

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Електронної та біомедичної інженерії

Кафедра __ Мікроелектроніки, електронних приладів та пристроїв

КВАЛІФІКАЦІЙНА РОБОТА **Пояснювальна записка**

рівень вищої освіти _____ другий (магістерський)

Способи розпізнавання образів за допомогою нейронних мереж у різних програмних середовищах (тема)

Виконав:
студент 6-го курсу, групи МНПм-21-1

Спеціальність 153 Мікро- та наносистемна техніка

Тип програми _____ освітньо-професійна

Освітня програма « Мікро та нанoeлектронні прилади та пристрої»

Керівник доцент кафедри МЕЕПП, кандидат фізико-математичних наук Глухов О. В.

Допускається до захисту

Зав. кафедри

_____ (підпис)

_____ (прізвище, ініціали)

2022 р.

Харківський національний університет радіоелектроніки

Факультет _____ Електронної та біомедичної інженерії _____

Кафедра _Мікроелектроніки, електронних приладів та пристроїв _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 153 Мікро- та наносистемна техніка _____

Тип програми _____ освітньо-професійна _____

Освітня програма _« Мікро та наноелектронні прилади та пристрої» _____

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Колисниченку Микиті Сергійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи Способи розпізнавання образів за допомогою нейронних мереж у різних програмних середовищах

затверджена наказом університету від _____ 14.11 _____ 2022 _р. № _1473Ст_____

2. Термін подання студентом роботи до екзаменаційної комісії _____ 20 __р.

3. Вихідні дані до роботи згортовка нейронна мережа, MATLAB

4. Перелік питань, що потрібно опрацювати в роботі _____

_1) Провести аналіз існуючих приладів _____

_2) реалізувати макет голографічного дисплею _____

_3) проаналізувати отримані результати _____

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри): схема

алгоритму _____ роботи _____ системи, _____ слайд-презентація(_____ б

слайдів) _____

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз технічного завдання	26.09.22	
2	Аналітичний огляд джерел	28.09.22	
3	Оформлення пояснювальної записки	05.10.22	
4	Вибір структури приладу, збірка	10.10.22	
5	Розробка та написання програм	19.10.22	
6	Оформлення креслеників	01.11.22	
7	Підготовка презентації	16.11.22	
8	Рецензування, нормоконтроль	05.12.22	
9	Захист роботи	15.10.22	

Дата видачі завдання _____ 20__ р.

Студент _____
(підпис)

Керівник роботи _____
(підпис) _____ (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи магістра: 67 стр, 38 рисунків, 5 таблиць, 15 посилань.

НЕЙРОННА МЕРЕЖА, ШТУЧНИЙ ІНТЕЛЛЕКТ, РОЗПІЗНАВАННЯ ОБРАЗІВ, MATLAB

Об'єкт дослідження – спосіб розпізнавання образів.

Мета роботи – вивчення принципу роботи нейронної мережі у розпізнаванні образів.

Метод – аналітичний

У світі штучні інтелектуальні системи завоювали велику популярність, завдяки багатим можливостям і ефективності використання. Існує безліч завдань, де необхідне практичне дозвіл із використанням інтелектуальних систем вже зараз: економіка та бізнес, робототехніка, геологорозвідка, математика, біофізика, авіоніка, безпека та охоронні системи, медицина та багато інші. Така велика кількість областей застосування, говорить про те, що інтелектуальні системи – це унікальний набір для вирішення питань аналізу та обробки великого обсягу даних, вирішення завдань різного рівня складності. Нині у системах штучного інтелекту активно використовуються штучні нейронні мережі. Нейронна мережа – це обчислювальна структура, що складається з безлічі елементів одного типу. Ці елементи виконують прості функції, а всі процеси, що відбуваються в штучній нейронній мережі можна асоціювати з процесами, що відбуваються в нервовій системі живих організмів.

У першій частині даної роботи були основні поняття штучних інтелектуальних систем. У другій частині роботи був проведений огляд термінології. Третя частина роботи включає інформацію про основні поняття щодо вибору мережі. У четвертій частині була решена задача із розпізнавання образу. П'ята частина це експеримент у MATLAB У цій роботі були розглянуті

механізми побудови голографічного дисплею та його основні технічні характеристики та вивчені основні принципи та використання мікроконтролеру Arduino Uno в сфері відображення інформації.

ABSTRACT

Explanatory note of the master's qualifying work: 67 pages, 38 drawings, 5 table, 15 references

NEURAL NETWORK, ARTIFICIAL INTELLIGENCE, PATTERN RECOGNITION, MATLAB

The object of research is the method of pattern recognition.

The purpose of the work is to study the principle of the neural network in pattern recognition. The method is analytical

In the world, artificial intelligent systems have gained great popularity, thanks to the rich possibilities and efficiency of use. There are many tasks where practical permission using intelligent systems is needed already: economics and business, robotics, geological exploration, mathematics, biophysics, avionics, safety and security systems, medicine and many others. Such a large number of areas of application suggests that intelligent systems are a unique set for solving issues of analysis and processing of a large volume of data, solving tasks of various levels of complexity. Currently, artificial neural networks are actively used in artificial intelligence systems. A neural network is a computing structure consisting of many elements of the same type. These elements perform simple functions, and all processes occurring in an artificial neural network can be associated with processes occurring in the nervous system of living organisms.

In the first part of this work, there were basic concepts of artificial intelligent systems. In the second part of the work, a review of the terminology was carried out. The third part of the work includes information on the basic concepts of network selection. In the fourth part, the problem of image recognition was solved. The fifth part is an experiment in MATHCAD. In this work, the mechanisms of building a holographic display and its main technical characteristics were considered, and the basic principles and use of the Arduino Uno microcontroller in the field of information display were studied.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	8
ВСТУП	Ошибка! Закладка не определена.
1 ОСНОВНІ ПОНЯТТЯ ШТУЧНИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ І НЕЙРОННИХ МЕРЕЖ.....	10
1.1 Історія розвитку нейронних мереж	10
1.2 Поняття «штучна нейронна мережа»	12
1.3 Біологічний нейрон та штучний нейрон	12
1.4 Архітектура з'єднань штучних нейронів.....	17
2 ТЕРМІНОЛОГІЯ.....	22
2.1 Згорткові структури нейронних мереж.....	24
2.2 Згортковий шар	27
2.3 Локальна з'єднаність	28
2.4 Просторова організація.....	28
2.5 Спільне використання параметрів.....	29
3 ИНФОРМАЦИОННЫЕ МОДЕЛИ НЕЙРОННЫХ СЕТЕЙ И РАСПОЗНАВАНИЯ ЭМОЦИЙ	31
3.1 Односрямовані багатошарові мережі.....	31
3.2 Рекурентна нейронна мережа	35
3.3 Мережі Кохонена	37
3.4 Алгоритм Кохонена	39
3.5 Радіальні нейронні мережі	41
3.6 Інтенсивність дії руху лицьових м'язів	45
4 РОЗВ'ЯЗАННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ НЕЙРОННОЮ МЕРЕЖЕЮ	48
4.1 Методика розпізнавання рухової активності особи	49
5 ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА НЕЙРОННОЇ МЕРЕЖІ У ПРОГРАМНОМУ СЕРЕДОВИЩІ MATLAB	55
ВИСНОВКИ.....	64
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	66
ДОДАТОК А.....	Ошибка! Закладка не определена.
ДОДАТОК Б	Ошибка! Закладка не определена.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ

БШМ – Багатошарова нейронна мережа

ЗНМ – Згорткова нейронна мережа

ЗШ – Згортковий шар

НМ – Нейронна мережа

РНМ – Рекурентная нейронна мережа

РОО – Рухательні одиниці особи

ШІ – Штучний інтелект

ВСТУП

У світі штучні інтелектуальні системи завоювали велику популярність, завдяки багатим можливостям і ефективність використання. Існує безліч завдань, де необхідне практичне дозвіл із використанням інтелектуальних систем вже зараз: економіка та бізнес, робототехніка, математика, біофізика, авіоніка, безпека та охоронні системи, медицина та багато інші . Така велика кількість областей застосування, говорить про те, що інтелектуальні системи – це унікальний набір для вирішення питань аналізу та обробки великого обсягу даних, вирішення завдань різного рівня складності. Нині у системах штучного інтелекту активно використовуються штучні нейронні мережі.

Нейронна мережа – це обчислювальна структура, що складається з безлічі елементів одного типу. Ці елементи виконують прості функції, а всі процеси, що відбуваються в штучній нейронній мережі можна асоціювати з процесами, що відбуваються в нервовій системі живих організмів.

Нейронні мережі нелінійні за своєю природою, у них відсутня явна залежність, що дозволяє одночасно використовувати розроблену технологію (Інформаційну модель нейромережі). Протягом багатьох років лінійне моделювання було основним методом моделювання, оскільки для нього добре розроблено процедури оптимізації. Актуальність магістерської дисертації полягає у використанні нейронних мереж для вирішення погано формалізованих завдань у інтелектуальний аналіз даних. Зростання обсягів інформації, а також розширення кола технічно складних завдань ухвалення рішень вимагають систематизації існуючих методів та розробки нових методик та алгоритмів розв'язання. У магістерській дисертації розглядається можливість застосування нейронної мережі при вирішенні задачі розпізнавання емоцій людини.

1 ОСНОВНІ ПОНЯТТЯ ШТУЧНИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ І НЕЙРОННИХ МЕРЕЖ

1.1 Історія розвитку нейронних мереж

Вперше про штучні нейронні мережі заговорили у 1940-х роках. Саме як науковий напрямок теорія нейронних мереж була предствленна у роботі Мак Каллока і Пітса в 1943 году. В даній роботі стверджувалося, що майже будь-яку логічну чи арифметичну функцію можна реалізувати за допомогою найпростішої нейронної мережі. Однією з основних робіт слід відзначити модель Д. Хебба, який у 1949 році сформулював закон, що став відправною точкою для навчання нейронних мереж. Хебб першим припустив, що навчання, насамперед, полягає у зміні сили синаптичних зв'язків. Теорія Хебба - типовий випадок самонавчання, де випробувана система навчається виконувати необхідне завдання без втручання експериментатора. Свій внесок вніс і М. Мінський, завдяки дослідженню низки завдань, у тому числі відомої задачі «Виключне «АБО». Ф. Розенблатт у 1958 році запропонував нейронну мережу, названу перцептроном, призначеним для класифікації об'єктів. При Навчання перцептрон отримував повідомлення від «вчителя». Завдяки повідомленню, можна визначити до якого класу належить цей об'єкт. Крім наш того, навчений перцептрон був здатний самостійно класифікувати об'єкти, що не використовувалися раніше, здійснюючи при цьому мале число помилок. Період затишшя у розвитку нейронних мереж припав на 1968-1985 р. З появою високопродуктивних персональних комп'ютерів стало доступним моделювати нейронні мережі. Настільною книгою фахівців, які цікавляться теорією нейронних мереж, стала робота Ф. Уоссермена «Нейрокомп'ютерна техніка» Після появи роботи Д. Хопфілда (1982) інтерес до нейронних мереж різко зріс. Хопфілд, ґрунтуючись на правилах навчання Хебба, показав, що завдання з нейронами можуть бути зведені до узагальнення ряду моделей, розроблених на той час у фізиці неупорядкованих систем. Потім у 1986 році з'явилася робота Вільямса, Румельхарта та Хінтона, яка відповідала на важливе питання

про поведінку навчання шаруватих нейронних мереж. Після цього, запропонований Хінтоном алгоритм зазнав безліч змін.

У 80-ті роки поступово сформувався міцний теоретичний фундамент, на основі якого сьогодні створюється більшість мереж. Розроблена теорія широко стала застосовуватися останні два десятиліття на вирішення прикладних завдань. Стали з'являтися фірми, що займаються розробкою програмного забезпечення для конструювання штучних нейронних мереж. У 90-ті роки нейронні мережі стали використовуватись у бізнесі, де вони показали колосальну ефективність при вирішенні багатьох завдань - від передбачення попиту на продукцію до аналізу платоспроможності клієнтів банку.

У 2007 році в університеті Торонто Джеффри Хінтон створив алгоритми глибокого вивчення нейронних мереж. Під час навчання нижніх шарів мережі, Хінтон використав обмежену машину Больцмана, яка представляє стохастичну рекурентну нейронну мережу. Після навчання мережі отриманий додаток міг швидко вирішувати поставлену завдання (наприклад, пошук облич на фотографії). Ця функція на сьогоднішній день вбудована у всі цифрові фотоапарати. Подібна технологія використовується інтернет - пошуковими системами при класифікації картинок. За оцінками фахівців, у галузі проектування нейронних мереж та нейрокомп'ютерів очікується технологічне зростання. Чимало нових можливостей було відкрито останніми роками, а роботи у цій галузі стають важливим внеском у науку, технології, економіку.

Незважаючи на те, що вивчення нейронного моделювання ведеться вже більше шістдесяти років, немає жодної області мозку, де процес обробки інформації було б зрозумілим остаточно. Також немає жодного нейрона, для якого можна було б визначити код передачі у вигляді послідовності імпульсів. В даний час існує велика кількість конфігурацій нейронних мереж, які відрізняються за принципами функціонування, та, отже, спрямовані різні завдання. Майбутнє нейрокомп'ютерних технологій буде пов'язане з новими

відкриттями в галузі нейронного моделювання – як тільки вдасться розгадати таємницю функціонування хоча б однієї області мозку, одразу стане зрозуміло багато про інші його області.

1.2 Поняття «штучна нейронна мережа»

Під штучною нейронною мережею розуміється математична модель, а також її програмна та апаратна реалізація, побудована за принципу біологічних нейронних мереж – нервових клітин живого організму. Це поняття виникло при спробі змодельювати процеси, протікають у мозку людини. Штучна нейронна мережа є системою простих процесорів (штучних нейронів), з'єднаних та взаємодіючих між собою. Кожен із процесорів мережі має справу з сигналами, які періодично надходять чи передаються іншим процесорам. Велика мережа здатна вирішувати найскладніші завдання у найкоротші терміни. З математичної точки зору нейронні мережі є спосіб розв'язання нелінійних завдань оптимізації. Кібернетика використовує теорію нейронних мереж у вирішенні завдань адаптивного управління, побудову алгоритмів для робототехніки.

У програмуванні нейронна мережа один із способів вирішення проблеми ефективного паралелізму.

Програмування нейронних мереж має на увазі саме навчання мережі, а не написання програмного коду. Саме завдяки навчанню мережа здатна виявляти залежності між даними (вхідними та вихідними), узагальнювати, спрощувати результати, використовувати знання для розбиття складних задач на простіші.

