмыскерлерді ұстаудағы жұмысын жеңілдету үшін де қолдануға болады, өйткені бет-әлпетті танытын бағдарламалар жасалып қойған.

Бұл жұмыста кері таралу әдісін қолданып, жасанды нейрондық желіні қолдану арқылы функцияны жуықтаймыз.



# 1 Нейрондық желілер

## 1.1 Жасанды нейрондық желінің анықтамасы

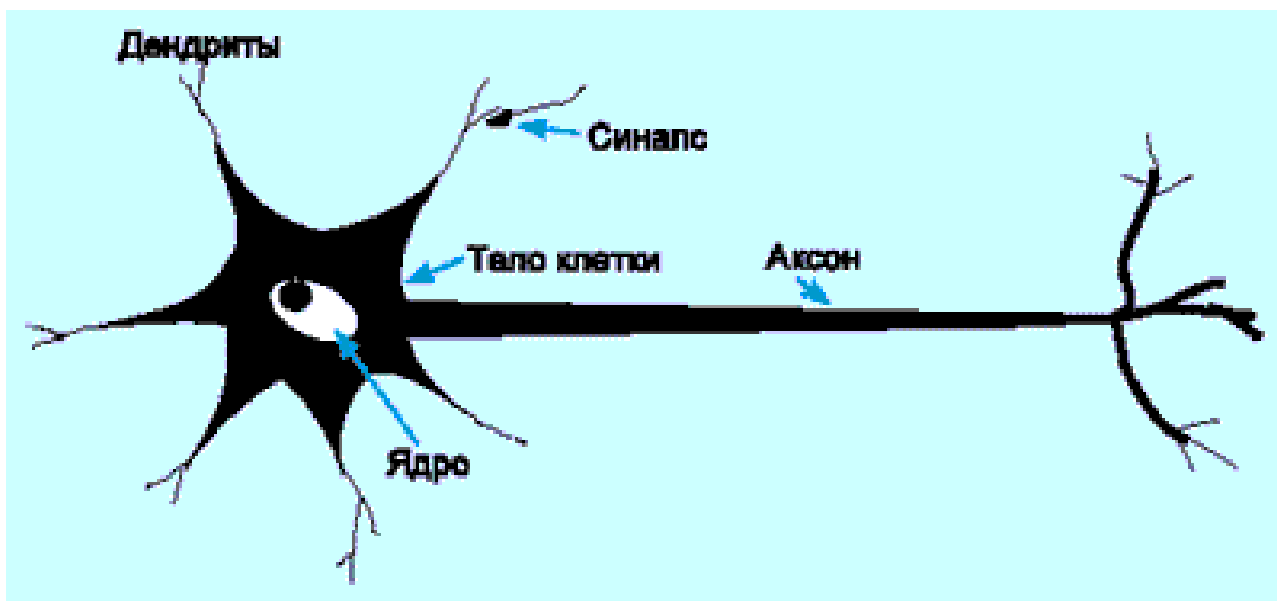
Нейрондық желілер жасанды интеллектті зертеулер барысында пайда болған, толығырақ айтқанда мидің төменгі деңгейлі құрылымын модельдеу арқылы биологиялық жүйке жүйелерінің қателерін анықтап, оларды түзеу болып табылады. 60-80 жылдар аралығындағы жасанды интеллекттің негізгі бағыты сараптамалық жүйелер болған. Мұндай жүйелер ойлау үдерісін жоғары деңгейде модельдеуге негізделген (атап айтқанда, біздің ойлау үрдісіміз таңбалармен айла-шарғы жасау негізінде құрылған деген идеяға негізделген). Көп ұзамай мұндай жүйелер кейбір салаларда пайдалы болғанымен, адам интеллектінің кейбір негізгі аспектілерін қамтымайтыны белгілі болды. Осыған көзқарастардың бірі ми құрылымын қайта жасай алмауы болған. Жасанды интеллектті құру үшін, ұқсас архитектурасы бар жүйені құру қажет.

Қызықты деректердің бірі, ми өте үлкен саннан тұрады (шамамен 10 000 000 000 нейрондардан). Нейронда – электрохимиялық сигналдарды беруге қабілетті арнайы жасушалар. Нейронның ақпарат кірісінің тармақталған құрылымының (яғни дендриттер), ядро және тармақталған шығыс (аксондар) болады. Аксондардың жасушалары синапстар арқылы басқа жасушалардың дендриттерімен байланысқа енеді. Екі жасуша байланысқа енгеннен кейін, нейрон аксон бойымен электрохимиялық сигнал жібереді. Синапстардың арқасында бұл электрохимиялық сигналдар басқа нейрондарға жетіп, белсеніске түседі. Нейрон тек оның ядросына дендриттерден келетін сигналдар арқасында берілген шегтен артық сигнал түскен кезде ғана белсендіріледі.

Нейрон қабылдаған сигналдардың қарқындылығы синапстардың белсенділігіне қатты байланысты. Нейрожүйелердің ең беделді зерттеушілерінің бірі Дональд Хебб оқыту ең алдымен синаптикалық байланыстардың «күшіндегі» өзгерістерден тұрады деп тұжырымдаған. Мысалға, Павловтың классикалық тәжірибесінде итін тамақтандырар алдын әр жолы қоңырауды соғатын және ит қоңыраудың соғуы тамақпен байланыстыруды тез үйренді. Осылайша ми қыртысының естуге жауапты аймақтары мен сілекей бездері арасындағы синаптикалық байланыстар күшейіп, қоңырау есту барысында иттің сілекейі шыға бастайтын.

Осылайша, өте қарапайым элементтердің өте көп санынан құрастырылғанын хабарлаймыз. Ал ми бұдан күрделі мәселелерді шешуге қабілетті. Әрине, мында біз мидің қиын әрі күрделі аспектілеріне тоқталған жоқпыз, алайда қызықтыратыны,

жасанды нейрондық желілер жоғарыда келтірілген модельден әлде қайда күрделі емес модельді пайдаланып тамаша нәтижелерге қол жеткізе алады.



1 - сурет – Биологиялық нейрон

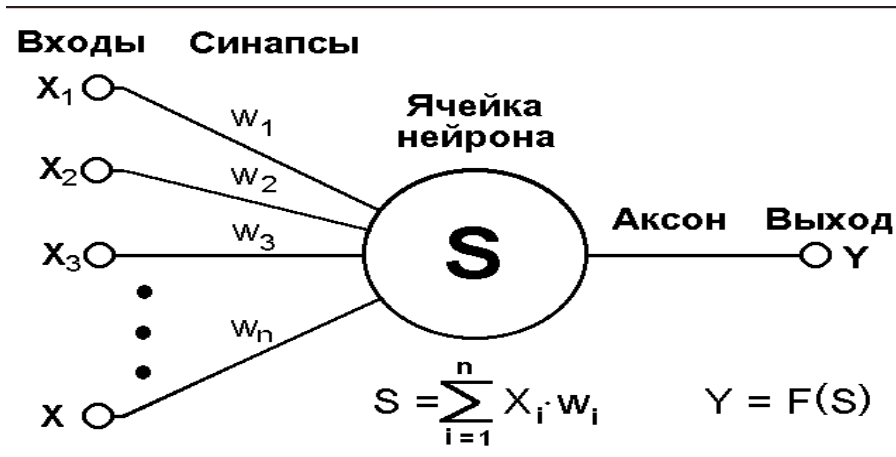
Ал ендігі кезекте жасанды нейрондық желіге тоқталсақ. Жасанды нейрондық желі – бұл тірі организмнің жүйке жасушаларының, биологиялық нейрондық желілерді ұйымдастырып, оның жұмыс жасауына негізделген математикалық модель, сонымен қатар ақпараттық және бағдарламалық іске асыру. Бұл тұжырым мида болып жатқан процестерді зерттеп, оны модельдеуге тырысқанда пайда болады. Ең алғашқы мұндай нейрондық желілерді Маккалоч пен Питстің ұйымдастыруынан болды. Кейіннен оқыту алгоритімдері пайда болғаннан кейін алынған модельдер практикалық мақсаттарда қолданыла бастады. Мысалға, математикалық есептерді болжауда, үлгілерді анықтауда, басқару есептерінде және т.с.с.

Әрбір нейрондық желілердің негізін қарапайым нейрондар құрайды. Көп жағдайда ми нейрондарын имитациялайтын бір типті элементтер құрайды. Әрбір синапстардың нейрондық топтары бар, яғни басқа нейрондардың шығыстары мен біріктірілген бір бағытты кіріс қосылымдары. Әрбір синапс түрі синаптикалық байланыс өлшемімен немесе оның салмағымен  $w_i$  сипатталады.

Әрбір нейронның ағымдағы күйі болады, ол әдетте оның кірістерінің өлшенген сомасы ретінде анықталады:



$$S = \sum_{i=1}^n X_i \cdot W_i \quad (1.1)$$

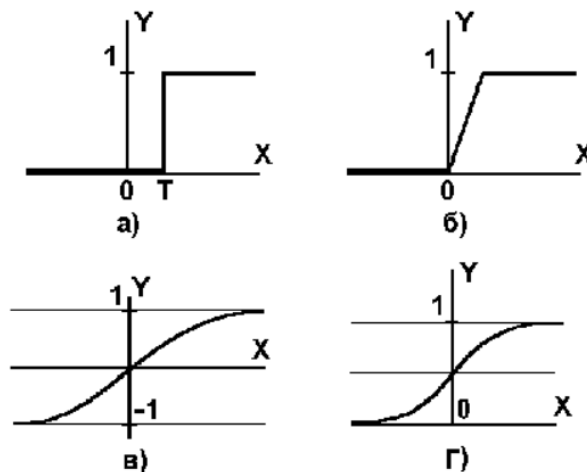


2 – сурет – Жасанды нейрон

Әрбір нейронның аксоны болады, олар берілген нейронның шығыс байланысы, яғни оның сигналы келесі синапстардың нейрондарына барады. Нейронның шығуы оның күйінің функциясы болып табылады:

$$y = f(s)$$

Сызықты емес  $f$  функциясы активтендіру деп аталады және әртүрлі формада болуы мүмкін. Толығырақ 3-суреттен байқай аламыз.



3 – сурет – Активтендіру функцияларының түрлері

- (3а) табалдырық (пороговый);
- (3б) бөлікті – сызықтық (кусочно – линейный);
- (3в, 3г) сигма тәріздес (сигмоид).

Сигналдарды өңдеу паралельді түрде жүргізіледі. Осынын арқасында көптеген нейрондарды біріктіру және әртүрлі қабаттардың нейрондарын, сонымен қатар кейбір конфигурацияларда бір қабаттың нейрондарын өзара байланыстыру арқылы қол жеткізіледі және барлық нейрондардың өзара әрекеттесуі қабат – қабат өңделеді.

Нейрондық желі құрылымын таңдау тапсырманың сипаттамалары мен күрделілігіне сәйкес жүзеге асырылады. Кейбір сыныптарға оңтайлы базалық конфигурациялар бар. Егер мәселені белгілі сыныптардың кез келгеніне дейін төмендету мүмкін болмаса, әзірлеушіге жаңа конфигурацияны синтездеу мәселесін шешуге тура келеді. Жасанды нейрондық желіні синтездеу мәселесі тапсырмаға өте тәуелді, жалпы егжей-тегжейлі ұсыныстар беру қиын. Көп жағдайда жасанды нейрондық желінің оңтайлы нұсқасы эксперименталды түрде алынады.

## **1.2 Нейрондық желінің архитектурасы. Көпқабатты перцептрон (MLP)**

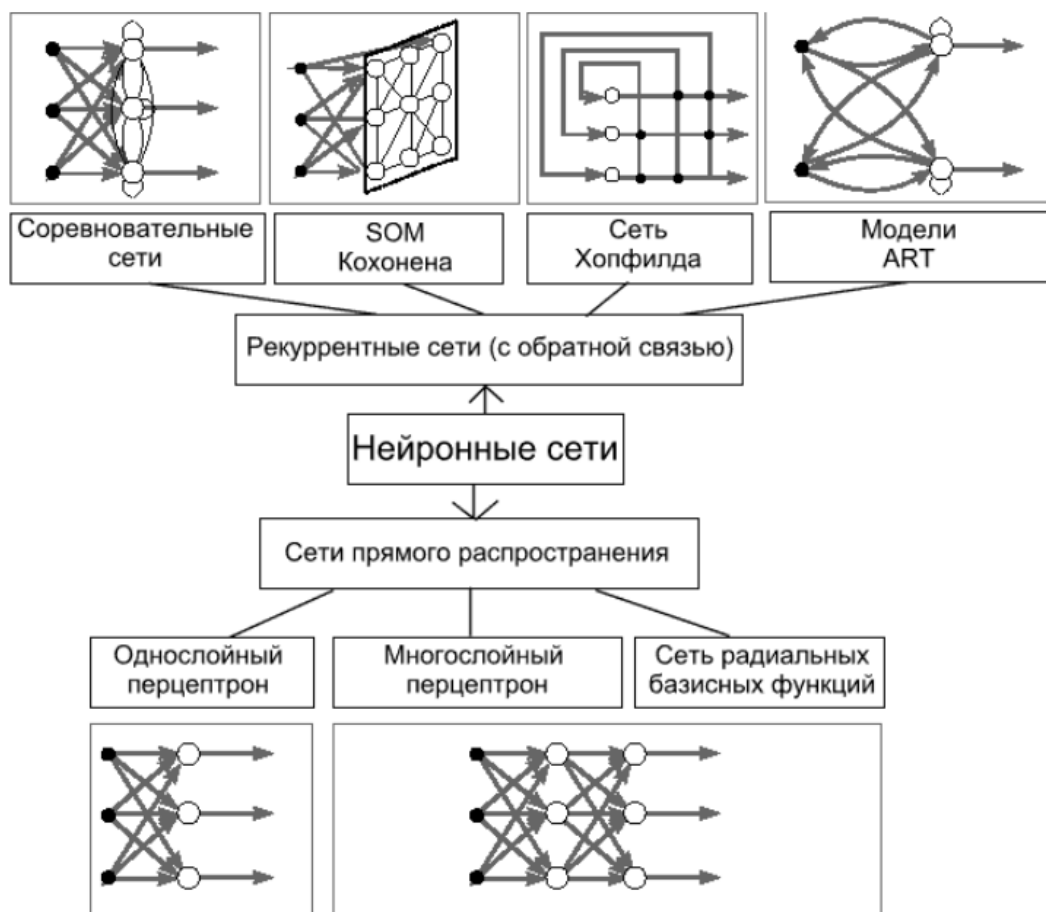
Жасанды нейрондық желіні жасанды нейрондар түйіндер болатын салмақты байланыстары бар бағытталған график ретінде қарастыруға болады. Байланыстардың архитектурасы бойынша жасанды нейрондық желілерді екі класқа топтастыруға болады (4-сурет): графиктерінде ілмектері жоқ алға жіберілетін желілер және қайталанатын желілер немесе оларды кері байланысы бар желілер деп атауға келеді.

Алға қарай бағытталған желілер статикалық болып еледі, себебі берілген кіріс үшін олар желінің алдыңғы күйінен тәуелсіз шығыс мәндерінің бір жинағын шығарады.

Қайталанатын желілер динамикалық болып табылады, өйткені кері байланыстың арқасында оларда нейрондардың кірістері модификацияланады, бұл желі күйінің өзгеруіне әкеледі.

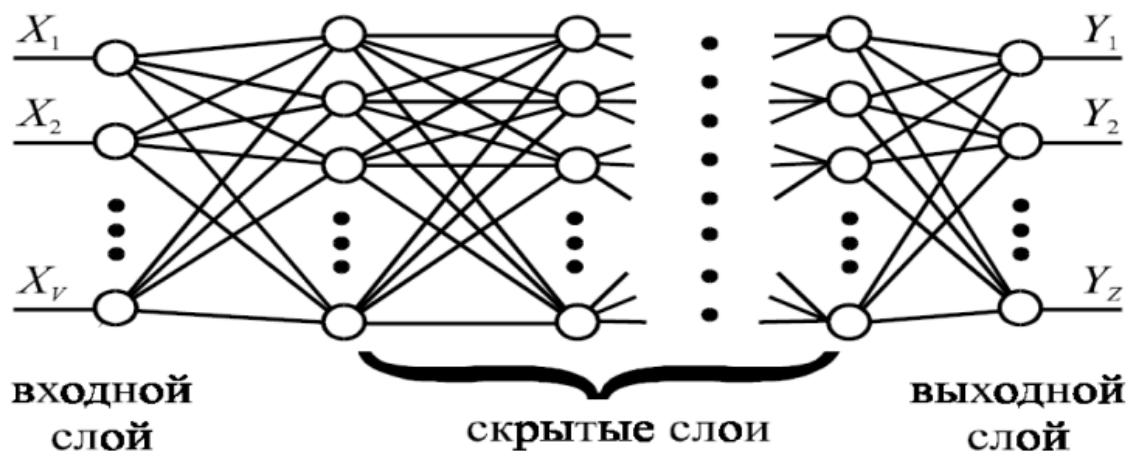
4 – суретте әрбір класстың типтік желілері көрсетілген.





4 - сурет – жасанды нейрондық желі қосылымдарының архитектурасы

Желілердің ең танымал класы – көпқабатты перцептрондар (5-сурет) деп аталатын алға жіберу желілері.



5 – сурет – Көпқабатты перцептрон

Көп қабатты перцептрон 3 немесе одан да көп қабаттан тұрады. Қабаттағы әрбір түйін келесі қабаттағы әрбір түйінге қосылып, желіні толығымен байланыстырады. Бұл архитектура сөзді тану және машиналық аударма тапсырмаларында қолданылады.

Бұл нейрондық желі архитектурасын 1986 жылы McClelland пен Rumelhart ұсынған және нейрондық желілер бойынша барлық дерлік оқулықтарда егжей-тегжейлі қарастырылған. Желінің әрбір элементі термин ретінде түзетілген кірістерінің өлшенген сомасын құрастырады, содан кейін осы элементтің шығыс мәнін шығару үшін осы белсендіру мәнін тасымалдау функциясы арқылы өткізеді. Элементтер сигналды тікелей жіберу арқылы қабат – қабат топологиясында ұйымдастырылған. Мұндай желіні енгізу-шығару моделі ретінде оңай түсіндіруге болады, онда салмақтар мен шекті мәндер модельдің бос параметрлері болып табылады. Мұндай желі кез келген дерлік күрделілік дәрежесіндегі функцияны модельдей алады, әр қабаттағы қабаттар мен элементтердің саны функцияның күрделілігін анықтайды.

## 2 Нейрондық желілерді оқыту

### 2.1 Кері таралу алгоритмі

Жасанды нейрондық желі контекстінде оқу процесін белгілі бір тапсырманы тиімді орындау үшін желі архитектурасы мен қосылым салмақтарын баптау ретінде қарастыруға болады. Әдетте, нейрондық желі қол жетімді жаттығулар жиынтығы негізінде қосылым салмақтарын реттеуі керек. Желінің өнімділігі салмақтар қайталанатын ретпен реттелетіндіктен жақсарады. Желілердің мысалдардан үйрену мүмкіндігі оларды сарапшылар тұжырымдаған жұмыс ережелерінің белгілі бір жүйесін ұстанатын жүйелермен салыстырғанда тартымды етеді.

Оқыту процесін жобалау үшін, ең алдымен, нейрондық желі жұмыс істейтін сыртқы ортаның моделі болуы керек – желіге қолжетімді ақпаратты білу. Бұл модель оқыту парадигмасын анықтайды. Екіншіден, желінің салмақ параметрлерін қалай өзгерту керектігін түсіну керек, яғни баптау процесін қандай оқу ережелері реттейтіндігін. Оқу алгоритмі салмақтарды реттеу үшін оқу ережелерін қолданатын процедураны білдіреді.

Оқытудың үш парадигмасы бар: «мұғаліммен», «мұғалімсіз» және аралас. Бірінші жағдайда нейрондық желіде әрбір енгізу мысалы үшін дұрыс жауаптар (желі шығыстары) болады. Салмақ желі белгілі дұрыс жауаптарға мүмкіндігінше жақын жауаптар беретіндей етіп реттеледі. Бақыланатын оқытудың жетілдірілген нұсқасы нейрондық желі шығысының дұрыстығына тек сыни баға беру белгілі, бірақ дұрыс шығыс мәндерінің өздері емес деп болжайды. Мұғалімсіз оқыту оқу жиынының әрбір мысалына дұрыс жауаптарды білуді талап етпейді. Бұл ішкі деректер құрылымын немесе деректер жүйесіндегі үлгілер арасындағы корреляцияны ашып, үлгілерді санаттауға мүмкіндік береді. Ал аралас оқытуда салмақтардың бір бөлігі бақыланатын оқыту арқылы анықталады, ал қалған бөлігі өздігінен білім алу арқылы алынады.