1.3 Біологічний нейрон та штучний нейрон

Нейрон (нервова клітка) складається з тіла клітини - соми, і двох типів зовнішніх деревоподібних відгалужень: аксона і дендритів. Тіло клітини вміщує ядро (nucleus), що містить інформацію про властивості нейрона, і

плазму, яка продукує необхідні для нейрона матеріали. Нейрон отримує сигнали (імпульси) від інших нейронів через дендрити (приймачі) і передає сигнали, що згенеровані тілом клітки, вздовж аксона (передавач), який наприкінці розгалужується на волокна. На закінченнях волокон знаходяться синапси.

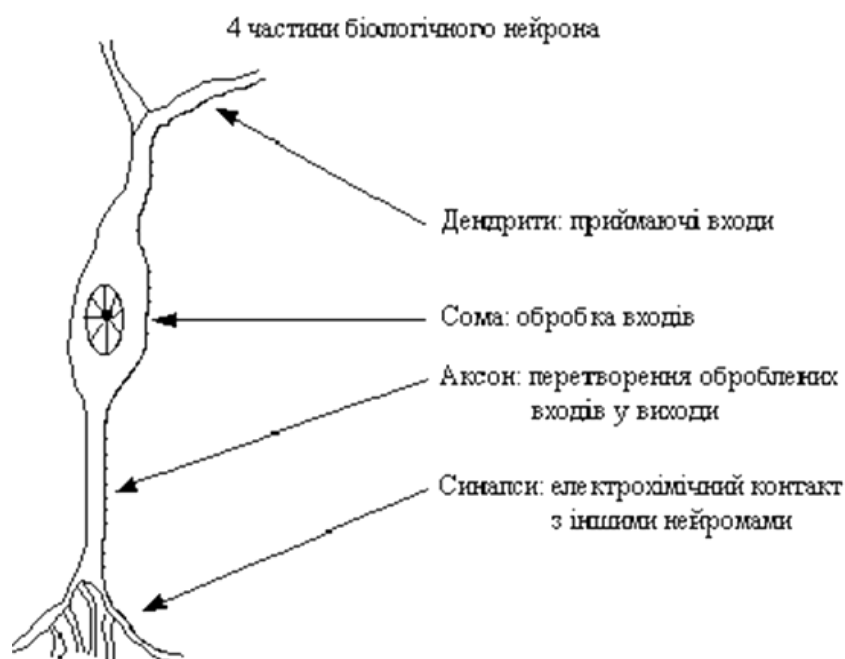


Рисунок 1.1 – Структура біологічного нейрону

Синапс є функціональним вузлом між двома нейронами (волокно аксона одного нейрона і дендрит іншого). Коли імпульс досягає синаптичного закінчення, там продукуються хімічні речовини - нейротрансмітери. Нейротрансмітери проходять через синаптичну щілину і в залежності від типу синапсу, збуджують або гальмують здатність нейрона-приймача генерувати електричні імпульси. Результативність синапсу налаштовується сигналами, які проходять скрізь нього, тому синапси навчаються у відповідності до активності процесів, у яких вони приймають участь. Нейрони взаємодіють за допомогою короткої серії імпульсів, повідомлення передається за допомогою частотно-імпульсної модуляції. Біологічні нейрони є структурно складнішими, ніж існуючі штучні нейрони. Оскільки нейрофізіологія надає науковцям розширене розуміння дії нейронів, а технологія обчислень постійно

вдосконалюється, розробники мереж мають необмежений простір для вдосконалення моделей біологічного мозку.

Штучний нейрон є базовим модулем нейронних мереж. Він моделює основні функції природного нейрона. При функціонуванні нейрон одночасно отримує багато вхідних сигналів. Кожен вхід має свою власну синаптичну вагу, яка надає входу вплив, необхідний для функції суматора елемента обробки. Ваги є мірою сили вхідних зв'язків і моделюють різноманітні синаптичні сили біологічних нейронів. Ваги суттєвого входу підсилюються і, навпаки, вага несуттєвого входу примусово зменшується, що визначає інтенсивність вхідного сигналу. Ваги можуть змінюватись відповідно до навчальних прикладів, топології мережі та навчальних правил.

Вхідні сигнали x_n зважені ваговими коефіцієнтами з'єднання w_n додаються, проходять через передатну функцію, генерують результат і виводяться.

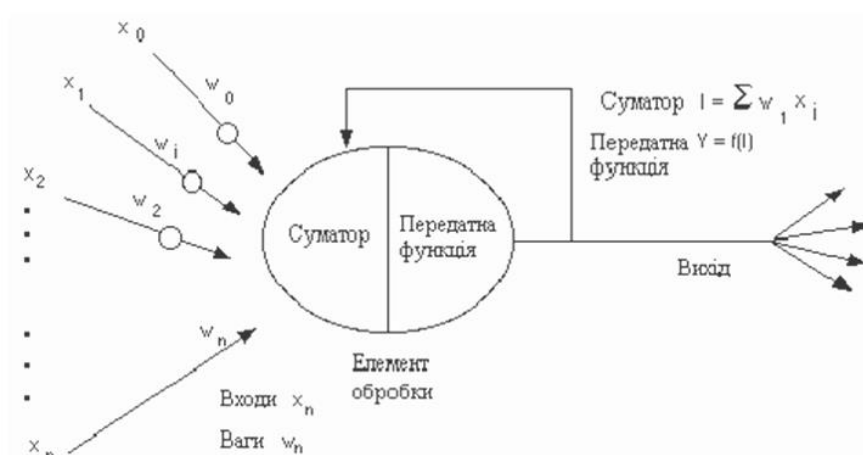


Рисунок 1.2 – Структура штучного нейрону

В програмних реалізаціях штучні нейрони називають «елементами обробки» або «процесорами» і вкладають в них більше можливостей, ніж в базовому штучному нейроні, що описаний вище.

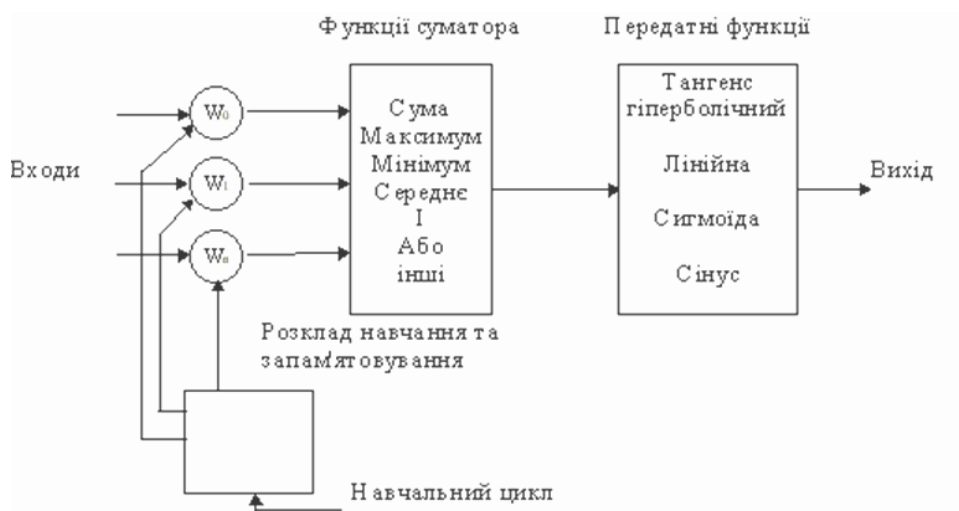


Рисунок 1.3 – Детальна структура штучного нейрону

Функція суматора може бути складнішою, наприклад, вибір мінімуму, максимуму, середнього арифметичного, добутку або обчислюватися за іншим алгоритмом.

Багато програмних реалізацій використовують власні функції суматора, що запрограмовані на мові вищого рівня (C, C++, TurboPascal). Перед надходженням до передатної функції вхідні сигнали та вагові коефіцієнти можуть комбінуватись багатьма способами. Алгоритми для комбінування входів нейронів визначають відповідно до мережної архітектури та парадигми. В деяких нейромережах суматор виконує додаткову обробку, так звану функцію активації, яка зміщує вихід функції суматора в часі. Цю функцію найкраще використовувати як компоненту мережі в цілому, ніж як компоненту окремого нейрона. Часто, ця функція є відсутньою. Результат функції суматора перетворюється у вихідний сигнал через передатну функцію. В передатній функції для визначення виходу нейрона загальна сума порівнюється з деяким порогом (зазвичай, це діапазон $[0, 1]$ або $[-1, 1]$ або інше) за допомогою певного алгоритму.

Переважно застосовують нелінійну передатну функцію, оскільки лінійні (прямолінійні) функції є обмеженими і вихід є пропорційним до входу.

Застосування лінійних передатних функцій було проблемою у ранніх моделях мереж, і їх обмеженість та недоцільність була доведена в книзі Мінські та Пейперта "Перцептрони".

В існуючих нейромережах як передатну функцію використовують сигмоїду, синус, гіперболічний тангенс тощо. На рис. 1.4 зображені типові передатні функції.

Жорстка порогова функція

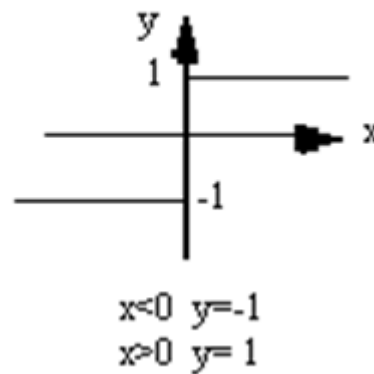


Рисунок 1.4 – Передаточна функція

Для простої передатної функції нейромережа може видавати 0 чи 1, 1 чи -1 або інші числові комбінації. Передатна функція в таких випадках є пороговою або «жорстким обмежувачем»

Лінійна з насиченням

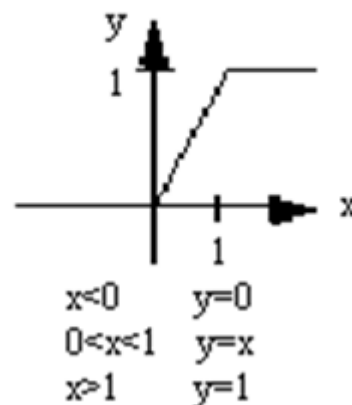


Рисунок 1.5 – Лінійна функція з насиченням

Передатна функція лінійна з насиченням віддзеркалює вхід всередині заданого діапазону і діє як жорсткий обмежувач поза межами цього діапазону. Це лінійна функція, яка відсікається до мінімальних та максимальних значень, роблячи її нелінійною.

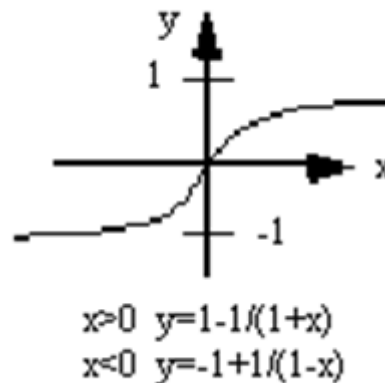


Рисунок 1.6 – Сигмоїда функція

Сигмоїда або S-подібна крива наближує мінімальне та максимальне значення у асимптотах. Вона називається сигмоїдою, коли її діапазон $[0, 1]$, або гіперболічним тангенсом, при діапазоні $[-1, 1]$. Важливою рисою сигмоїд є неперервність функцій та їх похідних. Застосування сигмоїдних функцій надає добрі результати і має широке застосування.

1.4 Архітектура з'єднань штучних нейронів

Штучні нейромережі конструюються з базового блоку - штучного нейрону. Іншою властивістю нейромереж є величезна кількість зв'язків, які пов'язують окремі нейрони. Групування нейронів у мозку людини забезпечує обробку інформації динамічним, інтерактивним та самоорганізуючим шляхом. Біологічні нейронні мережі з мікроскопічних компонентів існують у тривимірному просторі і здатні до різноманітних з'єднань. Але для реалізації

штучних мереж присутні фізичні обмеження. Об'єднуючись у мережі, штучні нейрони утворюють систему обробки інформації, яка забезпечує ефективну адаптацію моделі до постійних змін з боку зовнішнього середовища. В процесі функціонування мережі відбувається перетворення вхідного вектора сигналів у вихідний. Конкретний вид перетворення визначається архітектурою нейромережі, характеристиками нейронних елементів, засобами керування та синхронізації інформаційних потоків між нейронами.

Важливим фактором ефективності мережі є встановлення оптимальної кількості нейронів та типів зв'язків між ними. Для опису нейромереж використовують кілька усталених термінів, які в різних джерелах можуть мати різне трактування, зокрема структура нейромережі - спосіб зв'язків нейронів у нейромережі. Архітектура нейромережі - структура нейромережі та типи нейронів. Парадигма нейромережі - спосіб навчання та використання, іноді містить поняття архітектури. На базі однієї архітектури може бути реалізовано різні парадигми нейромережі і навпаки. Серед відомих архітектурних рішень виділяють групу слабозв'язаних нейронних мереж, у випадку, коли кожен нейрон мережі зв'язаний лише із сусідніми. В повнозв'язаних нейромережах входи кожного нейрона зв'язані з виходами всіх решти нейронів.

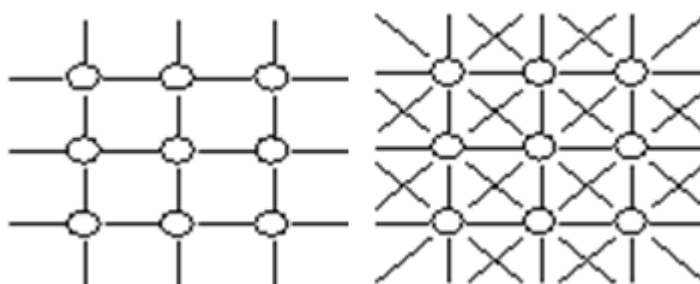


Рисунок 1.7 – Слабов'язна нейронна мережа

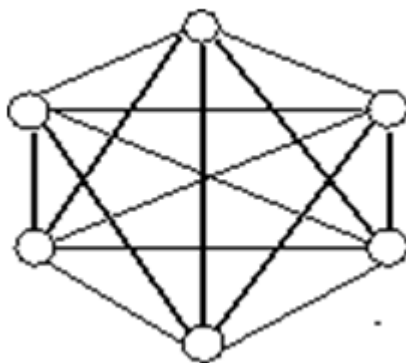


Рисунок 1.8 – Повнов'язна нейронна мережа

Самим поширеним варіантом архітектури є багат шарові мережі. Нейрони в даному випадку об'єднуються у прошарки з єдиним вектором вхідних сигналів. Зовнішній вхідний вектор подається на вхідний прошарок нейронної мережі (рецептори). Виходами нейронної мережі є вихідні сигнали останнього прошарку (ефектори). Окрім вхідного та вихідного прошарків, нейромережа має один або кілька прихованих прошарків нейронів, які не мають контактів із зовнішнім середовищем.

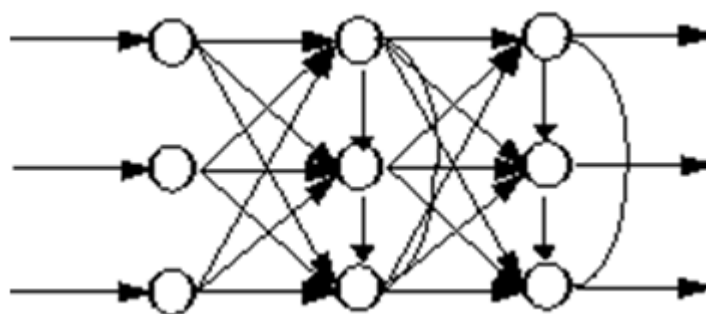


Рисунок 1.9 – Багат шаровий тип з'єднання нейронів

Зв'язки між нейронами різних прошарків називають проєктивними. Зв'язки між нейронами одного прошарку називають бічними (латеральними). На рисунку 2.3 показана типова структура штучних нейромереж. Хоча існують мережі, які містять лише один прошарок, або навіть один елемент, більшість

застосувань вимагають мережі, які містять як мінімум три типи прошарків - вхідний, прихований та вихідний. Прошарок вхідних нейронів отримує дані або з вхідних файлів, або безпосередньо з електронних датчиків. Вихідний прошарок пересилає інформацію безпосередньо до зовнішнього середовища, до вторинного комп'ютерного процесу, або до інших пристроїв. Між цими двома прошарками може бути багато прихованих прошарків, які містять багато нейронів в різноманітних зв'язаних структурах. Входи та виходи кожного з прихованих нейронів сполучені з іншими нейронами.

Важливим аспектом нейромереж є напрямок зв'язку від одного нейрону до іншого:

Зв'язки скеровані від вхідних прошарків до вихідних називаються аферентними. Зв'язки в зворотному напрямку називаються еферентними.

В більшості мереж кожен нейрон прихованого прошарку отримує сигнали від всіх нейронів попереднього прошарку чи від нейронів вхідного прошарку. Після виконання операцій над сигналами, нейрон передає свій вихід до всіх нейронів наступних прошарків, забезпечуючи передачу вперед (*feedforward*) на вихід. При зворотному зв'язку, вихід нейронів прошарку скеровується до нейронів попереднього прошарку

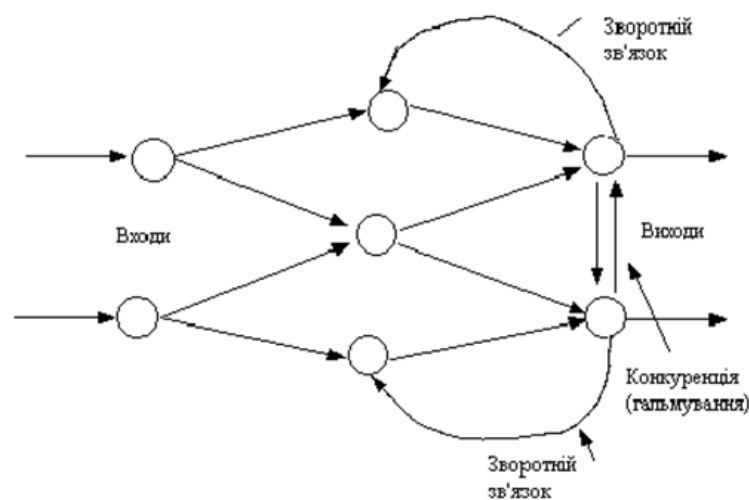


Рисунок 1.10 – Зворотній зв'язок нейронів

Напрямок зв'язків нейронів має значний вплив на роботу мережі. Більшість програмних нейромереж дозволяють користувачу додавати, вилучати та керувати з'єднаннями як завгодно. Корегуючи параметри, можна налаштувати зв'язки як на посилення так і на послаблення величини сигналів.

За архітектурою зв'язків, більшість відомих нейромереж можна згрупувати у два великих класи:

Мережі прямого поширення (з односкерованими послідовними зв'язками).

Мережі зворотного поширення (з рекурентними зв'язками).

2 ТЕРМІНОЛОГІЯ

Нейронна мережа, (НМ) – один із напрямків штучного інтелекту, мета якого змоделювати аналітичні механізми, що здійснюються людським мозком. Завдання, які вирішує типова нейронна мережа – класифікація, передбачення та розпізнавання. Нейронні мережі здатні самостійно навчатися і розвиватися, будуючи свій досвід на помилках. Нейронні мережі це послідовність обчислювальних одиниць - нейронів, з'єднаних між собою зв'язком далі - синапс. Структура нейронної мережі прийшла у світ програмування прямо з біології. Завдяки такій структурі машина знаходить здатність аналізувати і навіть запам'ятовувати різну інформацію. Також нейронні мережі здатні як аналізувати вхідну інформацію, а й відтворювати її з власної пам'яті. Іншими словами, нейронна мережа це машинна інтерпретація мозку людини, в якому знаходяться мільйони нейронів, що передають інформацію у вигляді електричних імпульсів.

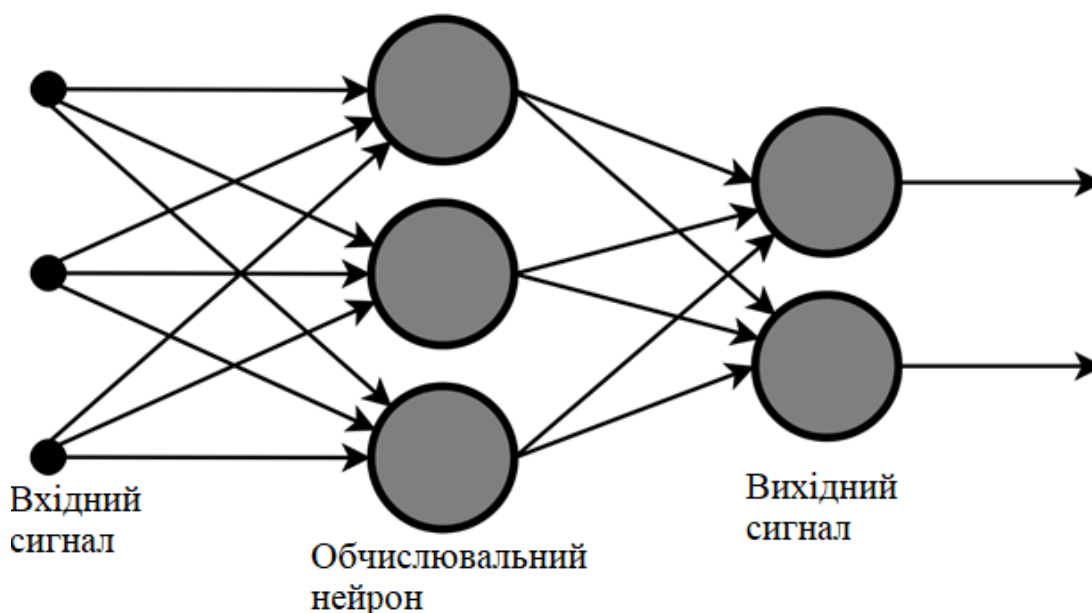


Рисунок 2.1 – Схема нейронної мережі

Нейрон – це обчислювальна одиниця, яка отримує інформацію, здійснює над нею

прості обчислення і передає її далі. Вони діляться на три основні типи: вхідний, прихований та вихідний: Також є нейрон зміщення та контекстний нейрон. У разі, коли нейронна мережа складається з великої кількості нейронів, вводять термін шару. Відповідно, є вхідний шар, який отримує інформацію, n прихованих шарів, які її обробляють та вихідний шар, який виводить результат. У кожного з нейронів є 2 основні параметри: вхідні дані - (input data) та вихідні дані (output data). Коли вихідний нейрон i передає сигнал вхідному нейрону j , в вихідний нейрон стає попередником, а вхідний нейрон стає наступником, цей зв'язок образує з'єднання та ваги ω_{ij} між обчислювальними одиницями. За період розвитку нейронні мережі поділилися на безліч типів, які переплітаються між собою в різних завданнях. На даний момент складно класифікувати будь-яку мережу лише за однією ознакою. Це можна зробити за принципом застосування, типом вхідної інформації, характером навчання, характером зв'язків, сферою застосування.

Таблиця 1.1 – Види та класифікації нейронних мереж

Нейронна мережа	Принцип застосування	Навчання з учителем + або – або змішане	Сфера Застосування
Перцептрон Розенблатта	Розпізнавання образів, прийняття рішень	+	Практично будь-яка окрім оптимізації
Хопфілда	Стиснення даних, асоціативна пам'ять	-	Будова комп'ютерних систем
Кохонена	Кластеризація, оптимізація, аналіз даних	-	Бази даних
Радіально- Базових функцій	Прийняття рішень, управління, прогнозування	з	Управлінські структури
Згорткова	Розпізнавання образів	+	Обробка графічних даних
Імпульсна	Прийняття рішення, розпізнавання образів	з	Протезування, робототехніка

2.1 Згорткові структури нейронних мереж

Один з найпопулярніших типів мережі, який часто використовується для розпізнавання тієї чи іншої інформації у фотографіях та відео, обробці мови, системах для рекомендацій.

Основні характеристики :

- відмінна масштабованість – проводять розпізнавання образів будь-якого дозволу (яке б воно не було велике).
- використання об'ємних тривимірних нейронів – усередині шару, нейрони пов'язані малим полем, іменовані рецептивним шаром.

Механізм просторової локалізації – сусідні шари нейронів пов'язані таким механізмом, за рахунок чого забезпечується робота нелінійних фільтрів та охоплення дедалі більшої кількості пікселів графічного зображення. Ідея складної системи цього типу нейронні мережі виникла при ретельному вивченні зорової кори, яка у великих півкулях мозку відповідає за обробку візуальної складової. Основний критерій вибору користь згорткового типу – вона у складі технологій глибокого навчання. Подібний тип з перцептроном, але різниця в тому, що тут використовується обмежена матриця ваг, що зсувається по шару, що обробляється, замість повнозв'язкової нейронної мережі.

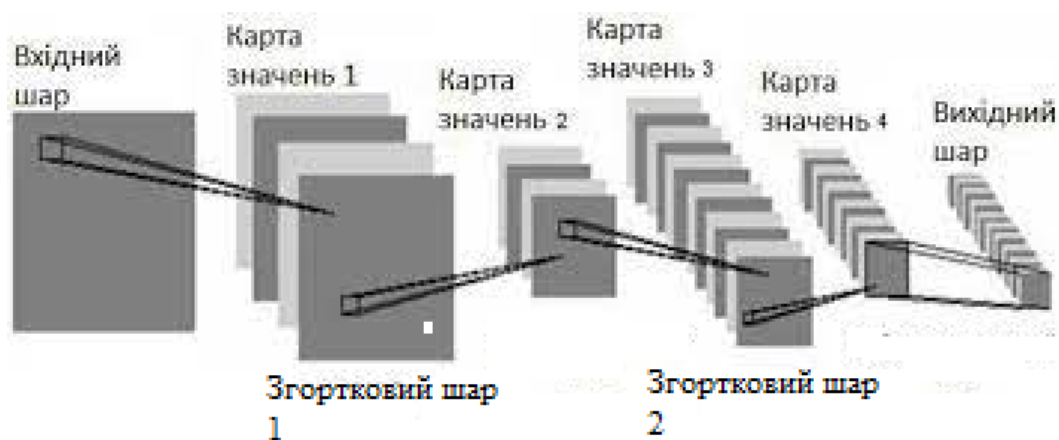


Рисунок 2.2 – Структурна схема ЗНМ

Згорткова нейронна мережа побудована за допомогою універсальних математичних операцій. Існують різні архітектури даного виду мережі, але кожна складається з 3 основних компонентів: згортковий шар, субдескриптивований шар, повнозв'язний шар. Структура мережі – одно направлена, без зворотних зв'язків, багато шарова.

Повнозв'язний шар - шар в якому кожен нейрон з'єднаний з усіма нейронами на попередньому рівні, причому кожен зв'язок має свій ваговий коефіцієнт.

Згортковий шар на відміну від повнозв'язну, з'єднаний лише з обмеженою кількістю нейронів попереднього рівня, тобто згортковий шар аналогічний застосуванню операції згортки, де використовується лише матриця ваг не великого розміру (ядро згортки), яку «рухають» по всьому оброблюваному шару. Ще одна особливість згорткового шару в тому, що він трохи зменшує зображення за рахунок крайових ефектів.

Субдескриптивований шар виконує зменшення розмірності даних (зазвичай в кілька разів). Це можна робити різними способами, але найчастіше використовується метод вибору максимального елемента - вся карта ознак поділяється на осередки, з яких вибираються максимальні за значенням. Даний тип мережі має такі переваги: невелика кількість вагів які потрібно налаштовувати, стійкість до спотворень, універсальність в розпізнаванні об'єктів.

Серед недоліків можна виділити наступне: містить занадто багато змінних параметрів багато параметрів які істотно впливають на результат, але вибираються емпірично.

Для оптимізації роботи згорткової нейронної мережі пропонується використовувати класичний генетичний алгоритм для навчання ваг в фільтрах мережі, що дозволяє уникнути зациклення в локальних мінімумах і підвищить середню ймовірність успішного вирішення завдання розпізнавання.