Оқыту теориясы мысалдармен оқытумен байланысты үш іргелі қасиеттерді қарастырады: сыйымдылық, үлгі күрделілігі және есептеу күрделілігі. Сыйымдылық желі қанша үлгіні есте сақтай алатынын және онда қандай функциялар мен шешім шекараларын құра алатынын білдіреді. Үлгі күрделілігі желіні жалпылау мүмкіндігіне жету үшін қажетті оқу мысалдарының санын анықтайды. Тым аз мысалдар желіні «шамадан тыс жаттықтыруға» әкелуі мүмкін, мұнда ол оқу жиынындағы мысалдарда жақсы жұмыс істейді, бірақ бірдей статистикалық үлестірімге жататын сынақ мысалдарында нашар. Негізгі оқыту ережелері 4 түрімен танымал: Больцман машинасы, Хебб ережесі, қателерді түзету және бәсекелестік оқыту.

Қатені түзету ережесі. Бақыланатын оқытуда әрбір енгізу мысалына қажетті нәтиже беріледі  $d$ . Нақты желі шығысы  $y$  қалағанымен сәйкес келмеуі мүмкін. Жаттығу кезінде қателерді түзету принципі қателіктің біртіндеп төмендеуін қамтамасыз ете отырып, салмақтарды өзгерту үшін  $(d-y)$  сигналын пайдалану болып табылады. Оқыту перцептрон қателескенде ғана жүзеге асады. Бұл оқыту алгоритмінің әртүрлі модификациялары белгілі. Қателерді түзетуге негізделген оқыту көпқабатты желіге арналған 2.1 бөлімінде толығырақ сипатталған.

Больцман жаттығулары. Бұл ақпараттық теориялық және термодинамикалық принциптерден туындайтын стохастикалық оқыту ережесі. Больцман оқыту негізінде құрылған нейрондық желі Больцман машинасы деп аталады.

Больцман машинасында барлық нейрондар екілік сигналдармен жұмыс істейтін қайталанатын құрылымдармен ұсынылған. Бұл олардың қосулы (+1 мәніне сәйкес) немесе өшірулі (1 мәніне сәйкес) күйінде болуы мүмкін екенін білдіреді. Мұндай машина  $E$  энергетикалық функциясымен сипатталады, оның мәні осы машинаны құрайтын жеке нейрондардың нақты күйлерімен анықталады. Оны келесідей формуламен байқасақ болады:

$$E = -\frac{1}{2} \sum_j \sum_{k(j \neq k)} w_{kj} x_k x_j \quad (2.1)$$

мұндағы  $x_j$  –  $j$  нейронының күйі,  $w_{kj}$  –  $j$  және  $k$  нейрондары арасындағы байланыстың синаптикалық салмағы.  $j=k$  шарты бұл желіде нейрондардың өздерімен кері байланыстың жоқтығын көрсетеді.

Больцман машинасының нейрондарын екі функционалдық топқа бөлуге болады: көрінетін және жасырын. Көрінетін нейрондар желі мен оның жұмыс ортасы арасындағы интерфейсті жүзеге асырса, ал жасырын нейрондар сыртқы ортаға байланысты жұмыс істейді. Больцман машинасының басты мақсаты көрінетін нейрондардың күйлері ықтималдықтың қажетті таралуын қанағаттандыратындай салмақтарды реттеу болып табылады. Больцмандық оқытуды қатені түзетудің ерекше жағдайы ретінде қарастыруға болады, онда қателік екі режимдегі күй корреляциясының алшақтығы ретінде түсініледі.

Хебб әдісімен оқыту. Хеббтің оқыту үлгісі барлық оқытудың ішіндегі ең алғашқысы және ең белгілісі делінеді. Бұл әдіс екі ережеге бағытталған:

- 1 Егерде синапстың екі жағындағы екі нейрон бір уақытта активтендірілсе, онда байланыстың күші арта түседі.



- 2 Ал егерде синапстың екі жанындағы нейрондары асинхронды түрде активтендірілсе, ол синапс әлсірей түседі, тіпті өлуіде мүмкін. Осылайша қызмет ететін синапс Хебб синапсы деп аталады.

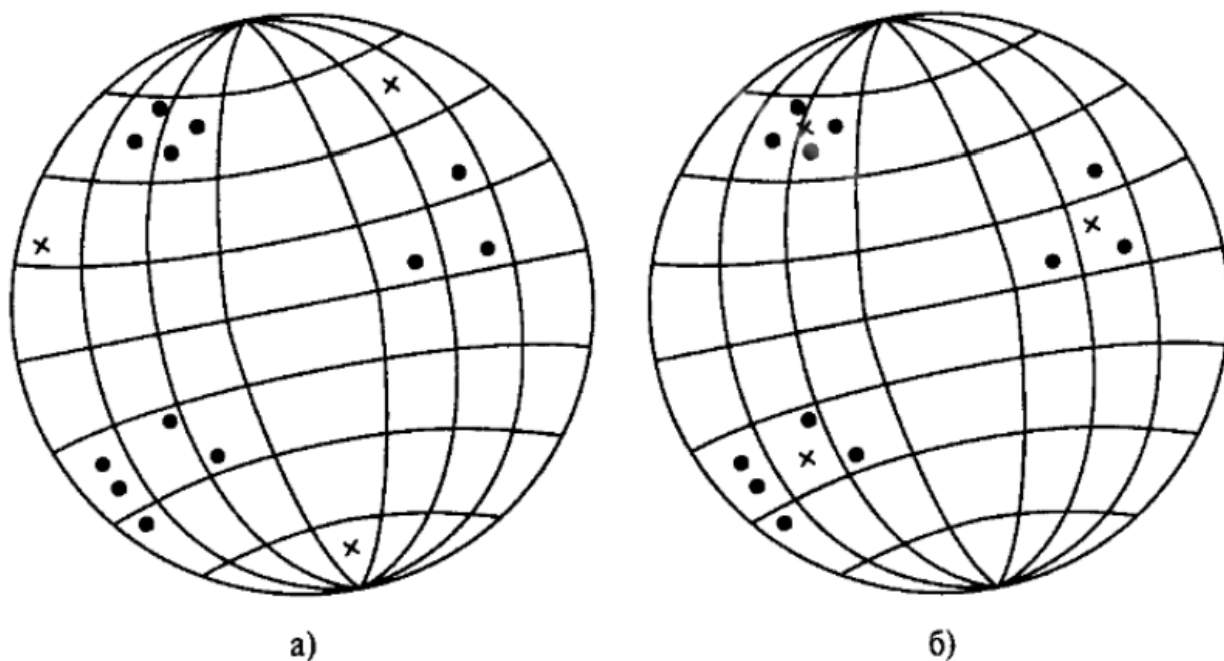
Хебб синапсы келесі 4 қасиетпен ерекшеленді:

- 1 Уақытқа байланысты. Хебб синапсы редсинаптикалық және постсинаптикалық сигналдардың пайда болу уақытының нақты уақытына байланысты.
- 2 Елді мекені (Локальность). Синапс өзінің табиғаты бойынша ақпараттық сигналдар (пресинапстық және постсинапстық элементтердің ағымдағы белсенділігін білдіретін) кеңістік-уақытша жақын орналасқан мәліметтерді тасымалдау түйіні болып табылады. Бұл жергілікті ақпарат берілген кіріс сигналына тән жергілікті синаптикалық модификацияларды орындау үшін Хебб синапсымен пайдаланылады.
- 3 Интерактивтілік. Хебб синапсындағы өзгерістер екі ұштағы сигналдармен анықталады. Бұл геббилік оқыту формасы пресинаптикалық және постсинаптикалық сигналдар арасындағы өзара әрекеттесу дәрежесіне байланысты екенін білдіреді (осы сигналдардың тек біреуінің негізінде болжау жасауға болмайды).
- 4 Корреляция. Хеббтің оқу постулатының бір түсіндірмесі синаптикалық байланыс тиімділігінің өзгеруінің шарты пресинаптикалық және постсинаптикалық сигналдар арасындағы тәуелділік болып табылады.

Бәсекелестік оқыту әдісі. Көптеген шығыс нейрондары бір уақытта жануы мүмкін Хеббилік оқытудан одан айырмашылығы, бәсекелестік оқытуда шығыс нейрондары бір – бірімен активтендірілу үшін бәсекелеседі. Бұл құбылыс жеңімпаз – бәрін алады ережесі ретінде белгілі. Ұқсас оқыту биологиялық нейрондық желілерде орын алады. Бәсекелестік арқылы оқыту кіріс деректерін кластерлеуге мүмкіндік береді: ұқсас мысалдар корреляцияға сәйкес желі бойынша топтастырылған және бір элементпен ұсынылған.

Жаттығу кезінде тек «жеңімпаз» нейронның салмақтары өзгертіледі. Бұл ереженің әсері желіде сақталған үлгіні өзгерту арқылы қол жеткізіледі (жеңімпаз нейронның қосылыс салмақтарының векторы), онда ол кіріс үлгісіне біршама жақындайды. 3 – суретте жарыс әдісі арқылы жаттығудың геометриялық иллюстрациясы берілген. Кіріс векторлары нормаланған және шар бетіндегі нүктелермен берілген. Үш нейронға арналған салмақ векторлары кездейсоқ мәндермен инициализацияланған. Жаттығудан кейінгі олардың бастапқы және соңғы мәндері сәйкесінше ба және бб суреттерінде Х белгісімен белгіленген.

Мысалдардың үш тобының әрқайсысы салмақ векторы анықталған топтың ауырлық центріне реттелген шығыс нейрондарының бірі арқылы анықталады.



6 - сурет – Бәсекелестік оқыту үдерісін геометриялық үлгісі.