Згорткові нейронні мережі дозволяють комп'ютерного зору працювати як з простими завданнями, так і зі складними продуктами і послугами,

починаючи від розпізнавання осіб і закінчуючи поліпшенням медичних діагнозів Ці мережі є біологічно натхненими варіантами багатошарових перцептронів, розробленими для імітації поведінки зорової кори. Ці моделі пом'якшують виклики, поставлені архітектурою БШП, використовуючи сильну просторово локальну кореляцію, присутню в природних зображеннях. На противагу до БШП, ЗНМ мають наступні відмітні ознаки: Тривимірні ємності нейронів. Шари ЗНМ мають нейрони, впорядковані в 3 вимірах: ширина, висота та глибина. Нейрони всередині шару є з'єднаними лише з невеликою областю попереднього шару, що називається рецептивним полем. Для формування архітектури ЗНМ складають різні типи шарів, як локально-, так і повноз'єднані. Локальна з'єднаність: відповідно до концепції рецептивних полів, ЗНМ використовують просторову локальність шляхом застосування схеми локальної з'єднаності між нейронами сусідніх шарів. Ця архітектура таким чином забезпечує, що навчені «фільтри» виробляють найсильніший відгук до просторово локального вхідного образу. Складання багатьох таких шарів веде до нелінійних фільтрів, що стають все глобальнішими (тобто, чутливими до більшої області піксельного простору), так що мережа спочатку створює представлення дрібних деталей входу, а потім з них збирає представлення більших областей. Спільні ваги: В ЗНМ кожен фільтр повторюється на всьому зоровому полі. Ці повторні вузли використовують спільну параметризацію (вектор ваги та упередженості) та формують карту ознаки. Це означає, що всі нейрони в заданому згортковому шарі реагують на одну й ту ж саму ознаку в межах свого рецептивного поля. Повторювання вузлів таким чином дозволяє ознакам бути виявленими незалежно від їхнього положення в зоровому полі, забезпечуючи таким чином властивість інваріантності відносно зсуву. Разом ці властивості дозволяють ЗНМ досягати кращого узагальнення на задачах бачення. Спільне використання ваг різко зменшує кількість вільних параметрів, яких вчиться мережа, знижуючи таким чином вимоги до пам'яті для роботи мережі та уможливлюючи тренування більших, потужніших мереж.

2.2 Згортковий шар

Згортковий шар є основним будівельним блоком ЗНМ. Параметри шару складаються з набору фільтрів для навчання (або ядер), які мають невеличке рецептивне поле, але простягаються на всю глибину вхідної ємності. Протягом прямого проходу кожен фільтр здійснює згортку за шириною та висотою вхідної ємності, обчислюючи скалярний добуток даних фільтру та входу, і формуючи 2-вимірну карту збудження цього фільтру. В результаті мережа навчається, які фільтри активуються, коли вона виявляє певний конкретний тип ознаки у певному просторовому положенні у вході. Складання карт збудження всіх фільтрів уздовж виміру глибини формує повну ємність виходу згорткового шару. Таким чином, кожен запис в ємності виходу може також трактуватися як вихід нейрону, що дивиться на невеличку область у вході, та має спільні параметри з нейронами тієї ж карти збудження.

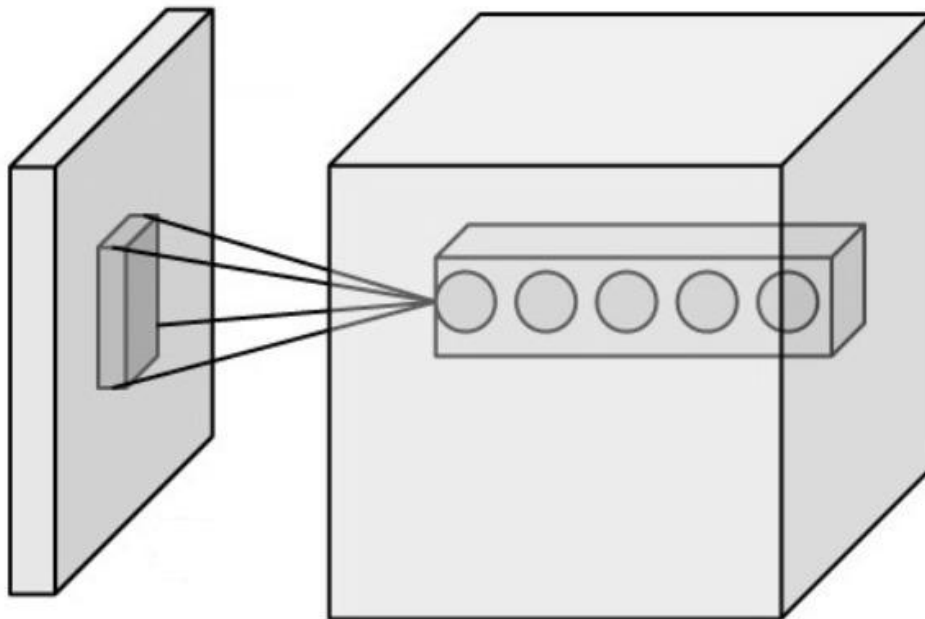


Рисунок 2.3 – Структура згорткового шару ЗНМ

2.3 Локальна з'єднаність

При опрацюванні входів високої розмірності, таких як зображення, недоцільно з'єднувати нейрони з усіма нейронами попередньої ємності, оскільки така архітектура мережі не бере до уваги просторову структуру даних. Згорткові мережі використовують просторово локальну кореляцію шляхом забезпечення схеми локальної з'єднаності між нейронами сусідніх шарів: кожен нейрон з'єднано лише з невеликою областю вхідної ємності.[4] Обшир цієї з'єднаності є гіперпараметром, що називається рецептивним полем нейрону. З'єднання є локальними в просторі (вздовж ширини та висоти), але завжди поширюються вздовж усієї глибини вхідної ємності. Така архітектура забезпечує, щоби навчені фільтри виробляли найсильніший відгук до просторово локальних вхідних образів.

2.4 Просторова організація

Розмір ємності виходу згорткового шару контролюють три гіперпараметри: глибина, крок та нульове доповнення. Глибина ємності виходу контролює кількість нейронів шару, що з'єднуються з однією й тією ж областю вхідної ємності. Ці нейрони вчаться активуватися для різних ознак входу. Наприклад, якщо перший згортковий шар бере як вхід сире зображення, то різні нейрони вздовж виміру глибини можуть активуватися в присутності різних орієнтованих контурів, або плям кольору. Крок контролює те, як стовпчики глибини розподіляються за просторовими вимірами (шириною та висотою). Коли кроком є 1, ми рухаємо фільтри на один піксель за раз. Це веде до сильного перекриття рецептивних полів між стовпчиками, а також до великих ємностей виходу. Коли ми робимо крок 2 (або, рідше, 3 чи більше), то фільтри, просуваючись, перестрибують на 2 пікселі за раз. Рецептивні поля перекриваються менше, й отримувана в результаті ємність виходу має менші просторові розміри.

Іноді зручно доповнювати вхід нулями по краях вхідної ємності. Розмір цього доповнення є третім гіперпараметром. Доповнення забезпечує контроль над просторовим розміром ємності виходу. Зокрема, іноді бажано точно зберігати просторовий розмір вхідної ємності.

Просторовий розмір ємності виходу може обчислюватися як функція від розміру вхідної ємності W , розміру ядрового поля нейронів згорткового шару K , кроку, з яким вони застосовуються S , і величини нульового доповнення P , що застосовується на краях. Формула для обчислення того, скільки нейронів «уміщається» до заданої ємності, задається як $\frac{(W-K+2P)}{S+1}$. Якщо це число не є цілим, то кроки встановлено неправильно, і нейрони не може бути розміщено вздовж вхідної ємності симетричним чином. Загалом, встановлення нульового доповнення в $P = \frac{(K-1)}{2}$, коли кроком є $S=1$, забезпечує, щоби ємності входу та виходу мали однаковий просторовий розмір. Хоча взагалі використання всіх нейронів попереднього шару не є абсолютно обов'язковим, наприклад, ви можете вирішити використовувати лише частину доповнення.

2.5 Спільне використання параметрів

Схема спільного використання параметрів застосовується в згорткових шарах для регулювання кількості вільних параметрів. Вона спирається на одне розумне припущення: якщо клаптикова ознака є корисною для обчислення в певному просторовому положенні, то вона також повинна бути корисною для обчислення й в інших положеннях. Іншими словами, позначаючи 2-вимірний зріз за глибиною як зріз глибини, ми обмежуємо нейрони в кожному зрізі глибини використанням одних і тих же ваг та упередженості.

Оскільки всі нейрони в одному зрізі поділяють спільну параметризацію, то прямий прохід у кожному зрізі глибини згорткового шару може бути обчислено як згортку ваг нейронів із вхідною ємністю (звідси й назва: згортковий шар). Таким чином, є звичним називати набори ваг фільтром (або

ядром), який згортається із входом. Результатом цієї згортки є карта збудження, і набір карт збудження для кожного з різних фільтрів складають до купи вздовж виміру глибини для отримання ємності виходу. Спільне використання параметрів сприяє інваріантності архітектури ЗНМ відносно зсуву. Іноді спільне використання параметрів може й не мати сенсу. Особливо в тому разі, коли вхідні зображення до ЗНМ мають певну особливу центровану структуру, в якій ми очікуємо зовсім різних ознак для навчання в різних просторових положеннях. Одним із практичних прикладів є коли вхід є обличчями, що було відцентровано в зображенні: ми можемо очікувати, що вчитимемося різних особливих ознак очей та волосся в різних частинах зображення. В такому разі є звичним пом'якшувати схему спільного використання параметрів, і натомість просто називати шар локально з'єднаним.

3 ИНФОРМАЦИОННЫЕ МОДЕЛИ НЕЙРОННЫХ СЕТЕЙ И РАСПОЗНАВАНИЯ ЭМОЦИЙ

3.1 Односпрямовані багат шарові мережі

Багат шарова нейронна мережа може бути формально визначена як сукупність простих обробних елементів, які називаються нейронами. Нейрони організовані за шарами та об'єднані односпрямованими зв'язками, які називаються синапсами. Зазвичай мережа складається з безлічі сенсорних елементів (вхідних вузлів), які утворюють вхідний шар; одного або кількох прихованих шарів обчислювальних нейронів та одного вихідного шару нейронів. Односпрямовані багат шарові нейронні мережі найбільш поширені, оскільки вони мають простий математичний опис. З кінця 70-х років багат шарові мережі отримали свій розвиток, завдяки зручним алгоритмам навчання. Нейронні мережі прямого поширення мають вхідний сигнал, який передається від шару до шару (від одних нейронів до інших). Саме такі мережі і названі багат шаровим перцептроном, що складається з вхідного шару, прихованих обчислювальних шарів усередині системи та вихідного шару нейронів.[5]

Багат шаровий перцептрон – це односпрямована мережа сигмоїдального типу. Сигмоїдальна функція - монотонно зростаюча скрізь диференційована S-подібна нелінійна функція з насиченням, яку дуже зручно використовувати у формальному нейроні як функцію активації. Сигмоїд дозволяє посилювати слабкі сигнали та не насичуватися від сильні сигнали.

Структурна схема багат шарового перцептрона (у разі двошаровий перцептрон) виглядає так (рис. 3.1)

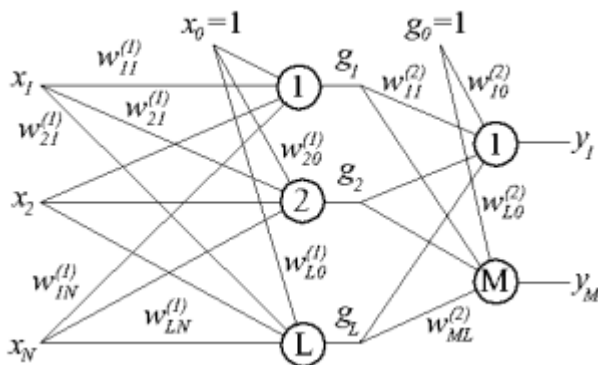


Рисунок – 3.1 Структурна схема багат шарового персеプトрону

де , g_l , $l=1,2,\dots,L$ – сигнали, що виходять із першого шару нейронів;

(m), $m=1,2,\dots$ - індекси у дужках угорі означають номер шару нейрона;

x_1, x_2, \dots, x_N – вхідні сигнали, які формують вхідний шар;

g_1, g_2, \dots, g_L – перший прихований шар;

y_1, y_2, \dots, y_M - вихідні сигнали, які формують шар на вихід.

Формули для розрахунку вихідних сигналів шарів нейронів:

$$g_l = \left(\sum_{j=0}^N w_{lj}^{(1)} \cdot x_j \right), l = 1, 2, \dots, L, \quad (3.1)$$

$$y_i = f \left(\sum_{l=0}^L w_{il}^{(2)} \cdot f \left(\sum_{l=0}^L w_{il}^{(2)} \cdot f \left(\sum_{j=0}^N w_{il}^{(1)} \cdot x_j \right) \right) \right), i = 1, 2, \dots, M. \quad (3.2)$$

Цільова функція для одичної навчальної вибірки $\langle X, D \rangle$ виглядає наступним чином:

$$E(W) = \frac{1}{2} \sum_{i=1}^M (y_i - d_i)^2, \quad (3.3)$$

Цільова функція для безлічі навчальних пар $\langle X^k, D^k \rangle$, $k=1,2,\dots,p$, має вигляд суми з усіх парам:

$$E(W) = \frac{1}{2} \sum_{k=1}^p \sum_{i=1}^M (y_i - d_i)^2. \quad (3.4)$$

Мета навчання багат шарового персептрону: при вхідному векторі X^k необхідно підібрати такі значення ваги мережі $W_{ij}^{(1)}$ та $W_{il}^{(2)}$, щоб вихідний вектор Y^k та вектор очікуваних значень D^k максимально збігалися.

Навчання багат шарового персептрона методом зворотного поширення помилки

Одним із методів навчання персептрону є алгоритм зворотного поширення помилки. Передбачається, що вектор матиме два проходу шарами мережі – один прямий, а другий зворотний. В першому випадку вектор починає рух із вхідного шару, далі рухається по мережі від шару до шару. При цьому всередині мережі генерується набір вихідних сигналів.

Під час зворотного проходу синаптичні ваги налаштовуються таким чином, що з бажаного виходу мережі віднімається фактичний, після чого надходить сигнал помилки. Далі сигнал у протилежному напрямку синаптичним зв'язкам поширюється по мережі. Саме тому даний метод отримав назву зворотного поширення помилки.

Суть даного методу розглянемо на прикладі двошарового персептрону. Максимально спростити приклад допоможе нагода одиначної навчальної вибірки, цільова функція буде мати вигляд згідно формулою (3.3). Для уточнення коефіцієнтів необхідна формула методу градієнта:

$$W^{k+1} = w^k - \eta \cdot \text{grad}E(w^k). \quad (3.5)$$

У реалізації цієї формули основну складність представляє розрахунок компонентів градієнта цільової функції:

$$E(W) = \frac{1}{2} \sum_{i=1}^M (f(\sum_{l=0}^L w_{il}^{(2)} \cdot g_l) - d_i)^2 = \frac{1}{2} \sum_{i=1}^M (f(\sum_{l=0}^L w_{il}^{(2)} \cdot f(\sum_{j=0}^N w_{il}^{(1)} x_j)) - d_i)^2. \quad (3.6)$$

Похідні цільової функції за вагами нейронів вихідного шару:

$$\frac{\partial E(W)}{\partial w_{il}^{(2)}} = (y_i - d_i) \cdot \frac{\partial f(u_i^{(2)})}{\partial u_i^{(2)}} \cdot g_l, \text{ де } u_i^{(2)} = \sum_{l=0}^L w_{il}^{(2)} \cdot g_l. \quad (3.7)$$

Якщо:

$$\delta_i^{(2)} = (y_i - d_i) \cdot \frac{\partial f(u_i^{(2)})}{\partial u_i^{(2)}}.$$

Тоді компонент векторного градієнту має вигляд:

$$\frac{\partial E(W)}{\partial w_{il}^{(2)}} = \delta_i^{(2)} \cdot g_l. \quad (3.8)$$

Компоненти для ваг передостаннього шару:

$$\frac{\partial E(W)}{\partial w_{il}^{(1)}} = \delta_i^{(1)} \cdot x_j. \quad (3.9)$$

Аналіз формул дозволяє сформулювати правило для розрахунку вектора градієнта $E(W)$ для похідних цільової функції для ваг кожному шарі:

$$\frac{\partial E(W)}{\partial w_{il}^{(k)}} = \delta_i^{(k)} \cdot x_j. \quad (3.10)$$

де x_j – сигнал на вході;

$\delta_i^{(k)}$ – похибка навчання.

Завдяки перенесенню похибки ($y_i - d_i$) з виходу мережі до попередніх шарів, метод отримав назву «зворотне поширення помилки». Модель багатошарової мережі сигмої дального типу та метод її навчання, незважаючи на свою відносну простоту та зручність, має ряд складнощів, які уповільнюють процес навчання, або зовсім не дозволяють мережі навчитися. Наприклад, значення ваги в результаті корекції можуть стати дуже великими величинами. Така ситуація призведе до того, що більшість нейронів на виході видасть величезні значення, але похідна у цих значеннях буде дуже мала. Процес може зупинитися, оскільки помилка, що посиляється назад в процесі навчання, пропорційна цій похідній. Крім того, необхідно вибрати такий розмір кроку, щоб він був кінцевим. У цьому питанні можна спиратися лише на досвід. Даний алгоритм має доказ збіжності. Розмір кроку безпосередньо буде впливати на збіжність: або вона буде занадто повільна, якщо розмір кроку дуже малий, або буде зазнавати постійної нестійкості через надто великого розміру кроку. Ще однією складністю є тимчасова нестійкість. Мережа може розпізнавати літери, але якщо, вивчивши одну, вона забуває іншу, то навчання не матиме сенсу. Процес навчання повинен проводитися так, щоб мережа була навчена на всій множині, без втрат тієї інформації, яка вже вивчена.

3.2 Рекурентна нейронна мережа

Рекурентна нейронна мережа найскладніший вид нейронних мереж, має зворотний зв'язок між елементами: від більш віддаленого до менш віддаленому. Саме зворотний зв'язок дозволяє мережі запам'ятовувати та навчатися. Однак потенціал такої мережі погано вивчений через складність аналізу. Рекурентна штучна нейронна мережа - одержувана з односпрямованого багатошарового персептрона мережу шляхом введення зворотних зв'язків із затримкою із виходу на входи мережі. Розглянемо один з різновидів таких мереж - це рекурентна мережа Елмана, що складається із двох шарів. Саме така структура дозволяє накопичувати та запам'ятовувати

інформацію. Сигнал надходить на вхід і проходить у прихований шар. Після перетворення прихованим шаром сигнал піде на вихід, а його копія потрапить на затримку. Потім до мережі надходить наступний сигнал, і в цей же час надходить копія попереднього сигналу.

Тут $v_l, l = 1, 2, \dots, L$ – вихідні сигнали для першого шару;

Стандартний вид цільової функції:

$$E^k(W) = \frac{1}{2} \sum_{i=0}^M (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{i=1}^M (e_i^k)^2, \quad (3.11)$$

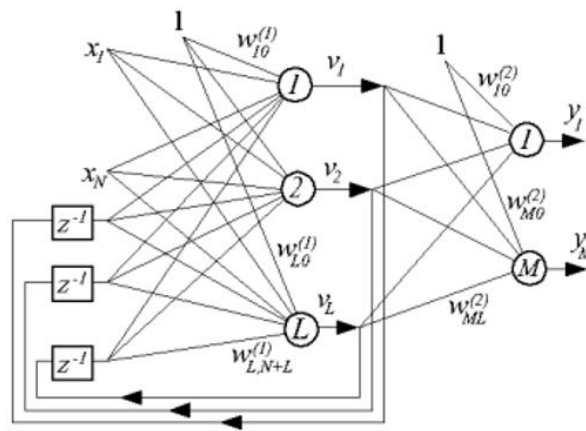


Рисунок 3.2 - Рекурентна нейронна мережа

Формула Елмана виглядає так:

$$E^k(W) = \frac{1}{2} \sum_{i=0}^M (y_i^k - d_i^k)^2 = \frac{1}{2} \sum_{i=1}^M (e_i^k)^2, \quad (3.12)$$

$$\frac{\partial E(W)}{\partial w_{it}^{(k)}} = e_n^{(k)} \cdot \frac{\partial f_2(g_n^k)}{\partial g_{it}^k} \cdot v_m^k, \quad (3.12)$$

Навчання рекурентної мережі Елмана можна задати наступним алгоритмом:

- визначити початкові значення ваги, присвоїти $k=1$;
- визначити всі поточні сигнали мережі для поточного моменту k ;
- для всіх терезів нейронів першого шару розрахувати значення;
- розрахувати усі компоненти вектора цільової функції;

- для всіх шарів скоригувати ваги нейронів;
- збільшити k на 1, перейти до пункту 2.

Мережі Елмана переважно застосовуються в системах управління рухомими об'єктами. Вони не забезпечують високу точність рішення, оскільки наявність зворотного зв'язку у прихованому шарі не дозволяє точно обчислити градієнт функціоналу. Однак такі мережі здатні обробляти послідовності образів, враховуючи зв'язок між елементами послідовності; мають високу здатність до узагальнення, яка набагато вище, ніж у багат шарового перцептронну. Рекурентні мережі застосовуються в задачах обробки звуку та мовлення, аналізі електронних схем, завдання комп'ютерного зору та обробки сигналів.

3.3 Мережі Кохонена

Для вирішення різних завдань класифікації використовується окремий клас нейронних мереж, названий мережами Кохонена. Основне завдання класифікації полягає у поділі об'єктів на класи, за якими – або ознак. Основою такого розподілу є вектор параметрів об'єкта. Такі мережі ще називають мережами, що самоорганізуються, на основі конкуренції. Нейронні мережі не вимагають «вчителя», а діють за принципом "переможець отримує все". З математичної точки зору це означає, що є двошарова мережа, в якій кожен нейрон має всі компонентами вхідного вектора X розмірністю N . На виході максимальний сигнал стане одиничним, решта звернеться в нуль.

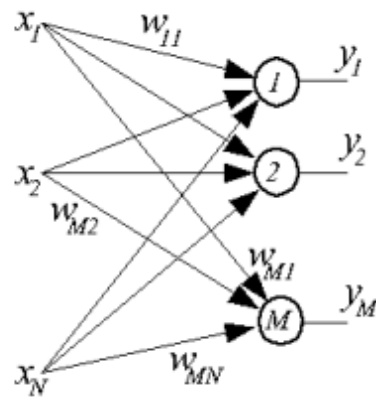


Рисунок 3.3 – Мережа Кохонена

Мережа побудована на основі нейронів типу WTA (Winner Takes All), які виглядають так (рисунок 3.4)

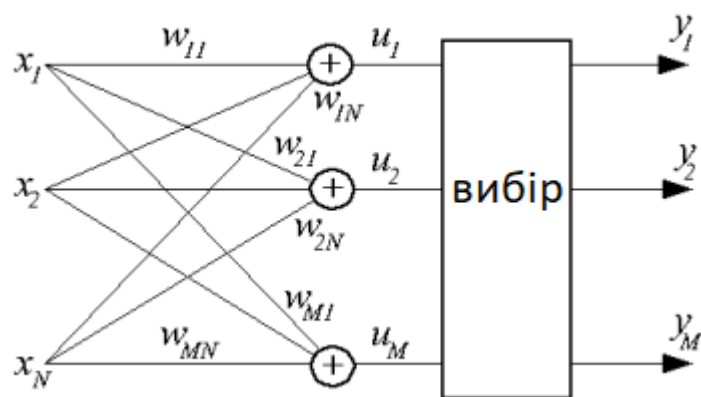


Рисунок 3.4 – Нейронна мережа «Переможець отримує все»

Вхідні сигнали всім конкуруючих нейронів однакові, вихідні сигнали вважаються за формулою:

$$u_i = \sum_{j=0}^N w_{ij} \cdot x_j, \quad (3.13)$$

Після порівняння результатів, вибирається нейрон, що переміг (найбільше значення u_i). На виході сигнал y_i отримує значення 1, у разі перемоги, всі інші сигнали нейронів набувають значення 0. Ваги вхідних зв'язків утворюють вектор $W_i = [w_{i1}, w_{i2}, \dots, w_{iN}]^T$. Дуже часто потрібна

нормалізація значень вхідного вектора, адже крім зв'язків, представлених у схемі, мають місце зв'язки, які вказують на ступінь "сусідства" між нейронами. Процес навчання виглядає так: X_k – навчальний вектор, що подається на вхід мережі; $d(X_k, W_i)$ – відстань між зазначеними векторами, визначається для кожного нейрона. Переможцем стане той нейрон, для якого це відстань виявиться найменшою. Навколо нейрона, що переміг, утворюється Sw_k – околиця із сусідніх нейронів з відомою «відстанню» до переможця. Ваги переможця та його сусідів з околиці уточнюються за правила Кохонена:

$$W_i^{k+1} = W_i^k + \mu_i^k \cdot (X^k - W_i^k), \quad (3.14)$$

де- μ_i^k - коефіцієнт навчання.

Зі збільшенням відстані від i -ого нейрона до переможця, значення коефіцієнта зменшуватиметься. Ваги нейронів поза околицями Sw_k не змінюватимуться. Розмір околиці та коефіцієнт навчання зменшуються з часом.

3.4 Алгоритм Кохонена

У нейронних мережах, запропонованих Т. Кохоненом у 1982 році, нейрони утворюють одномірний ланцюжок, при цьому кожен із нейронів має сусідів. ліворуч і праворуч. У більш складному випадку нейрони можуть утворити так звану сітку, де кожен нейрон матиме по одному сусіду праворуч, ліворуч, зверху та знизу. Ще один випадок, це коли кожен нейрон може мати шість сусідів на площині (як циферблат годинника 2,4,6,8,10,12), то є сітка гексагональна. За алгоритмом Кохонена налаштовується корекція терезів і виражається формулою:

$$W_i^{k+1} = W_i^k + \mu_i^k \cdot G^k(i, X^k) \cdot (X^k - W_i^k), \quad (3.15)$$

де $G_k(i, X_k)$ – функція «сусідства», що визначається формулою Гауса:

$$G^k(i, X_k) = \exp\left(-\frac{d^2(i, X_k)}{2 \cdot (\sigma^k)^2}\right), \quad (3.16)$$

де $d(i, X_k)$ – відстань між переможцем та i -им нейроном у k -му циклі навчання;

u_i та (σ^k) – коефіцієнт ширини Гауса, зменшуються у процесі навчання.

В результаті навчання сусідні нейрони стають представниками навчальних даних. Гідність мереж Кохонена полягає наочності подання даних шляхом одновимірної або двовимірної візуалізації. Мережі Кохонена також називають картами Кохонена.

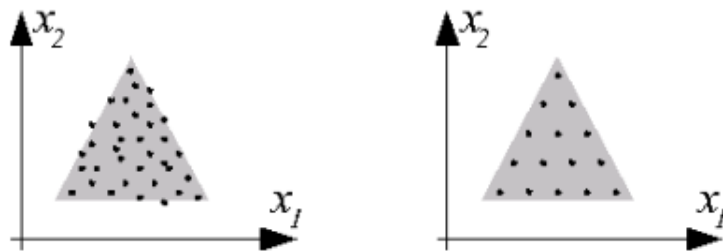


Рисунок 3.5 - Самоорганізація нейронної мережі на основі конкуренції

На рисунку 3.5 представлений приклад навчання мережі з самоорганізацією на основі конкуренції. Використовується мережа з 15 нейронами та двокомпонентним вхідним вектором $X=[x_1, x_2]^T$. На лівому рисунку представлені дані у навчальній вибірці, на правому розподілі нейронів у навченій мережі. Завдання зводиться до такого розбиття простору на області, при якому мінімізується середньоквадратична помилка.

Нейронні мережі Кохонена – це двошарові мережі, що складаються з вхідного шару нейронів та шару активних нейронів (шар Кохонена), який може бути одновимірним, тривимірним та двовимірним.

Визначення вагів нейронів для шару Кохонена засноване на

застосування алгоритмів самонавчання чи кластеризації. Як і будь-яка модель, мережа Кохонена також має ряд складнощів при реалізації:

- мережі Кохонена працюють лише з числовими даними,
- спочатку необхідно самостійно задавати кількість кластерів, якщо використовується метод кластеризації,
- необхідна мінімізація розмірів мережі,
- не завжди оптимальне рішення може бути знайдено.