Оқыту жылдамдығы параметрі 0 болмаса, желі ешқашан оқуды тоқтатпайтынын байқауға болады. Кейбір кіріс үлгісі оқу процесіндегі келесі итерацияларда басқа шығыс нейронды белсендіруі мүмкін. Бұл оқыту жүйесінің тұрақтылығы туралы сұрақты тудырады. Жаттығу үлгісіндегі мысалдардың ешқайсысы оқыту процесінің соңғы итерацияларынан кейін оның санат мүшелігін өзгертпесе, жүйе тұрақты болып саналады. Тұрақтылыққа жетудің бір жолы - оқыту жылдамдығы параметрін 0-ге дейін біртіндеп төмендету. Дегенмен, оқытудың бұл жасанды тежелуі жаңа деректерге бейімделу қабілетіне қатысты пластикалық деп аталатын басқа проблеманы тудырады. Бәсекеге қабілетті оқытудың бұл ерекшеліктері пластикалық тұрақтылық дилеммасы ретінде белгілі.

Кері таралу алгоритмі (back propagation) көпқабатты нейрондық желілерді оқыту әдістерінің бірі болып табылады.

Кері таралу алгоритмімен оқыту желінің барлық қабаттары арқылы екі өтуді қамтиды: алға және артқа. Алға өтуде кіріс векторы нейрондық желінің кіріс деңгейіне беріледі, содан кейін желі арқылы қабаттан қабатқа таралады. Нәтижесінде шығыс сигналдарының жиынтығы түзіледі, бұл желінің берілген кіріс кескініне нақты жауабы болып табыла алады. Алға өту кезінде желінің барлық

синаптикалық салмақтары бекітіледі. Кері немес артқа өту кезінде барлық синаптикалық салмақтар қатені түзету ережесіне сәйкес жаңартылады, яғни толығырақ айтқанда: желінің нақты шығысы қажетті шығыстан алынып тасталады, нәтижесінде қате сигналы пайда болады. Бұл сигнал кейіннен бүкіл желі бойынша синаптикалық байланыстардың бағытына қарама-қарсы бағытта таралады. Демек, кері таралу алгоритмі атауы. Синаптикалық салмақтар желілік шығыс сигналын қалағанға барынша жақындату үшін реттеледі [3].

Кері таралу алгоритмінде қате бетінің градиент векторы есептеледі. Бұл вектор берілген нүктеден бет бойымен ең қысқа түсу бағытын көрсетеді, сондықтан оның бойымен «аздап» қозғалсақ, қателік азаяды. Мұндай қадамдардың тізбегі ең аз дегенде бір немесе басқа түрге әкеледі. Бұл жерде белгілі бір қиындық – қадамдардың ұзақтығы туралы мәселе.

Үлкен қадам ұзындығымен конвергенция жылдамырақ болады, бірақ шешімнің үстінен секіру қаупі бар немесе егер қателік беті ерекше өңделген пішінге ие болса, ол дұрыс емес бағытта кетуі мүмкін [1]. Нейрондық желіні жаттықтыру кезіндегі бұл құбылыстың классикалық мысалы – алгоритмнің тік беткейлері бар тар саймен өте баяу қозғалуы, бір жағынан екінші жағына секіру. Керісінше, кішкене қадам сізді дұрыс жолға түсіруі мүмкін, бірақ ол көп қайталауды қажет етеді. Іс жүзінде қадам өлшемі еңістің тіктігіне пропорционалды болып қабылданады, осылайша алгоритм оқу жылдамдығы деп аталатын кейбір тұрақтымен минимумға жақын баяулайды. Оқыту жылдамдығын дұрыс таңдау нақты тапсырмаға байланысты және әдетте эмпирикалық түрде орындалады; бұл константа да уақытқа байланысты болуы мүмкін, алгоритм ілгерілеген сайын азаяды.

Қателік функциясын келесі формуламен анықтауға болады:

$$E = \frac{1}{2} \sum_{i=1}^p (Y_i - y_i)^2, \quad (2.2)$$

яғни, мындағы  $Y_i$  – қалаған желі шығысы,  $y_i$  – нақты желі шығысы.

$i=1 \dots p$ ,  $p$  – кіріс мәндерінің саны.

Қате бетінің градиент векторын келесі формуламен анықтасақ болады:

$$\nabla E = \left( \frac{\partial E}{\partial E w_1}, \frac{\partial E}{\partial E w_2}, \dots, \frac{\partial E}{\partial E w_n} \right), \quad (2.3)$$

яғни, мындағы  $n$  – желідегі салмақтардың саны.

Әрбір желі салмағы келесі формуладағы тең мәнге жаңартылады:

$$\Delta w_i = -\mu \frac{\partial E}{\partial w_i}, \quad (2.4)$$

яғни,  $\mu$  – нейрондық желіні үйрену жылдамдығы.

Келесі кезекте көпқабатты перцептронның аппроксиматор және классификатор ретіндегі мүмкіндіктерін зерттеп өтейік.  $n$  кірісі,  $m$  шығысы және олардың арасында  $l$  жасырын элементтері бар үш деңгейлі нейрондық желіні енгізіп, содан кейін салмақтың екі қабатын қарастырып және оны құру қажет: кірістерден жасырын элементтерге және шығысқа дейін, яғни ( $W_1, W_2$ ).

Кері таралу алгоритмінің мақсаты көп қабатты құрылымның барлық қабаттарын реттеу болып табылады. Бір жасырын қабаты және бір шығысы бар желі мысалында алгоритм жұмысын қарастырайық (сурет 7). Нейрондық желімен белгіленген кіріс сигналдарының түрлендірулері келесі формулалармен анықталады:

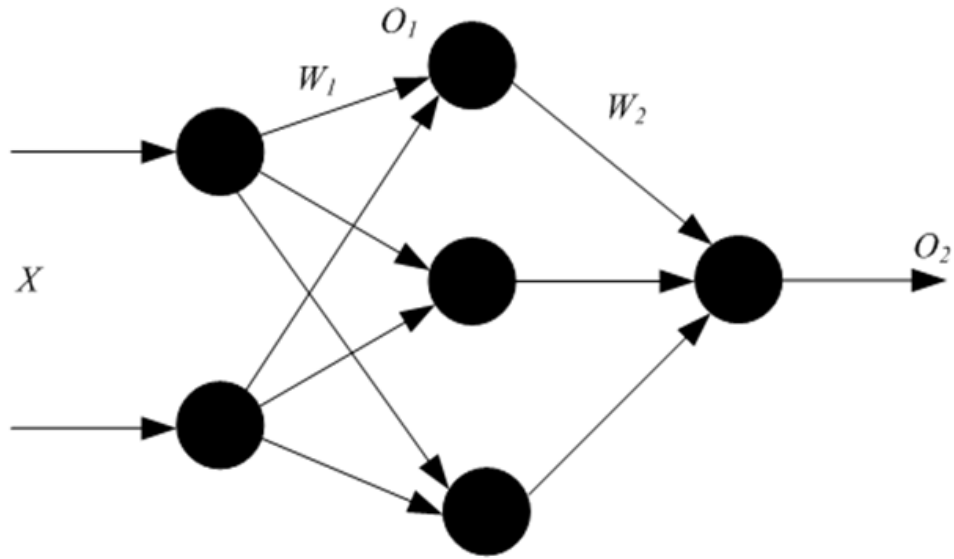
$$\begin{aligned} F(\langle W, X \rangle) &= 1 / (1 + \exp(-W^T X)); \\ O_1 &= 1 / (1 + \exp(-W_1^T X)); \\ O_2 &= 1 / (1 + \exp(-W_2^T O_1)). \end{aligned} \quad (2.5)$$

Жалпы қателік функциясы барлық қабаттардың салмағына байланысты, біздің жағдайда  $W_2$  векторында және  $W_1$  матрицасында:

$$E(W_1, W_2) = 1/2 (Y - 1 / (1 + \exp(-W_2^T O_1)))^2,$$

мұндағы  $Y$  - оқу жиынында көрсетілген шығыс.





7 - сурет – Көпқабатты перцептрон

Енді ішінара туындыларды пайдаланып әрбір салмақтың өсуін анықтауымыз керек:

$$\frac{\partial E(W_2, W_1)}{\partial W_2},$$

$$\frac{\partial E(W_2, W_1)}{\partial W_1}. \quad (2.6)$$

Көп қабатты архитектура үшін әр қабаттың салмақтық матрицасына қатысты қатенің ішінара туындылары күрделі туынды формуласы арқылы анықталады. Бірполярлы сигма тәрізді жағдайда салмақтарды өзгерту ережесі келесідей болады:

$$W_2 = W_2 + h \cdot (Y - O_2) \cdot O_2 \cdot (1 - O_2) \cdot O_1;$$

$$\partial = (Y - O_2) \cdot O_2 \cdot (1 - O_2);$$

$$W_1 = W_1 + h \cdot \partial \cdot W_1 \cdot (1 - O_1) \cdot O_1 \cdot X. \quad (2.7)$$

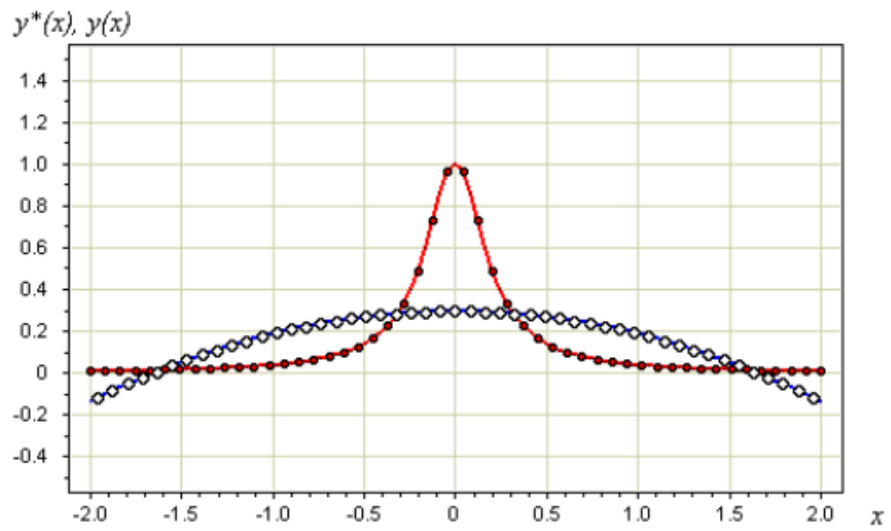
Осылайша, кері таралу әдісі аралық қабаттардағы қажетті мәндер көрсетілмегенімен, аралық қабаттардың салмақтарын өзгертуге мүмкіндік береді.

## 2.2 Кері даярлау және жалпылау.

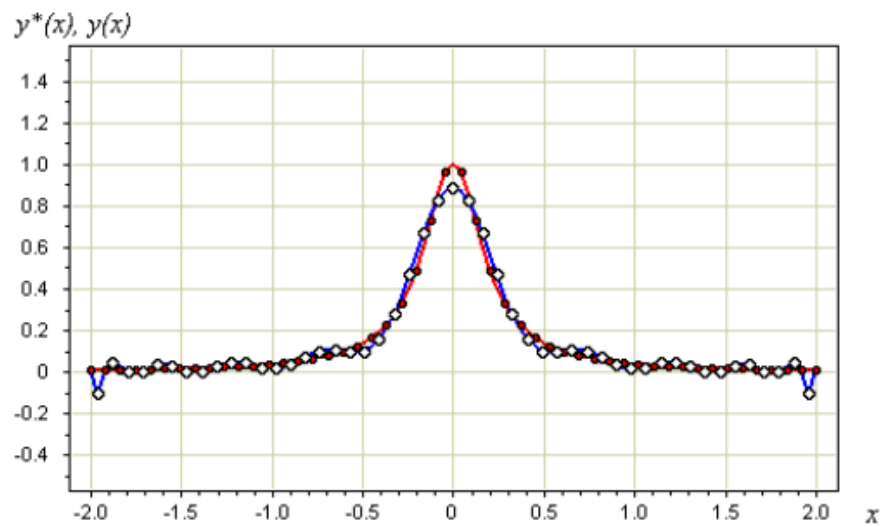
Көрсетілген тәсілдің ең күрделі қиындықтарының бірі - осылайша біз шынымен азайтуымыз керек қате қатені - желіге мүлдем жаңа бақылаулар берілген кезде күтуге болатын қатені азайтамыз. Басқаша айтқанда, нейрондық желінің нәтижені жаңа бақылауларға жалпылау мүмкіндігі болғанын қалаймыз. Шындығында, желі жаттығулар жиынтығындағы қатені азайтуға үйретілген, ал идеалды және шексіз үлкен жаттығулар жиынтығы болмаған жағдайда, бұл бұрынғы нұсқадағы қате бетіндегі «нақты» қатені азайтумен бірдей емес. құбылыстың белгісіз моделі.

Бұл айырмашылық артық орнату немесе тым жақын орнату мәселесінде айқын көрінеді. Бұл құбылысты нейрондық желі үшін емес, көпмүшелердің көмегімен жуықтау мысалын қолдану арқылы көрсету оңайырақ болады - және құбылыстың мәні мүлдем бірдей. Көпмүше (немесе көпмүше) – тәуелсіз айнымалының тек тұрақтылары мен бүтін дәрежелерін қамтитын өрнек. Көпмүшелердің графиктері әртүрлі пішіндерге ие болуы мүмкін және көпмүшенің дәрежесі неғұрлым жоғары болса (және, осылайша, құрамында неғұрлым көп термин болса), соғұрлым бұл пішін күрделірек болуы мүмкін. Егер бізде кейбір деректер болса, біз оған полиномдық қисық (модель) орнату мақсатын қоя аламыз және осылайша бар қарым-қатынасқа түсініктеме аламыз. Біздің деректеріміз шулы болуы мүмкін, сондықтан біз ең жақсы үлгі барлық қол жетімді нүктелерден дәл өтетін қисықпен берілген деп болжауға болмайды. Төменгі ретті көпмүше деректерге сәйкес келу үшін жеткілікті икемді болмауы мүмкін, ал жоғары ретті көпмүше тым икемді болуы мүмкін, деректерді мұқият қадағалайды, бірақ нақты қатынастың пішініне ешқандай қатысы жоқ икемді пішінді алады.

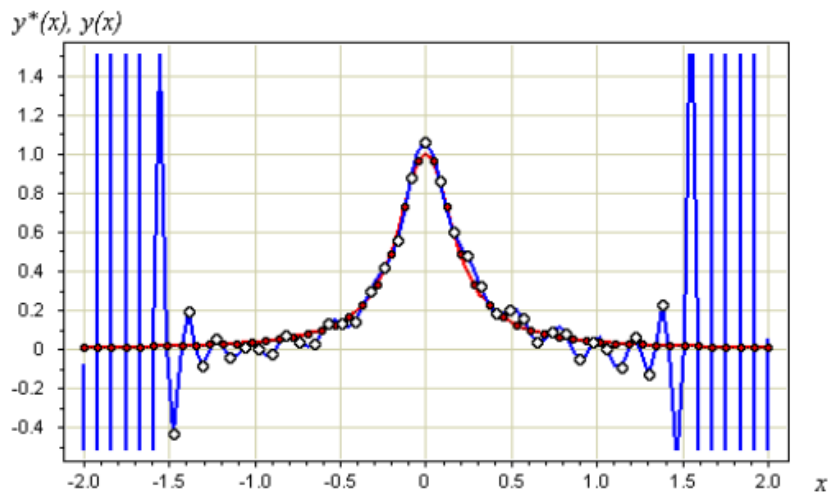
Нейрондық желі дәл осындай қиындыққа тап болады. Салмақтары көбірек желілер күрделі функцияларды модельдейді, сондықтан шамадан тыс орнатуға бейім. Салмақ саны аз желі бар тәуелділікті модельдеу үшін жеткілікті икемді болмауы мүмкін. Мысалы, аралық қабаттары жоқ желі нақты сызықтық функцияны модельдейді.



8 - сурет – Нейрондардың жеткіліксіз саны. Желінің дайындығы төмен.



9 - сурет – Оптималды желінің күрделілігі.



10 - сурет – Нейрондардың артық саны. Желі шамадан тыс дайындалған және тұрақсыз.

Әрқашан дерлік күрделі желі кішірек қателік береді, бірақ бұл модельдің жақсы сапасын көрсетпейді, керісінше шамадан тыс қондыруды көрсетеді.

Желінің дұрыс күрделілігін таңдау үшін басқарудың кросс-валидация механизмін қолдану қажет. Біз жаттығулардың кейбір бақылауларын сақтаймыз және оларды кері таралу алгоритмін оқытуда пайдаланбаймыз. Оның орнына, алгоритм іске қосылғанда, олар нәтижені тәуелсіз бақылау үшін пайдаланылады. Жұмыстың ең басында оқыту және бақылау жиынтықтарындағы желілік қателік бірдей болады (егер олар айтарлықтай ерекшеленетін болса, онда барлық бақылаулардың екі жиынтыққа бөлінуі гетерогенді болуы мүмкін). Желі үйренген сайын жаттығу қатесі табиғи түрде азаяды және жаттығу нақты қате функциясын азайтқанша, сынақ жиынындағы қате де азаяды. Егер басқару қатесі азаюды тоқтатса немесе тіпті көбейе бастаса, бұл желінің деректерді тым жақындата бастағанын және жаттығуды тоқтату керектігін көрсетеді. Оқыту процесі кезінде шамадан тыс дәл жуықтау құбылысы артық фитинг деп аталады. Бұл орын алса, әдетте жасырын элементтердің және/немесе қабаттардың санын азайту ұсынылады, себебі желі бұл тапсырма үшін тым қуатты. Егер, керісінше, желі бар тәуелділікті модельдеу үшін жеткілікті бай болмаса, онда қайта даярлау, ең алдымен, орын алмайды және екі қате де – оқыту және тестілеу – кішігірім жеткілікті деңгейіне жете алмайды.

Жергілікті минимумдармен сипатталған мәселелер және желі өлшемін таңдау тәжірибеде нейрондық желілермен жұмыс істегенде, әдетте, әртүрлі желілердің үлкен санымен тәжірибе жасауға, кейде олардың әрқайсысын бірнеше рет



жаттықтыруға (сондықтан) әкеледі. жергілікті минимумдармен жаңылыспау үшін) және алынған нәтижелерді салыстыру. Нәтиже сапасының негізгі көрсеткіші бақылау қателігі болып табылады. Сонымен қатар, жалпы ғылыми принципке сәйкес, оған сәйкес, басқалары тең болған жағдайда, басқару қателері шамамен бірдей екі желіден қарапайым модельге артықшылық беру керек, кішірекін таңдау мағынасы бар.

Қайталанатын эксперименттердің қажеттілігі басқару жиынтығының модель таңдауда шешуші рөл атқара бастауына, яғни оқу процесінің бір бөлігіне айналуына әкеледі. Бұл оның модель сапасының тәуелсіз критерийі ретіндегі рөлін әлсіретеді - көптеген эксперименттермен басқару жиынтығында жақсы нәтиже беретін «сәтті» желіні таңдау қаупі бар. Соңғы модельге тиісті сенімділік беру үшін олар жиі (кем дегенде оқу деректерінің көлемі мүмкіндік бергенде) мұны жасайды: олар тағы біреуін резервте қалдырады - бақылаулардың сынақ жинағы. Қорытынды модель оқыту және бақылау жиындарында қол жеткізілген нәтижелердің оқу процесінің артефактілері емес, шынайы екендігіне көз жеткізу үшін осы жиынтық деректері бойынша сыналады. Әрине, өз рөлін жақсы атқару үшін тест жинағы тек бір рет пайдаланылуы керек: егер ол оқу процесін реттеу үшін қайта пайдаланылса, ол шын мәнінде бақылау жинағына айналады.

### 3 Функцияның жуықтауы

#### 3.1 Нейрондық желілер арқылы функцияларды даярлау

Аппроксимация (функцияны жуықтау) – мәндер кестесіне ең жақсы сәйкес келетін функцияны табу әдісі. Жақындау кезінде функция түрі таңдалады және бұл функцияның параметрлері жуықтау функциясының мәндері кесте мәндеріне жақын болатындай етіп анықталады.

Нақтырақ айтар болсақ  $f(x)$  функциясын жуықтау (аппроксимациялау), яғни берілген мәнге жақын болатын  $g(x)$  функциясын табу, жақындату болып табылады. Бұл екі функцияның ( $f(x)$  және  $g(x)$ ) жақындығы үшін бірнеше критерилер әзірленген.

Аппроксимацияның негізгі міндеті берілген нүктелердің айналасында немесе үздіксіз функцияның графигі айналасында ең жақын өтетін жуықтау функциясын құру болып табылады. Мұндай мәселенің пайда болуы бастапқы деректерде қателердің болуымен байланысты. Бұл жағдайда интерполяцияда орындалғандай, бастапқы деректердің барлық нүктелері арқылы функцияны дәл орындау мүмкін емес.