3.5 Радіальні нейронні мережі

Мережа радіальних базисних функцій - нейронна мережа прямого поширення сигналу, що містить проміжний (прихований) шар 41 симетричні нейрони. Такий нейрон перетворює відстань від даного вхідного вектора до відповідного йому "центру" нелінійного закону. Радіальні нейронні мережі будуються за допомогою радіальних нейронів. На околицях свого центру функція активації цих нейронів має ненульові значення. Це називається локальною апроксимацією. Апроксимація - науковий метод, який полягає у заміні одних об'єктів іншими близькими до початкових, але спрощених у сенсі. З математичної точки зору багатошарові мережі сигмоїдального типу виконують апроксимацію функції кількох змінних $X \in R^M$ безліч вихідних змінних $Y \in R^M$. Сигмоїдальна функція виконує роль функції активації нейронів. Оскільки вона має ненульове значення по всьому діапазоні вхідних даних, то у перетворенні вхідних даних у вихідні всередині мережі, беруть участь усі її нейрони. Радіальна мережа має два шари: перший шар складається з радіальних нейронів, а другий (вихідний) складається з одного нейрона або з кількох лінійних. На рисунку 3.6 представлена схема радіальної мережі із одним вихідним нейроном

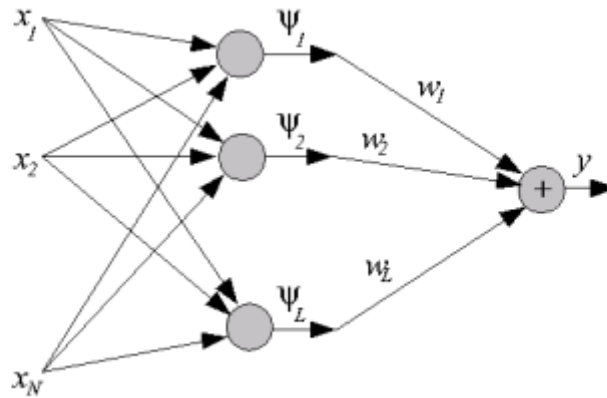


Рисунок 3.6 - Структурна схема радіальної нейронної мережі

$\Psi_i(\|X - C_i\|)$ - базисна функція або функція активації i -ого радіального нейрону. Апроксимація виконується мережею згідно з виразом, для спрощення $w_0 = 0$:

$$y = \sum_{i=1}^L w_i \cdot \Psi_i(\|X - C_i\|). \quad (3.17)$$

Завдання навчання: необхідно підібрати такі значення L , C_i та w_i , які доставляють мінімум цільової функції:

$$E = \frac{1}{2} \cdot \sum_{k=1}^p (\sum_{i=1}^L w_i \cdot \Psi(\|X^k - C_i\|) - d^k)^2, \quad (3.18)$$

де p -кількість навчальних вибірок.

Нехай G -матриця Гріна. Якщо параметри радіальних функцій відомі, то завдання навчання буде аналогічне рішення системи лінійних алгебраїчних рівнянь

$$G \cdot W = D,$$

де $W = [w_1, w_2, \dots, w_L]$ T - вектор ваг;

$$D = [d_1, d_2, \dots, d_p];$$

T - вектор очікуваних значень вихідного сигналу мережі.

Як радіальна функція найчастіше використовується функція Гауса:

$$\Psi(X) = \Psi(\|X - C_i\|_2) = \exp\left(-\frac{\|X - C_i\|_2^2}{2\sigma_i^2}\right), \quad (3.19)$$

де σ_i – параметр, який задає ширину функції.

Процес навчання радіальної мережі включає два етапи:

- для кожного радіального нейрона необхідно підібрати параметри функції Ψ ;
- для вихідного шару нейронів потрібно підібрати ваги.

Другий етап дещо простіше: якщо використовувати матрицю Гріна G , то обчислювальні витрати полягають у розрахунку зворотної матриці; якщо за формулою Гауса, то тут процес рішення знову розпадається надві підзадачі:

а) визначення центрів C_i ;

Мається на увазі визначення величини, для якої основними вимогами є повнота покриття області визначення та рівномірність розподілу даних.

б) розрахунок параметра σ_i .

Після того, як будуть визначені всі центри радіальної функції, вибираються параметри, що визначають величину області охоплення. Щоб максимально спростити завдання потрібно як значення σ_i прийняти евклідову відстань від центру C_i до найближчого нейрона. Радіальні нейронні мережі полегшують вибір початкових умов процесу навчання, забезпечують швидке досягнення оптимального рішення. Отримана архітектура радіальних мереж має структуру, аналогічну багат шаровій структурі сигмоїдальних мереж з одним прихованим шаром. Роль прихованих нейронів у ній грають базисні радіальні функції, що відрізняються своєю формою від сигмоїдальних функцій. Попри зазначену подібність, мережі цих типів важливі відрізняються один від одного. Такі мережі мають більш слабку здатність до узагальнення, в порівнянні з сигмоїдальними мережами. Цей недолік компенсується за рахунок розбиття навчальних класів на підкласи. При цьому кожен підклас матиме свій «центр», довкола якого реалізується радіальна базова функція. Крім того, алгоритм навчання радіальних мереж унікальний. За наявності тільки одного

прихованого шару та тісного зв'язку нейрона з областю навчальних даних, точка ще на початковому етапі навчання виявляється ближчим до оптимального рішення. У таблиці 1 представлений порівняльний аналіз другого розділу виділеними інформаційними моделями нейронних мереж. У таблиці наведено основні інформаційні моделі нейронних мереж (по горизонталі), по вертикалі перераховані основні завдання, які можна вирішити з використанням штучних нейронних мереж.

Таблиця 3.1 - Таблиця значень

	Односпрямовані багат шарові мережі	Рекурентні нейронні сіті (Мережа Елмана)	Мережі Кохонена	Радіальні нейронні сіті
Асоціативна пам'ять	+	+	-	+
Стиснення інформації	+	+	-	-
Прогнозування	+	+		+
Оптимізація	-	+	+	
Класифікація	+	-	+	+
Кластеризація	-	-	+	+

Асоціативна пам'ять - це особливий вид машинної пам'яті, використовується в програмах дуже швидкого пошуку. Стиснення інформації - алгоритмічний стиск інформації для зменшення займаного обсягу. Застосовується для раціонального використання пристроїв зберігання та передачі даних.

Прогнозування - оцінка можливих шляхів розвитку на підставі накопиченого досвіду та поточних припущень. Оптимізація-знаходження найкращого рішення задачі при заданих умовах та обмеженнях.

Класифікація – розв'язання задачі шляхом розбиття об'єктів на класи за будь-якими ознаками. Кластеризація – завдання, в якому потрібно розбити задану

вибірку об'єктів на непересічні підмножини, звані кластерами, так, щоб кожен кластер складався зі схожих об'єктів.

3.6 Інтенсивність дії руху лицьових м'язів

Інтенсивність руху м'язів визначається наступним чином: номеру одиниці зі СКЛД додаються літери з А по Е залежно від інтенсивності руху (від мінімальної до максимальної).

Значення:

- а) А - слабо помітне;
- б) В – незначне;
- в) С – помітне чи яскраво виражене;
- г) D – різке чи вкрай помітне;
- д) Е – виражене найвищою мірою.

Двигуни відносно вертикальної осі особи можуть бути: двосторонні, симетричні, односторонні, ліві, праві.

Список основних рухових одиниць представлений у таблиці 3.2:

Таблиця 3.2 – Таблиця дій

Номер одиниці дії	Назва в СКЛД	М'язові основи
0	Нейтральне обличчя	-
1	Підняти внутрішню частину брови	Лобова(медіальна частина)
2	Підняти зовнішню частину брови	Лобова (медіальна частина)
3	Опустити брови	м'яз, що опускає надперенесення, м'яз, опускаюча брова.

Продовження таблиці 3.2

1	2	3
4	Підняти верхню повіку	м'яз очної ямки, що піднімає верхнє повіко
5	Підняти щоки/щоку	круговая мышца глаза
6	Натягнути повіку/століття	круговий м'яз ока мімічний м'яз, що закриває очну щілину і розширює слізний мішок
7	Підняти крила носа	м'яз, що піднімає верхню губу і крило носа
8	Підняти верхню губу	м'яз, що піднімає верхню губу
9	Поглибити носогубну складку	мімічний м'яз, що поглиблює носогубну складку
10	Підняти куточки губ	великий вилицевий м'яз
11	Более отточено поднятъ уголки губ	м'яз, що піднімає кут рота, мімічний м'яз, що тягне кут рота догори
12	Поява ямочок біля губ	щічний м'яз - м'яз щічної області особи, що відтягує кут рота в сторони, притискає щоку до зубів, стискає щоки

Визначення інтенсивності руху лицьових м'язів, а також знання основних мимічних проявів на обличчі людини є невід'ємною частиною розробки інтелектуальної системи. Розробнику необхідно описати кожен кодований мимічний прояв у геометричному та математичному уявленні. Даний крок є найскладнішим, так як вимагає обробки великої кількості інформації та достатнього кількості знань у вибраній галузі.

4 РОЗВ'ЯЗАННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ НЕЙРОННОЮ МЕРЕЖЕЮ

Розв'язання задачі розпізнавання емоцій відноситься до задачі класифікації, тобто. нейронна мережа має віднести отриманий набір даних до емоцій, що відповідає заданому набору параметрів. Розглянемо математичний опис завдання розпізнавання: Нехай дано безліч M зображень осіб (емоція, наприклад здивування) $\{w_1, \dots, w_m\}$, кожне з яких має вектор значень ознак (мімічні ознаки) $X_i = (x_{i1}, \dots, x_{im})$, $i=1, \dots, m$, $x_j, j = 1, n$, де n - кількість ознак. Векторні ознаки віднесені експертами до деяких класам $\Omega_l, l = 1, \dots, k, M = \bigcup_{l=1}^k \Omega_l$ наступним чином: $\Omega_1 = \{w_1, \dots, w_{m_1}\}, \Omega_2 = \{w_{m_1+1}, \dots, w_{m_1+m_2}\}, \dots, \Omega_k = \{w_{m_{k-1}+1}, \dots, w_m\}$, де $m_1 + m_2 + \dots + m_k = m, m = |M|$.

Вся вибірка поділена на два непересічні підмножини: навчальне та тестове. Після навчання штучної нейронної мережі перевіряється якість її навчання на тісовому множині. штучна нейронна мережа прямого поширення із сигмоїдальною функцією активації $\sigma(x) = \frac{1}{1+e^{-x}}$ нейронів прихованого шару та лінійної функцією активації вихідного шару

Нейронна мережа має бути налаштована так, щоб при подачі вектора Признпків X_i на вхід, віднесеного до класу Ω_i на виході з номером i мережа видавала значення «1», але в інших виходах «0». Це досягається налаштування мережі методом зворотного розповсюдження помилки.

З урахуванням особливостей формування класів для розпізнавання осіб цікавить імовірнісна нейронна мережа. У ній зразки класифікуються на основі оцінок їхньої близькості до класів з урахуванням особливостей імовірнісного розподілу значень ознак. Для кожного класу на основі навчальних даних визначається функція густини розподілу ознаки, що характеризується математичним очікуванням та дисперсією.

Функція активації нейрона, що відповідає j -му класу має вигляд:

$$f_i = \exp(-\sum_{j=1}^n (w_{ij} - \frac{x_i}{\sigma_{ij}})^2),$$

де σ_{ij} – середньо квадратичне відхилення i -го ознаки j -го класу,

w_{ij} - математичне очікування i -го ознаки j -го класу.

У випадку, якщо на вхід вектор ознак був поданий невідомий зразок, то вибирається клас, якому відповідає максимальне вихідне значення.

4.1 Методика розпізнавання рухової активності особи

Для вирішення задачі розпізнавання обрано інформаційну модель нейронної мережі на основі багатошарового персептрона зі зворотним поширенням помилки,

- для створення прототипу емоцій (метрик) було обрано систему кодування лицьових рухів П. Екмана;

- для дослідження було обрано 6 базових емоцій: радість, смуток, гнів, огида, подив, страх. Такий набір дозволяє охоплювати максимальна кількість реальних виразів осіб. Кожна емоція має прототип виразу, наприклад, формула здивування:

$$1+2+5B+26.$$

У цій формулі рухові одиниці виражені числами, кожне з чисел характеризує мімічні прояви, яка задіює невелику частину м'язів обличчя. Латинська "B" позначає інтенсивність руху. Нейромережва модель навчається на великій кількості зображень або безперервному відеопотоку в режимі реального часу

Слід зазначити, що деякі з кодів інтенсивності дії є важко реєстрованими, наприклад коди A і B мають незначну різницю у зовнішньому прояві міміки:

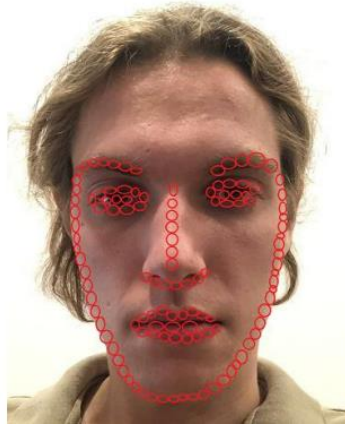


Рисунок 4.1 – Здивування з інтенсивністю А5

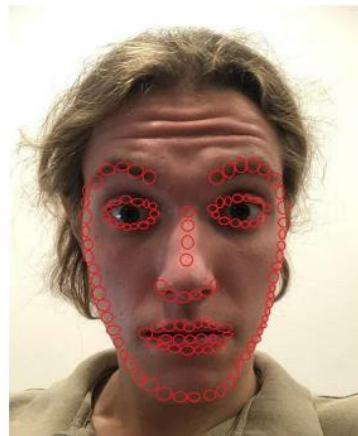


Рисунок 4.2 – Здивування з інтенсивністю А6

Двигательні одиниці особи (далі РОО) можна умовно розділити на три групи де перша група: статичні РОО - можливе визначення фотографії, 2 група: динамічні РОО – необхідно відслідковувати зміну положення ключових точок на обличчі або мати середнє значення відстані за цій РОО, 3 група: порожні РОО - беруть активну участь у прояві емоцій, проте не реєструються алгоритмами пошуку (ямочки на щоках).

На обличчі людини можна виділити до 80 значних областей на обличчі. Як правило це межі очей, рота та брів. Вилицеві м'язи не є важливим ознакою висловлювання тій чи іншій емоції. Кожна з емоцій має динаміку «переживання».

Розглянемо складові груп, зазначені в таблиці нижче

Таблиця 4.1 – Складові за емоціями

	Статика	Динаміка	Порожні
Радість	12	6	-
Здивування	1,25	2,5,26	27
Гнів	4,25	5,23,24,26	7,10,22
Огида	15	9,26,10	16,17
Печаль	1,4,15	-	11
Страх	1	2,5,20,38	-

Система ключових метрик для розпізнавання:

- висота рота: верхня губа – низ нижньої губи;
- висота відкритого рота: низ верхньої губи верх нижньої губи;
- куточок губ донизу: куточок рота – верх нижньої губи;
- куточок губ нагору: куточок рота – верх верхньої губи;
- ширина рота: лівий куточок рота – правий куточок рота;
- висота підборіддя: низ нижньої губи – підборіддя;
- ширина ока: верх ока – низ ока;
- висота брови: верхній центр брови – середина очей;
- внутрішній куточок брови: внутрішній кут брови – внутрішній кут очей;
- зовнішній куточок брови: зовнішній кут брови - зовнішній кут ока.

Метрики під номерами 3, 4, 7–10 є несиметричними та розраховуються для лівої та правої половини окремо. Нормування показників здійснюється на відстані між зіницями. Воно мінімально відрізняється у різних людей і добре стандартизує навчальну вибірку.