Аппроксимация объектінің сандық сипаттамаларын және сапалық қасиеттерін зерттеуге мүмкіндік береді, мәселені қарапайым немесе ыңғайлырақ объектілерді зерттеуге дейін азайтады. Мысалы, сипаттамалары оңай есептелетін немесе қасиеттері бұрыннан белгілі.

Түпнұсқалық және жуықтау функцияларының жақындығы сандық өлшеммен – жуықтау критеріімен анықталады. Ең көп қолданылатын критерий – бұл «эксперименттік» мәндерден есептелген мәндердің квадраттық ауытқуларының қосындысына тең берілген нүктелердегі жақындық. Бұл критерий «квадрат» деп аталады. Оны келесідей формуламен көре аламыз:

$$R = \sum_{i=1}^n \beta_i (y_i - g_i)^2. \quad (3.1)$$

Мұнда,  $y_i$  –  $i$ -ші нүктедегі функцияның берілген кестелік мәндері;  $g_i$  –  $i$ -ші нүктеде жуықтау функциясының мәндері;  $\beta_i$  –  $i$ -ші нүктенің салыстырмалы маңыздылығын,  $R$  критерийін минимизациялау кезінде  $\beta_i$ -нің ұлғаюы, ең алдымен,  $i$ -ші нүктедегі ауытқудың төмендеуіне әкелетіндей етіп көрсететін салмақ коэффициенттері.

Жоғарыда атап өткен квадраттық критерийдің бірқатар артықшылықтары бар. Соның ішінде дифференциалдау мен көпмүшелік жуықтау функциялары бар жуықтау есебінің бірқатар шешімі ретінде қарастырылады.

Аппроксимациялау функциясының негізгі екі шешімі бар:

- 1) көрсетілгеннен аспайтын қатесі бар қолда бар деректерді сипаттайтын жуықтау функциясын алу;
- 2) ең жақсы қателікпен берілген құрылымның жуықтау функциясын алу.

Көбінесе бірінші мәселе әртүрлі жуықтау функцияларын іздеу және одан кейін ең жақсысын таңдау арқылы екіншісіне дейін төмендейді.

Жоғарыда көрсетілгендей, жуықтауды нүктелердің бар жиынынан функционалдық қатынасты табуға дейін азайтуға болады. Бұл жағдайда функционалдық тәуелділік нейрондық желі негізінде, яғни нейрондық желілерді белсендіру функцияларының комбинациясы арқылы ұсынылатын болады.

Көпқабатты перцептронды оқыту көбінесе кері таралу алгоритмі немесе оның модификациясы арқылы жүзеге асады.

Қатені кері тарату – жасанды нейрондық желіні оқытудың жүйелі әдісі. Оның маңызды математикалық негізі бар. Кейбір шектеулерге қарамастан, бұл алгоритм жасанды нейрондық желілерді қолдануға болатын мәселелер аймағын едәуір кеңейтті.

Әдіс жасанды нейрондық желіні оқыту үлгісінің мысалдарын қарап, осы үлгілерді пайдалана отырып оқытуға негізделген. Жаттығу жиынынан әрбір мысал үшін жасанды нейрондық желі шығысының қажетті мәні белгілі болып саналады, яғни кері таралу әдісі бақылаудағы оқыту болып табылады.

Мұндай оқыту жасанды нейрондық желінің оқытылған параметрлерін – синапстық салмақтардың  $w$  мәндерін, белсендіру функцияларының параметрлерін реттеу арқылы бүкіл жаттығу жиынында қателік функциясын  $E$  азайтудың оңтайландыру мәселесін шешу ретінде ұсынылуы мүмкін. Бұл функцияны келесідей формуламен бақылай аламыз:

$$E = \frac{1}{2} \sum_i^n (t_i - y_i)^2, \quad (3.2)$$

мұндағы,  $t_i$  –  $i$ -ші үлгідегі қажетті шығыс мәні;  $y_i$  – нақты (жасанды нейрондық желімен есептелген) мән;  $n$  – оқу жинағының үлгілерінің саны.

Кері таралу әдісіндегі қателік функционалды  $E$  минимизациясы градиентті оңтайландыру әдістерінің бірін қолдану арқылы жүзеге асырылады, мысалға, «ең тік түсу» немесе оның модификациялары. Бұл жағдайда синапстардың салмақтары

кателік функцияның ең үлкен өсу бағытына қарама-қарсы бағытта өзгереді, яғни антиградиент  $E$  бағыты бойынша:

$$W(t+1) = W(t) - \eta \frac{\partial E}{\partial W}, \quad (3.3)$$

мұндағы,  $0 < \eta \leq 1$  – «оқу қадамы», бірнеше әдістерде итеративті түрде таңдалатын оқу жылдамдығын анықтайтын пайдаланушы анықтайтын параметр.

Қатені кері тарату әдісі екі нұсқада жүзеге асырылады. Бірінші нұсқада оқытылған параметрлер бүкіл оқу үлгісін жібергеннен кейін қайта есептеледі және қатенің пішіні бар:

$$E = \frac{1}{2} \sum_i^p (t_i - y_i)^2. \quad (3.4)$$

Екінші тәсілде қате әрбір мысалдан кейін қайта есептеледі:

$$E_i = \frac{1}{2} (t_i - y_i)^2, \quad (3.5)$$

$$E(W) = \frac{1}{2} \sum_{i,p} (t_i - y_i)^2. \quad (3.6)$$

Екінші тәсілде формула осы күйде болса:

$$\Delta W_{ij}(t) = -\eta \frac{\partial E}{\partial W_{ij}}, \quad (3.7)$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}(t). \quad (3.8)$$

Онда:

$$\frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial y_{ij}} \cdot \frac{\partial y_i}{\partial S_i} \cdot \frac{\partial S_j}{\partial W_{ij}}, \quad (3.9)$$

$y_i(S_i)$  – белсендіру функциясы. Үшін:



$$y_j = \bar{f}(S) = \frac{1}{1 + e^{-S_j}}, \quad (3.10)$$

$$\begin{aligned} \bar{f}'(S) &= \frac{\partial}{\partial S_j} \left( \frac{1}{1 + e^{-S_j}} \right) = \frac{1}{(1 + e^{-S_j})^2} (-e^{-S_j}) = \\ &= \frac{1}{1 + e^{-S_j}} * \frac{-e^{-S_j}}{1 + e^{-S_j}} = y_j(1 - y_j), \end{aligned} \quad (3.11)$$

$$\frac{dy_i}{dS_j} = (1 - y_j)^2; \quad (3.12)$$

Үшінші факторды қарастырайық:

$$S_j = \sum W_{ij} \cdot y_i^{n-1}; \quad (3.13)$$

$$\frac{\partial S_i}{\partial W_{ij}} = y_i^{(n-1)}. \quad (3.14)$$

Осылайша k-ден астам қосындылау n-ші қабаттың нейрондары арасында болатынын көрсетуге болады.

$$\frac{\partial E}{\partial y_i} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{dS_k} \cdot \frac{\partial S_k}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{dS_k} \cdot W_{jk}^{(n+1)}, \quad (3.15)$$

Жаңа белгіні енгізсек:

$$\delta_j = \frac{\partial E}{\partial y_k} \cdot \frac{dy_j}{dS_j}, \quad (3.16)$$

$$\delta_j^{n-1} = \left[ \sum_k \delta_k^n \cdot W_{jk}^n \right] \cdot \frac{dy_j}{dS_j}. \quad (3.17)$$

Ішкі нейрон үшін (3.17)

$$\delta_j^n = (d_j^n - y_j^n) \cdot \frac{dy_j}{dS_j}. \quad (3.18)$$

Сыртқы нейрон үшін (3.18)

$$\Delta W_{ij}^n = -\eta \delta_j^{(n)} \cdot y_j^{n-1}; \quad (3.19)$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}. \quad (3.20)$$

Жоғарыда көрсетілген (3.5) – (3.20) формулаларын ескере отырып, кері таралу әдісімен оқыту келесі алгоритмге сәйкес жүзеге асырылады:

- 1 Оқытылған параметрлердің бастапқы мәндері инициализацияланады. Әдетте бастапқы мәндер шағын кездейсоқ шама болып табылады.
- 2 Жаттығу үлгісінің бір мысалы жасанды нейрондық желілердің кірістеріне беріледі. Жасанды нейрондық желілердің қалыпты жұмыс режимінде барлық нейрондардың шығыс мәндері есептеледі. Яғни нақтырақ айтқанда сигналдар кірістерден шығыстарға дейін таралады.
- 3 Шығу қабатының нейрондары үшін  $\delta_j^n$  мәндері сыртқы нейрон үшін формула бойынша есептеледі (3.18).
- 4  $\delta_j^n$  мәндері (3.17) формула бойынша барлық ішкі нейрондар үшін есептеледі.
- 5 (3.19) формуласын қолданып, барлық қосылымдар үшін  $\Delta W_{ij}^n$  салмақ коэффициенттерінің өсулері есептеледі.
- 6 (3.20) формула арқылы синапстардың салмағын реттей аламыз:  
 $W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}$ .
- 7 Е қателік функциясының мәні жеткілікті аз болғанша жаттығу үлгісінің әрбір мысалы үшін 2 мен 6 қадамдарды қайталап отырамыз.

7 қадамда көрсетілген циклды тоқтатудың баламалы шарты жасанды нейрондық желілерді қолайлы дәлдікке үйрету мүмкін еместігі болып табылады. Бұл шешім  $\Delta W_{ij}^n$  салмақ коэффициенттерінің өсімдері тым аз болатынына және осы нүктедегі итерациялар саны көрсетілгеннен көп болатынына негізделген.

Бұл жағдай келесі екі жағдайда туындауы мүмкін:

- 1 Оңтайландыру алгоритмі Е қатесінің функционалдық экстремумына жетті. Бұл мүмкіндікті болдырмау үшін оқытуды оқытылған параметрлердің басқа бастапқы конфигурациясынан бастау қажет.
- 2 Жасанды нейрондық желілердің архитектурасы мәселені шешу үшін жеткіліксіз. Егер желіні қайталап оқыту әлі де табысты оқуға әкелмесе, онда нейрондар санын немесе тіпті қабаттар санын көбейту арқылы желі қиындауы керек.