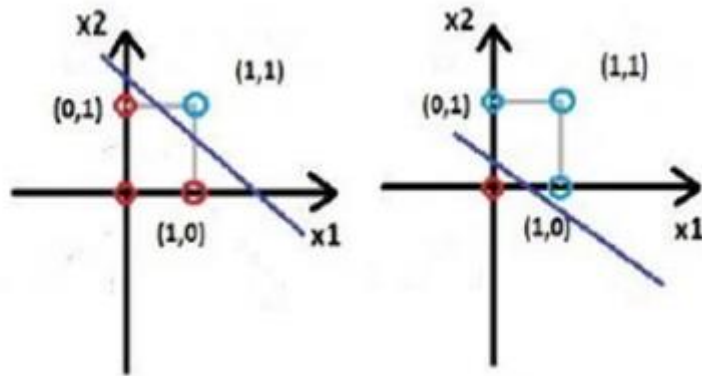


Рисунок 4.8 – Гіперплощини

Як приклад у таблиці істинності показані значення вихідного сигналу для логічних функцій «І» та «АБО» залежно від комбінації вхідних сигналів. Таблиці відповідностей для координат на гіперплощинах:

Таблиця 4.2 - Відповідно до значень логічне «І»

Вход 1	Вход 2	Выход
0	0	0
0	1	0
1	0	0
1	1	1

Таблиця 4.3 - Відповідності значенням логічне «АБО»

Вход 1	Вход 2	Выход
0	0	0
0	1	1
1	0	1
1	1	1

Перцептрон може класифікувати лише образи, які розділені за допомогою гіперплощини. Сформулюємо алгоритм зворотного поширення помилки:

-Ініціалізація ваг зв'язків: невеликі випадкові значення та максимальна середньоквадратична помилка, потім подати на вхід мережі вхідний вектор;

здійснити поширення сигналів відповідно до прямих зв'язками; обчислення помилки та помилка вихідного шару нейронів; та обчислення помилки внутрішнього шару нейронів; оновлення ваги зв'язків кожного шару.

5 ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА НЕЙРОННОЇ МЕРЕЖІ У ПРОГРАМНОМУ СЕРЕДОВИЩІ MATLAB

Навчання нейронної мережі проводилося в інтерактивному середовищі програмування та чисельних розрахунків, а також візуалізації результатів - MATLAB. За допомогою MATLAB можна аналізувати дані, розробляти алгоритми, створювати моделі та додатки. Дані для нейронної мережі із застосуванням багат шарового персептрону:

Кількість вхідних нейронів: 16 (матриця 4×4)

Вихідних нейронів: 6 (кількість емоцій);

Нейрони у прихованому шарі: 11.

Кількість нейронів у прихованому шарі дорівнює половині суми вхідних і вихідних нейронів. Для кожного шару мережі є вектор зсувів. Позначимо вектор прихованого шару b_1 (включає 11 нейронів).

Вихідний шар включає 6 нейронів - b_2 .

Результат суми усунення кожного нейрона є аргументом функції активації нейрона Візьмемо кількість спостережень приблизно $3/5$ від максимально можливого. Оскільки 16 бінарних операцій складають вхідний вектор, то число можливих комбінацій складе $2^{16} = 65536$. Навчальна послідовність має становити приблизно 65536 вхідних образів. Для навчання нейронної мережі спочатку необхідно отримати зміщення та значення ваг, які мінімізують помилку навчання.

Алгоритм виконується строго за складеним планом та на кожному циклі виконують певний набір дій:

- елементи навчальної послідовності подаються на вхід мережі один за іншим;
- елементи навчальної послідовності порівнюються з цільовими (бажаними) значеннями;
- обчислюється функціонал помилки;
- значення функціоналу та його градієнта використовуються для

коригування ваг;

- операції повторюються доти, доки не буде виконано певну кількість циклів або поки помилка не перестане зменшуватись.

Градiєнтний алгоритм навчання є найшвидшим з наявних алгоритмів у програмному пакеті MATLAB, тому для навчання мережі його застосування раціонально.

Щоб сформувати значення вихідного вектора, потрібно вибрати таку функцію активації, для якої діапазон вихідних сигналів був би визначено від 0 до 1. Для такого випадку виберемо логарифмічну сигмоїдальну функцію активації. Ця функція також підходить для нейронів прихованого та вихідного шарів.

Навчання за методом зворотного розповсюдження помилки, вимагає створення двошарової нейронної мережі прямої передачі сигналу з сигмоїдальними функціями активації. Запишемо у командному рядку програми

```
net=new_F([01;01;01;01;01;01;01;01;01;01...33[0-1]],[11,6],{'logsig','logsig'},
trainrp');
```

Пояснення:

- мережа має один вектор входу з 16 елементами, для яких допустима межа значень та сама [0 1];
- у першому прихованому шарі мережа має нейронів, у вихідному шарі 7 нейронів;
- використані функції активації: logsig- обчислює вихід шару за його входу, вказана для кожного шару; trainrp – функція навчання та модифікації ваг та зсувів мережі, відповідно до заданого алгоритму навчання.

Налаштування параметрів алгоритму навчання:

```
net.trainparam.goal=1e-5;
```

```
net.trainparam.show=50;
```

```
net.trainparam.epochs=5000;
```

goal – граничне значення критерію навчання, помилка мережі;

epochs – максимальна кількість циклів навчання;

show-інтервал виведення інформації.

```
inputs = [0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1; 0 ... 1];
```

```
targets = [0 ... 1; 0...0; 0...0; 1 ... 0];
```

Підсумок: виконано налаштування мережі, сформовано навчальну послідовність, вибрано алгоритм навчання.

Процес навчання мережі:

У командний рядок MATLAB введемо наступний вираз:

```
net = train (inputs, targets);
```

Запуск процесу навчання нейронної мережі (рис. 7)

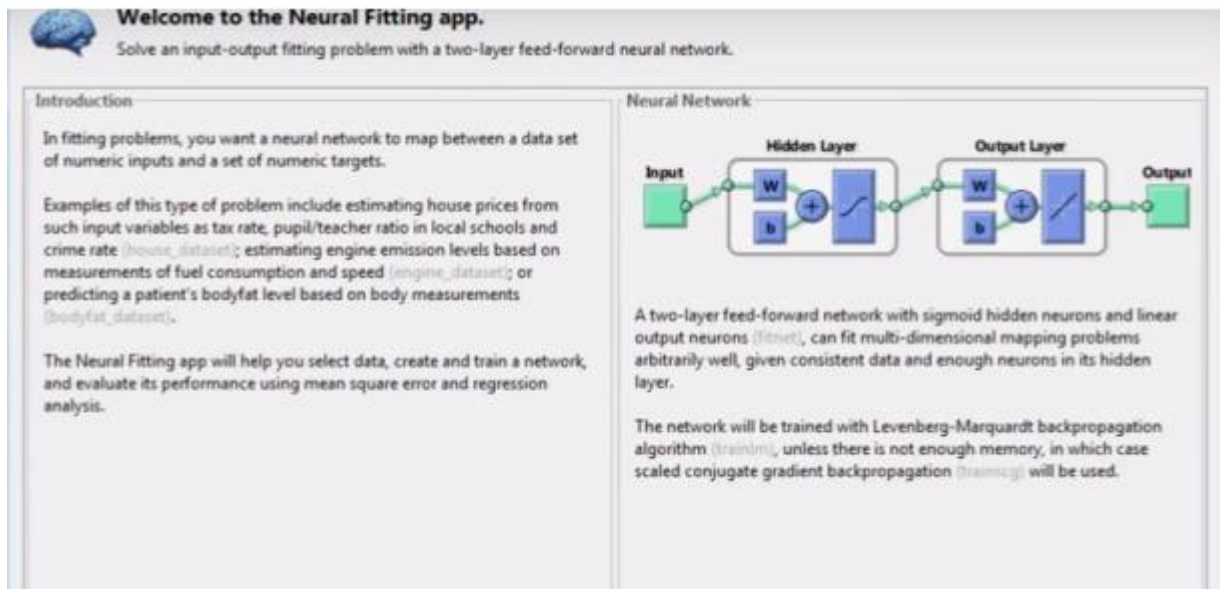


Рисунок 5.1 – Запуск процесу навчання у MATLAB

Вибір файлу з даними для входу та виходу (рисунок 5.2):

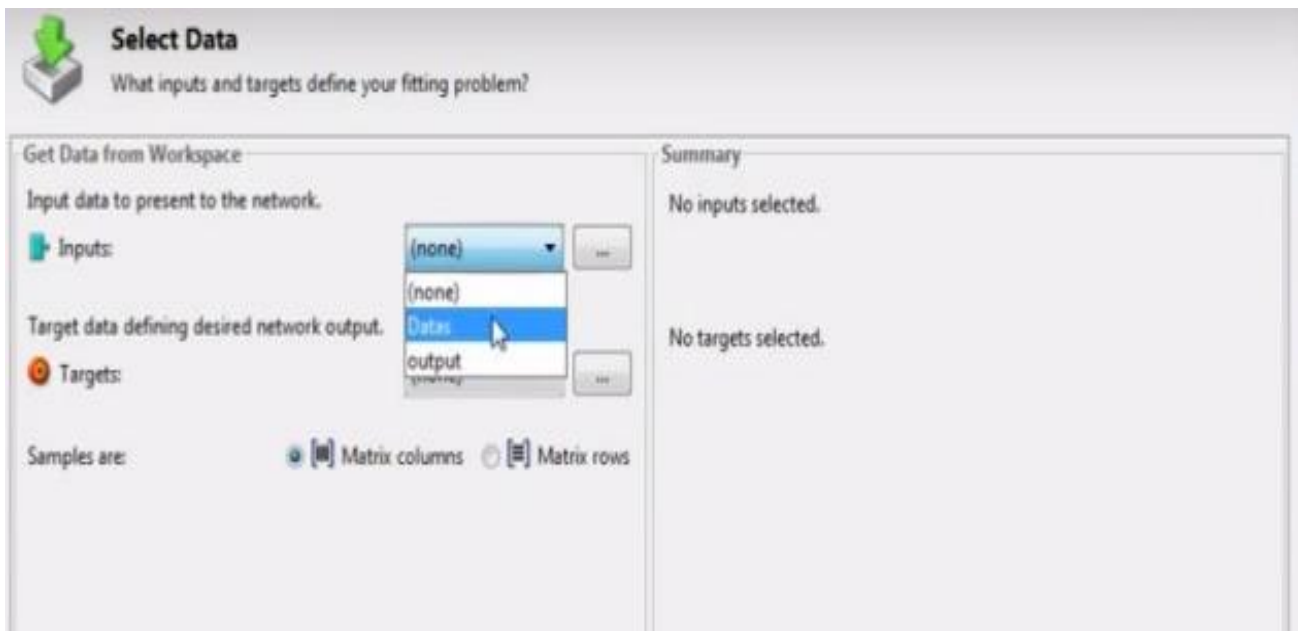


Рисунок 5.2 – Вибір даних із сформованого файлу

Вибір кількості нейронних шарів (рисунок 5.3):

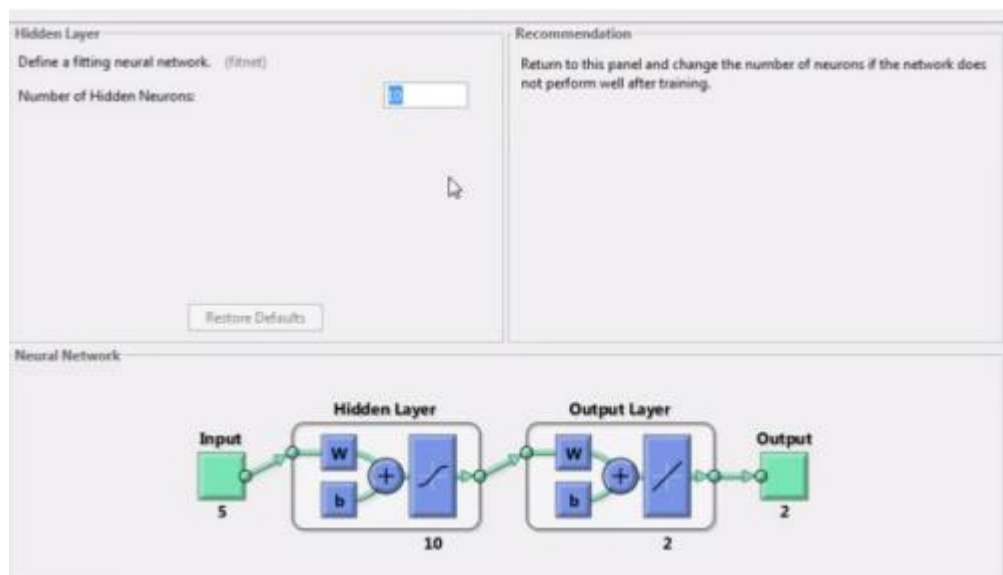


Рисунок 5.3 – Вибір шарів нейронної мережі

Вибір алгоритму навчання нейронної мережі (рис. 5.4):

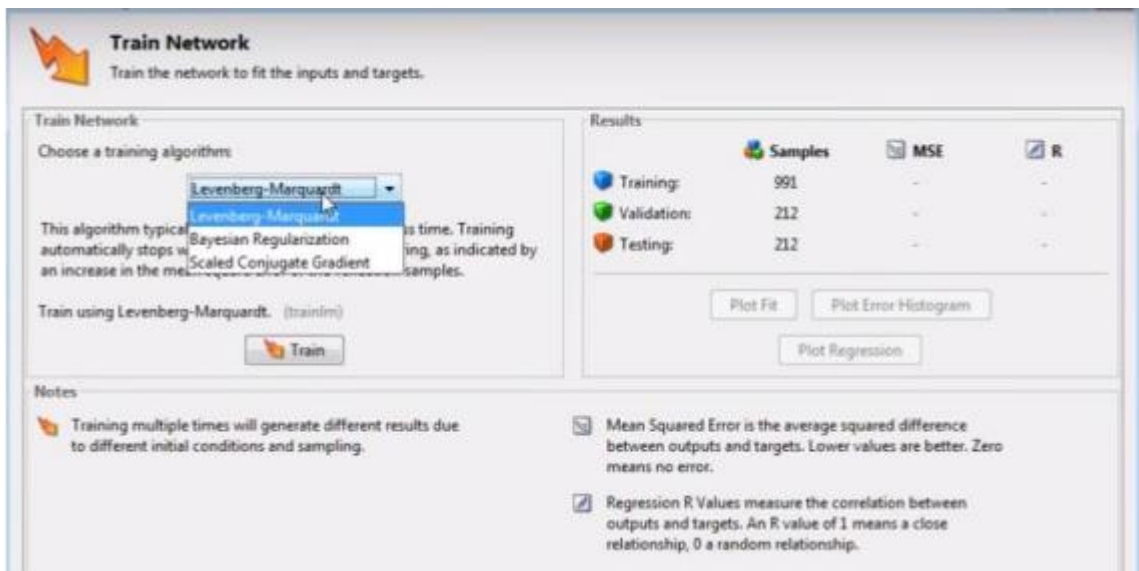


Рисунок 5.4 – Алгоритм навчання нейронної мережі

Процес навчання нейронної мережі (рис. 5.5):