### 3.2 Matlab жүйесінде көпқабатты желіні енгізу.

1989 жылы Л.Фунахаши дәлелдеген теорема бойынша, ерікті түрде күрделі функцияны ерікті активтендіру функциялары бар екі қабатты нейрондық желі арқылы жақындатуға болады. Сондықтан классикалық түрде жуықтау есебін шешу үшін көпқабатты перцептрон қолданылады және бұл бір немесе бірнеше айнымалылардың функциялары болуыда мүмкін.

Мысал ретінде сызықтық шуы бар бір айнымалы шудың функциясын қарастырайық.

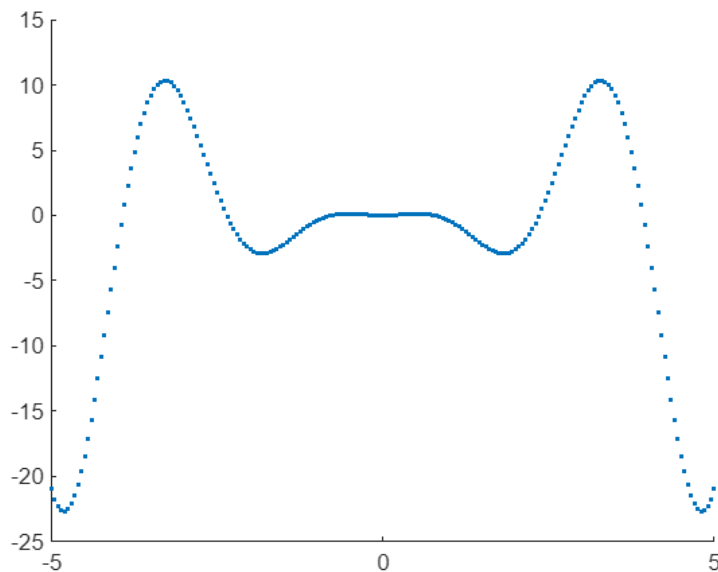
$$y = x^2 \sin\left(2x + \frac{\pi}{2}\right) \quad (3.21)$$

Практика жүзінде көрсету үшін, үлгіні жасау MatLab бағдарламасын қолдана отырып жасаймыз (11 - сурет).

```
% Функция үшін деректер жиынын құру
L = 200;
% X мәндерін құру
X = -5:(10/L):5;
% Сызықтық шуы бар шулы функция нүктелерін есептеу
Noise = 0.5*rand(1,(L+1))*2-1.0;
Y = X.^2.*sin(2*X+pi/2)+Noise;
```

11 - сурет – Жасанды нейрондық желілерді оқыту үшін бастапқы деректерді генерациялау

Жасалған нүктелер жинағы суретте көрсетілгендей болады (12 - сурет).



12 - сурет – Жасанды нейрондық желілерді оқытуға арналған функция графигі

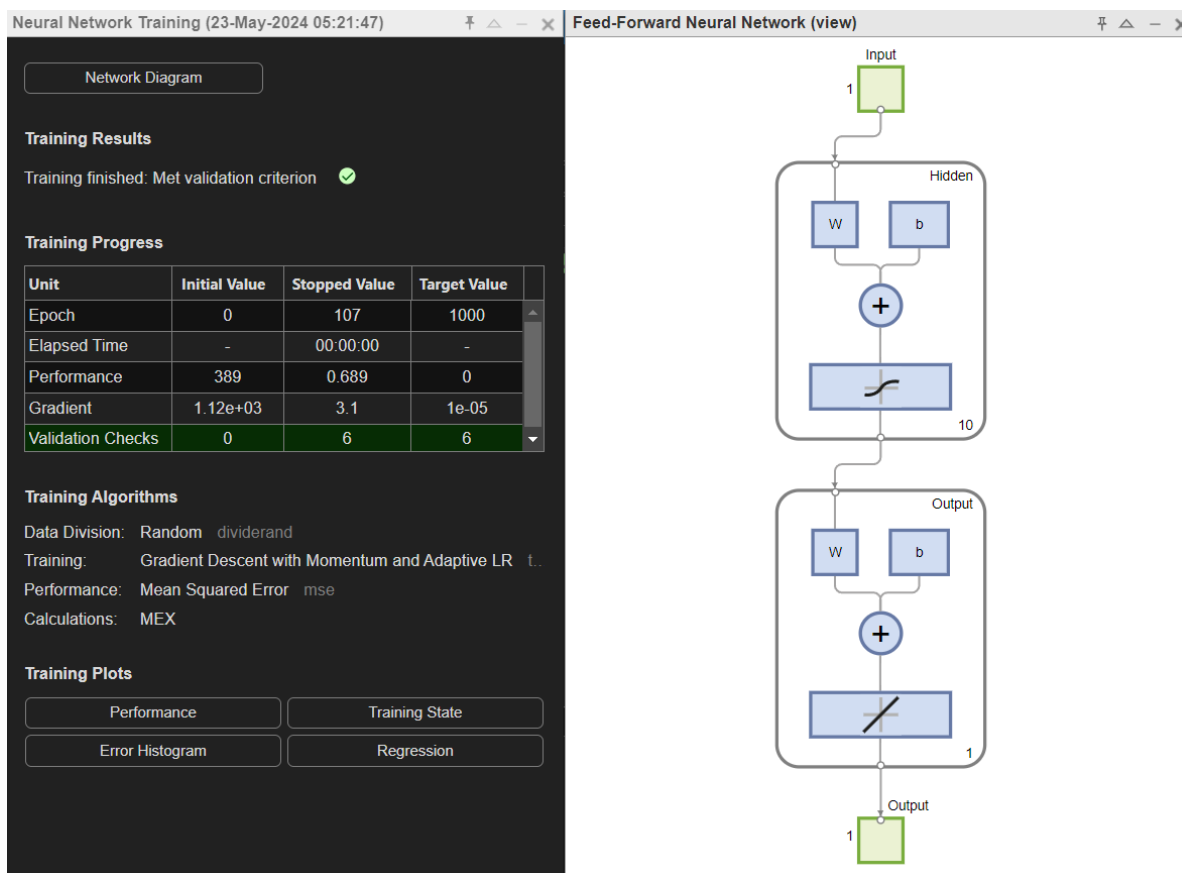
Жасанды нейрондық желілерді `feedforwardnet` пәрменін пайдалана отырып, жасалған деректерді жуықтаймыз, `train` пәрменін пайдаланып шыққан нейрондық желіні жаттықтырамыз, содан кейін оны `sim` функциясының көмегімен мәндердің жаңа диапазонында сынап көреміз (13 - сурет).

```
hold on;
% Жаттығу үлгісін визуализациялаймыз
plot (X,Y,'LineStyle','none','Marker','.', 'MarkerSize',5);
% Жасанды нейрондық желіні құрайық
Net = feedforwardnet(10,'traingdx');
% Жасанды нейрондық желіні оқыту
Net = train(Net,X,Y);
% Жасанды нейрондық желіні жаттығу жиынтығынан кеңірек аралықта сынаймыз
X_rez = -10:0.1:10;
Y_rez = sim(Net,X_rez);
% Жуықтау(аппроксимация) нәтижелерін визуализациялаймыз
plot(X_rez,Y_rez,'LineStyle','none','Marker','.', 'MarkerSize',15);
hold off
```

13 - сурет – Жасанды бағыттаушы нейрондық желілерді құру және оқыту.

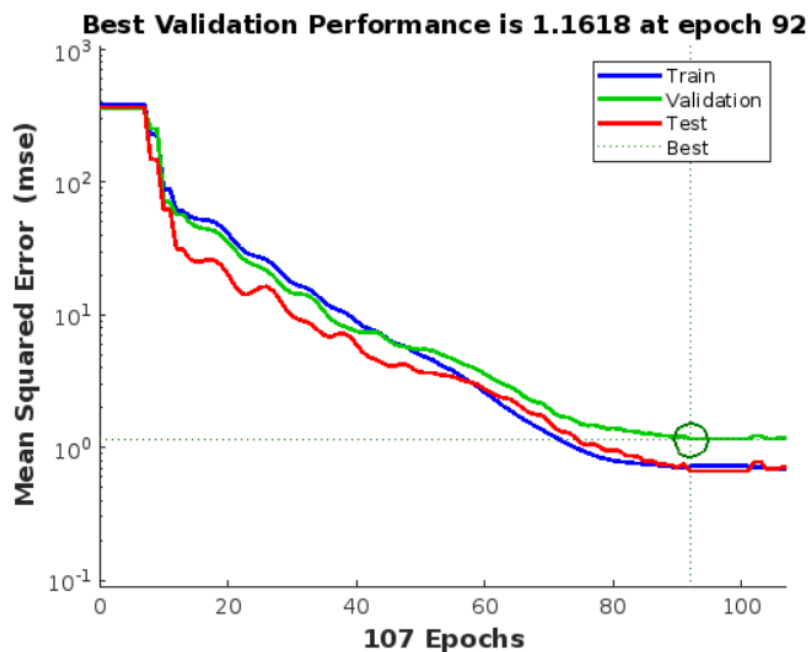
Оқыту процесі кезінде MatLab нейрондық желінің интерфейстік формасын көрсетеді (14 - сурет). Жүйенің архитектурасы көрінеді – қабаттардың саны (2) және

олардағы нейрондар, белсендіру функцияларының түрлері (“tansig” және “purelin”), оқыту алгоритмі (модифицирленген градиенттің түсуі), қате функциясының түрі (орташаквадраттық), сонымен қатар оқу үлгерімінің параметрлері бейнеленген.



14 - сурет – MatLab-та нейрондық желінің интерфейстік формасы

Performance түймешігінде жаттығу циклары бойынша желі қатесінің динамикасының графигі бар (15 - сурет).