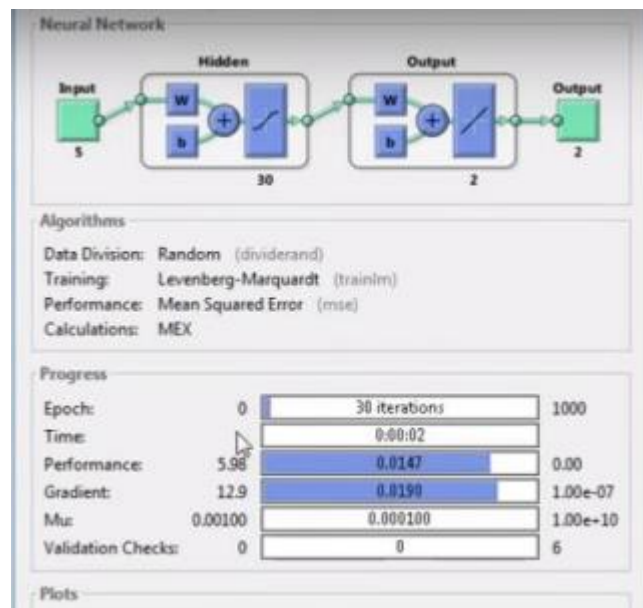


Рисунок 5.5 - Навчання нейронної мережі

Вікно з навчальними здібностями мережі (рис. 5.6):

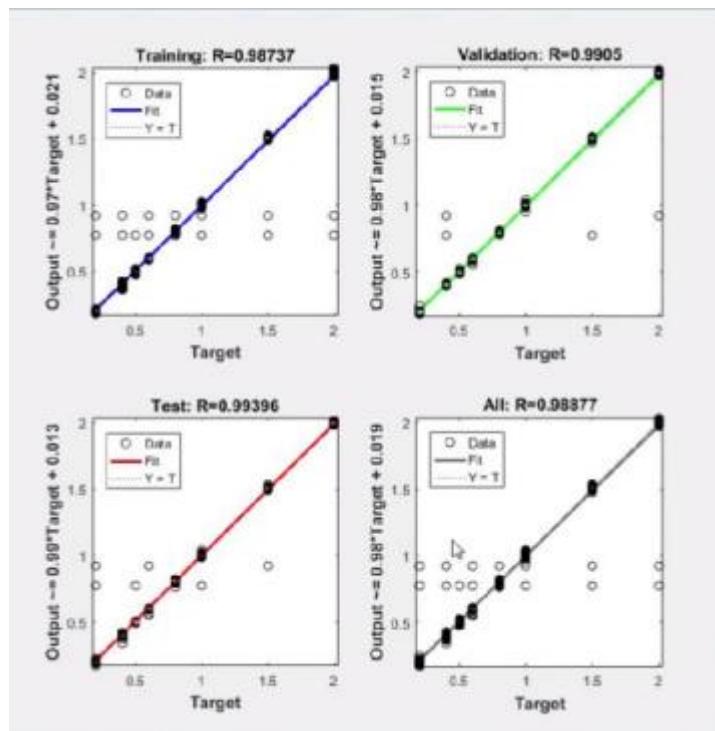


Рисунок 5.6- Графіки навчання нейромережі

У ході експерименту емоція вважалася розпізнаною, якщо на виході нейронна мережа мала найбільше значення. В результаті моделювання двошарової нейронної мережі були отримані дві матриці вагових коефіцієнтів IW для прихованого шару, LW для вихідного шару і два вектори зсувів $b1$ і $b2$ відповідно. Для нейромережі необхідно побудувати математичну модель та провести аналіз її функціонування представлено на рисунках.5.7, 5.8.

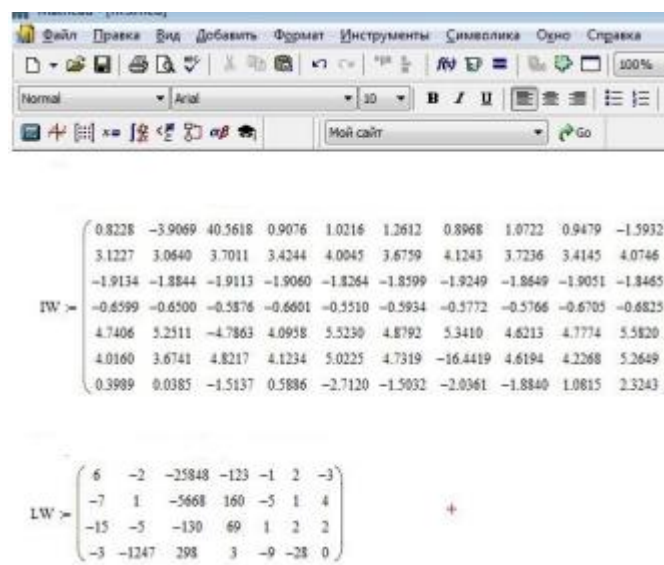


Рисунок 5.7 – Матриці коефіцієнтів шарів нейронів

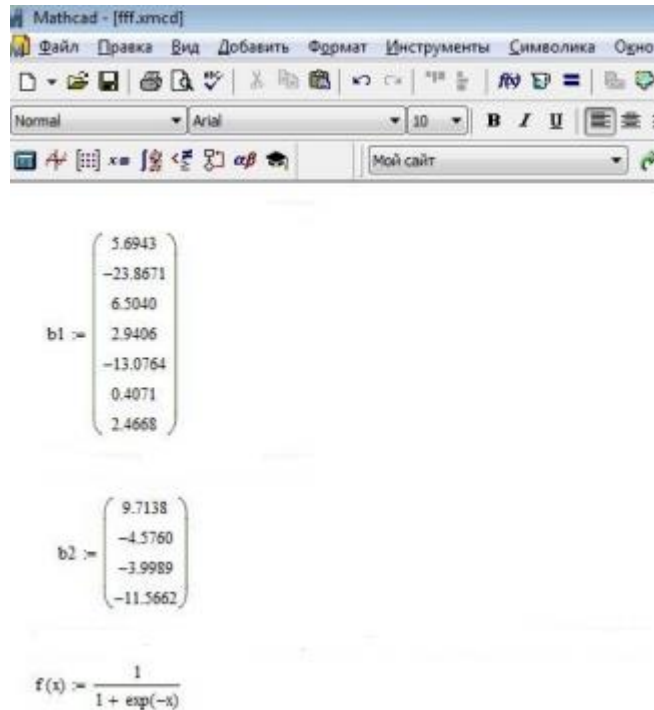


Рисунок 5.8 – Векторні зсуви нейронної мережі

Розрахунок вектора аргументів для функцій активації нейронів представлений на рисунках 5.9 та 5.10:

$$A = \begin{pmatrix} \sum s_1^{(0,j)} \\ \sum s_1^{(1,j)} \\ \sum s_1^{(2,j)} \\ \sum s_1^{(3,j)} \\ \sum s_1^{(4,j)} \\ \sum s_1^{(5,j)} \\ \sum s_1^{(6,j)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0.045 \\ 7.67 \times 10^{-3} \\ 0.275 \\ 0.999 \\ 1 \\ 0.995 \end{pmatrix}$$

Рисунок 5.9 – Розрахунок A вектора аргументів

$$Y = \begin{pmatrix} \sum s_2^{(0,j)} \\ \sum s_2^{(1,j)} \\ \sum s_2^{(2,j)} \\ \sum s_2^{(3,j)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0.00002 \\ 0.97674 \\ 0 \end{pmatrix}$$

Рисунок 5.10 – Розрахунок Y вектора аргументів

Результат: у процесі моделювання нейронної мережі у програмній середовищі MathCAD та подачі на вхід вектор, який відповідає емоції "здивування", мережа згенерувала правильний вихідний сигнал.

Інтенсивність емоції мала найбільше значення. Підсумок аналізу можна записати в наступному вигляді (рис.5.11):

$$\begin{array}{l}
 S1 = P * IW + b1 \\
 A = f(S1) \\
 S2 = A * LW + b2 \\
 Y = f(S2)
 \end{array}
 \qquad \text{або}
 \qquad
 Y = f(f(P * IW + b1) * LW + b2)$$

Рисунок 5.11 – Модель визначення емоції «здивування»

P- вектор входу з 11 елементів;

IW-матриця вагових коефіцієнтів синапсів (зв'язків першого шару зі входами);

b1- вектор зсувів першого шару нейронів;

b2- вектор зсувів вихідного шару;

f(x)-логарифмічна сигмоїдальна функція активації для всіх нейронів мережі;

Y-вихідний вектор сигналів нейронної мережі (складається з 6 елементів);

S1, S2 - аргумент для функцій активації 1 і 2 шари відповідно.

Висновки за главою: Завдання розпізнавання з використанням моделі багатошарового перцептрон може мати коректне рішення. Складності реалізації та навчання нейронної мережі можуть виникнути при формуванні початкових даних. Для запуску процедури розпізнавання потрібна попередня обробка зображення та розкладання кожного елемента особи на окремі складові. Модель багатошарового перцептрона виявилася адекватною для обраної системи тестування Крім того, багатошаровий перцептрон може використовувати різні алгоритми навчання, завдяки гнучкій структурі та високу здатність до навчання. Розв'язання задачі розпізнавання емоцій

проводилося за заданим алгоритму. Програмні пакети MATLAB та MathCAD мають широкий набір інструментів, що допомагають реалізувати математичну модель завдання, а також допомагають навчити та налаштувати нейронну мережу. Дані програмні продукти швидко обчислюють громіздкі висловлювання, що дозволяє прискорити процес отримання відповіді

ВИСНОВКИ

Реалізація інтелектуальної системи, здатної розпізнавати емоції людини – це рішення технічно та математично складного завдання, що вимагає поглибленого вивчення предметної галузі, власного аналізу, вміння та навичок при обробці та виборі експериментальних даних, глибокі знання в галузі дискретної математики, геометрії, програмування, а також психології та багатьох інших галузях науки.

Використання нейронних мереж при розробці систем розпізнавання, дозволяє структурувати поганоформалізовані дані, прискорити процес обробки даних, оцінити адекватність середовища, в яке вміщено інформаційну модель нейронної мережі.

Метрична система дозволяє створити прототипи емоцій, завдяки яким у розробника з'являється можливість «взаємодії» з нейронною мережею, тобто. навчання з урахуванням виявлених прототипів емоцій.

Модель на основі багатошарового перцептрон оптимальна за внутрішньою структурою та способом управління інформаційними потоками між нейронами Крім того, така модель здатна мінімізувати кількість вхідних елементів. Модель на основі багатошарового перцептрон є універсальною моделлю та підходить для вирішення завдань різного рівня складності, у тому числі і для розв'язання задачі розпізнавання емоцій На підставі вивченого матеріалу зі штучних інтелектуальним системам було виявлено найефективніші системи та точні методи розпізнавання, відмічені успішні приклади реалізації. Система, яка використовувалася у цій роботі була самонавчання.

Огляд основних видів інформаційних моделей нейронних мереж дозволив виділити дві універсальні моделі, застосовні для широкого кола завдань: радіальні нейронні мережі прямого поширення та односпрямовані багатошарові мережі. Вибір інформаційної моделі для вирішення практичного завдання було зроблено на користь односпрямованих

багатошарових мереж (модель багатошарового персептрону). Ця модель підходить за внутрішньою структурою та математичним описом та передбачає різні підходи до розв'язання задачі.

Існуючі методи розпізнавання емоцій у магістерській дисертації було розглянуто теоретично. Серед вивчених методів розпізнавання було виділено метод із використанням штучних нейронних мереж.

У практичній частині роботи був розглянутий математичний апарат реалізації алгоритму розпізнавання: вибір інваріантів особи, адекватної метрики, розпізнавання образів нейронною мережею. Також була представлена методика розпізнавання рухової активності особи та вирішено завдання розпізнавання емоцій у програмних продуктах аналізу та обробки даних MATLAB та MathCAD.

На підставі отриманих результатів можна зробити висновок про те, що обрана інформаційна модель штучної нейронної мережі відповідає моделі реальної системи та враховує всі її характеристики.

Всі поставлені завдання були виконані в повному обсязі, отже, головної мети було досягнуто.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. D. Serre (2002). *Matrices: Theory and Applications*. New York: Springer Verlag.
2. E. Eiben, R. Hinterding (1999). "Parameter control in evolutionary algorithms," *IEEE Trans. Evol. Computer*
3. Josef Tvrđíka, Ivan Křivýa, Ladislav Mišíkb. (2007). Adaptive population-based search: Application to estimation of nonlinear regression parameters
4. O. Bychkov (2020). Using neural networks application for the font recognition task solution.
5. D. Horii, A. Ito and T. Nose (2021). Analysis of Feature Extraction by convolutional neural network for speech emotion recognition.
6. G. Lou and H. Shi (2020). Face image recognition based on convolutional neural network.
7. Y. Zhang, Z. Xue (2010). RBF Neural Network Application to Face Recognition.
8. M. Z. Alom, M. Alam, T.M. Taha and K. M. Iftakharuddin (2017). Object recognition using cellular simultaneous recurrent networks and convolutional neural network.
9. M.W. Kurzynski and E. Puchala (1992). Algorithms of the multiperspective pattern recognition.
10. J. Chen, (2020). "Image Recognition Technology Based on Neural Network".
11. Nunes, C., & Pádua, F. (2021). A Convolutional Neural Network for Learning Local Feature Descriptors on Multispectral Images.
12. M. A. Dede, E. Aptoula, and Y. Genc, (2019). "Deep network ensembles for aerial scene classification".
13. T. Mouats, N. Aouf, D. Nam, and S. Vidas (2019) "Performance evaluation of feature detectors and descriptors beyond the visible".
14. V. Balntas, E. Riba, D. Ponsa, and K. Mikolajczyk (2016) "Learning local feature descriptors with triplets and shallow convolutional neural networks."

15.H. Le, C. Smailis, L. Shi, and I. Kakadiaris (2020) "EDGE20: A cross spectral evaluation dataset for multiple surveillance problems"