15 - сурет – Жаттығу барысы кестесі

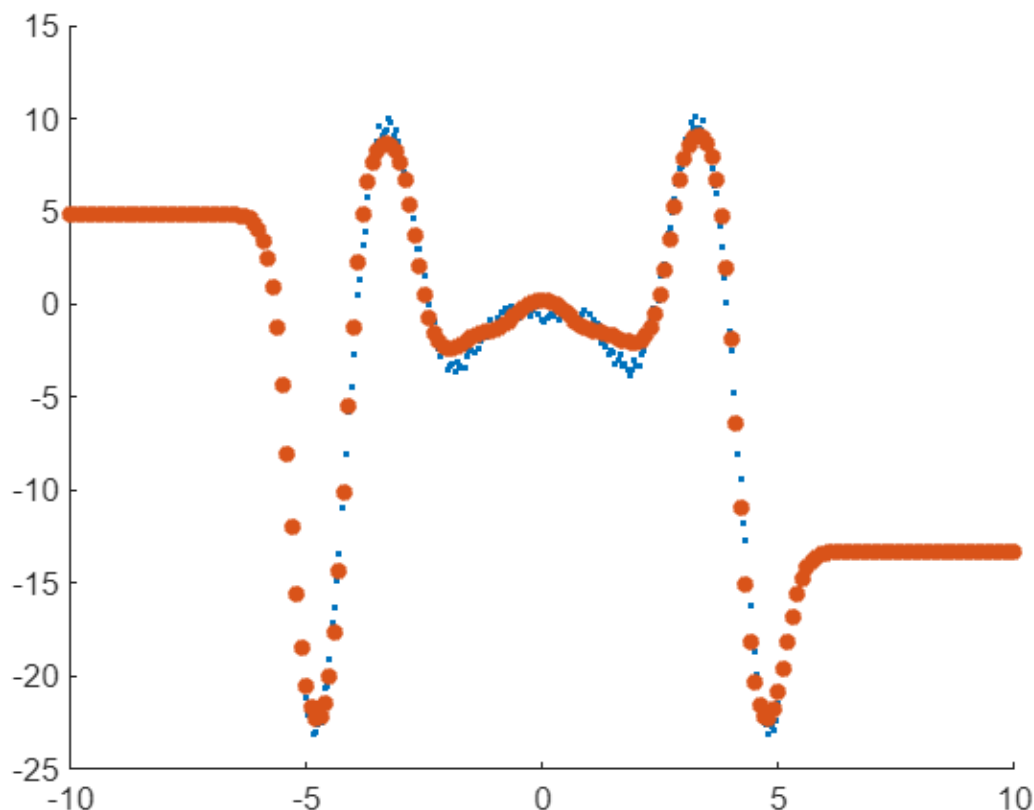
Жасанды нейрондық желінің қателік функционалдық мәні алты дәйекті дәуірге жоғарылағанда оқу процесі әдепкі бойынша тоқтайды.

Бұл мысалда нәтиже келесі себептерге сәйкес келеді:

1. Соңғы орташа квадрат қатесі (MSE) аз.
2. Validation және Test жиындарының қателігі ұқсас сипаттамаларға ие.
3. Қайта даярлау болған жоқ (тоқтату нүктесінен кейін 201-ші дәуірге дейін тестілеу кешенінің MSE ғана артады).

16 - суретте жуықталған функцияны, `sim` функциясы арқылы желіні есептеу нәтижесін көре аламыз. 13 суретте көрсетілгендей, желілік тестілеу оқу жинағына қарағанда кеңірек диапазонда жүзеге асырылып жатыр.

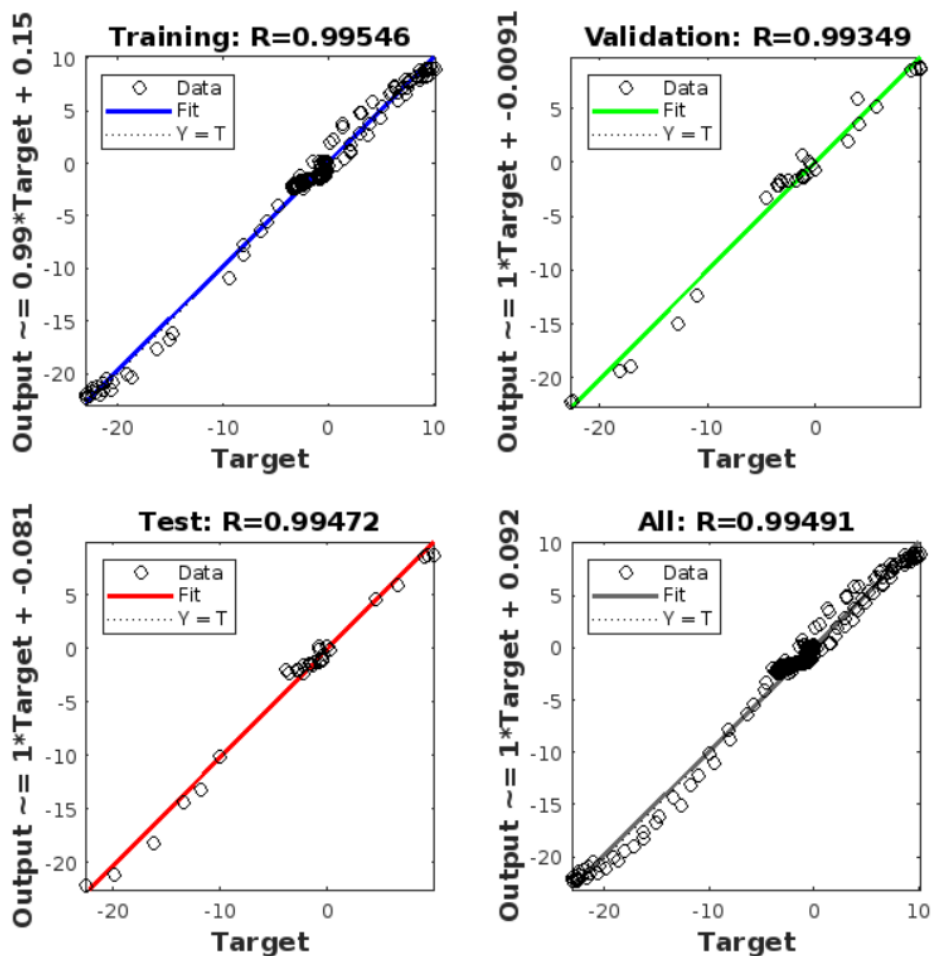




16 - сурет – Жуықталған (Аппроксимированная) функция.

Оқу үлгісінің диапазонында жасанды нейрондық желілерді болжау бастапқы функцияның нүктелік бұлтына өте жақын екенін көруге болады. Сондай-ақ, жақын маңда (жаттығу үлгілерінің диапазонының шамамен 10%) жасанды нейрондық желілердің де бастапқы функцияның әрекетін жаңғыртып, кейбір экстраполяциялық қабілеті бар, ал кеңірек жақын жерде ол түзу сызыққа жақындайтынын көруге болады.

Жасанды нейрондық желілерді оқыту нәтижесін бағалаудың қосымша құралы жасанды нейрондық желілердің шығыс мәндерінің (Output) регрессия функцияларын құру (17 - сурет) болып табылады, оларда көрсетілген мақсатты мәндерден (Target) жаттығу жиынтығы. Бұл диаграмма Регрессия түймесі арқылы шақырылады (14 - суреттен байқай аламыз).



17 - сурет – Есептеу және оқыту нәтижелерінің корреляциясы.

Құрылған жасанды нейрондық желілер бастапқы деректерге қанағаттанарлықсыз жақындаса, келесі әдістерді қолдана отырып, шешім іздеуге болады:

- 1 желінің құрылымын - қабаттар мен нейрондар санын арттыру;
- 2 оқыту үлгісінің көлемін ұлғайту;
- 3 егер бар болса, енгізу параметрлері кеңістігінің өлшемін ұлғайту;
- 4 жасанды нейрондық желілер үшін оқыту алгоритмін ауыстыру.

Біздің мысалда Net желісі жеткілікті түрде дайындалған.

### 3.3 Жұмыстың орындалу тізімі

- 1 Берілген функция үшін оның кестелік мәндерін құрастырыңыз, яғни нақтырақ айтқанда мәндер саны жуықталған функцияның кестелік мәндермен визуалды сәйкес келуі үшін жеткілікті болуы керек.
- 2 Осы кесте мәндерін жуықтау және жуықталған функцияны табу үшін нейрондық желіні пайдаланыңыз. Train функциясын пайдаланып нейрондық желіні жаттықтырыңыз.
- 3 Графикте кесте мәндерін нүктелермен белгілеп жуықталған функциясын сызық түрінде көрсетіңіз.
- 4 Жуықтау нәтижесінің қатесінің сандық мәнін тауып, экранға шығарыңыз.

## ҚОРЫТЫНДЫ

Жалпы дипломдық жұмысты жазу барысында келесі зерттеулерді анықтадық:

- Нейрондық желілер жайлы жалпылама түсінік;
- Жасанды нейрондық желінің әртүрлі архитектуралары;
- Кері таралу әдісімен нейрондық желілерді оқыту;
- Нейрондық желі арқылы функцияларды жуықтау.

Дипломдық жұмысты орындау, жазу барысында Matlab жүйесімен қарапайым кодтарды теріп, нейрондық желілердің графиктері мен параметрлерімен танысып өткен болатынбыз. Matlab жүйесінде кері таралу әдісін қолдана отырып, нейрондық желілерді қолдану арқылы функцияларды жуықтап қарастырылған болатын.

Сонымен қатар 3.3 жоспарда айтылып өткендей, осы тізімді ұстана отырып Matlab жүйесінде практиккалық тұрғыда бір айнымалыны қолдана отырып нейрондық желіні жуықтау функциясымен қарастырған болатынбыз.

## ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

- 1 Хайкин С. Нейронные сети. Полный курс. М.: Вильямс, 2006. – 1104 с.
- 2 Уварова А.В., Подколзин В.В. Аппроксимация параметров сети на основе неполной информации о предметной области: Математические методы и информационно-технические средства: материалы XI Всероссийской научнопрактической конференции, г. Краснодар Краснодарский университет МВД России, 2015. – с. 303-308.
- 3 Информационно-аналитический ресурс, посвященный машинному обучению. URL: <http://www.machinelearning.ru> / (дата обращения 02.12.2017).
- 4 Документация библиотеки SFML. URL: <https://www.sfml-dev.org> / (дата обращения 20.10.2017).
- 5 Christopher M. Bishop. Neural Networks for Pattern Recognition, 1995-482с.
- 6 Онлайн Matlab <https://matlab.mathworks.com/>
- 7 Моделирование нейронных сетей в системе MatLab / М. Г. Доррер ; СибГУ им. М. Ф. Решетнева. – Красноярск, 2021. – 98 с.