

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО–НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра інженерії програмного забезпечення

Пояснювальна записка

до магістерської роботи

на ступінь вищої освіти магістр

на тему: «**РОЗРОБКА ТА НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ
РОЗПІЗНАВАННЯ СИМВОЛІВ**»

Виконав: студент 6 курсу, групи ПДМ–61

спеціальності

121 Інженерія програмного забезпечення

(шифр і назва спеціальності)

Буков Р. Д.

(прізвище та ініціали)

Керівник Щербина І.С.

(прізвище та ініціали)

Рецензент _____

(прізвище та ініціали)

Київ – 2021

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра Інженерії програмного забезпечення

Ступінь вищої освіти - «Магістр»

Спеціальність – 123 «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Інженерії програмного забезпечення

_____ О.В. Негоденко

“ ____ ” ____ 20 ____ року

ЗАВДАННЯ

НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Букову Ростиславу Дмитровичу

(прізвище, ім'я, по батькові)

1. Тема роботи: «Розробка та навчання нейронної мережі для розпізнавання символів»

Керівник роботи к. т. н., доцент кафедри ІІЗ Ірина Сергіївна Щербина,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затвердені наказом вищого навчального закладу від “13” жовтня року №230.

2. Строк подання студентом роботи 24.12.2020.

3. Вихідні дані до роботи:

3.1. Вимоги до кваліфікаційної роботи магістра з актуальних завдань спеціальності;

3.2. Нормативні матеріали (стандарти, Гости);

3.3. Технічні вимоги;

3.4. Науково-технічна література з питань, пов'язаних з темою роботи.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити):

4.1. Порівняльний аналіз результатів, отриманих іншими авторами;

4.2. Методика дослідження;

4.3. Результати дослідження;

4.4. Висновки

5. Перелік графічного матеріалу

1. Актуальність теми

2. Об'єкт та предмет дослідження

3. Мета дослідження

4. Поставлена мета передбачає розв'язання таких завдань

5. Методи дослідження

6. Практичне значення одержаних результатів

-
7. Наукова новизна одержаних результатів
 8. Схема алгоритму нейронної мережі для розпізнавання символів
 9. Алгоритм розпізнавання символів
 10. Висновки
-

6. Дата видачі завдання 02.11.2020

КАЛЕНДАРНИЙ ПЛАН

№ з / п	Назва етапів магістерської роботи	Строк викона ння етапів робо ти	Приміт ка
1	Підбір науково-технічної літератури	02.11.20	Виконано
2	Огляд існуючих рішень та літератури	03.11.20	Виконано
3	Аналіз методів побудови нейронних мереж	12.11.20	Виконано
4	Розробка алгоритму нейронної мережі	16.11.20	Виконано
5	Вступ, висновки, реферат	08.12.20	Виконано
6	Розробка обов'язкових демонстраційних матеріалів	11.12.20	Виконано
7	Здача роботи	24.12.20	Виконано

Студент _____
(підпис) (прізвище та ініціали)

Керівник роботи _____
підпис) (прізвище та ініціал

РЕФЕРАТ

Дипломний проект на тему «Розробка та навчання нейронної мережі для розпізнавання символів» складається зі вступу, основної частини із трьох розділів, висновків, переліку джерел посилання, додатків.

Загальний обсяг роботи становить 86 сторінок. Список джерел посилання складається з 43 найменувань.

Мета роботи—аналіз особливостей розробки та навчання нейронної мережі для розпізнавання символів.

Об'єкт дослідження—процес проектування та навчання нейронної мережі для розпізнавання символів.

Предмет дослідження - специфіка розробки та навчання нейронної мережі для розпізнавання символів.

Наукова новизна одержаних результатів: Результати дослідження пропонують альтернативний метод застосування штучного інтелекту під час розпізнавання символів.

Практичне значення одержаних результатів полягає в тому, що дослідження ґрунтується на результатах поглибленого вивчення особливостей проектування та навчання нейронної мережі для заданих умов, а також в розробці практичних рекомендацій щодо використання нейронних мереж для розпізнавання символів.

Ключові слова: РОЗПІЗНАВАННЯ СИМВОЛІВ, ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ, НЕЙРОННІ МЕРЕЖІ, ШТУЧНИЙ ІНТЕЛЕКТ, МЕТОДИ НАВЧАННЯ.

ЗМІСТ

ВСТУП.....	9
1. ОГЛЯД ПРЕДМЕТНОЇ ГАЛУЗІ І ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ.....	12
1.1. Визначення поняття «нейронної мережі» та його сутність.....	12
1.2. Класифікація нейронних мереж.....	19
1.3. Алгоритми навчання нейронних мереж.....	25
1.4. Постановка задачі дослідження.....	31
Висновки до розділу 1.....	31
2. ОСОБЛИВОСТІ НАВЧАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ.....	32
2.1. Специфіка застосування нейронних мереж.....	32
2.2. Огляд програмних засобів, що використовують нейронні мережі.....	37
2.3. Формалізація кластерного аналізу та адаптація методів сегментації до умов задачі.....	41
Висновки до другого розділу 2.....	51
3. ПРОЕКТ РІШЕННЯ З РЕАЛІЗАЦІЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ЗАДАНИХ УМОВ.....	52
3.1. Методика синтезу штучних нейронних мереж для розпізнавання символів.....	52
3.2. Комбінування двох мереж з однаковою структурою.....	60
3.3. Рекомендації з використання нейронних мереж для розпізнавання символів.....	76
Висновки до розділу 3.....	80
ВИСНОВКИ.....	81
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	85
ДОДАТКИ.....	89

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

БЗ – база знань;

ЗНМ – згоркова нейронна мережа;

ІБД – інтегрована база даних;

ІІТ – інтелектуальні інформаційні технології;

ЕОД – електронний обмін даними;

НМ – нейронна мережа;

ПК – персональний комп'ютер,

РНМ – рекурентна нейронна мережа;

СРС – система розпізнавання символів,

ШІ – штучний інтелект;

ШНМ – штучні нейронні мережі.

ВСТУП

Актуальність дослідження. На сьогоднішньому етапі розвитку суспільства інформаційні технології стають головною причиною значного зростання актуальності сфер наукової діяльності, що пов'язані з математичним моделюванням процесів та явищ. Моделювання реальних об'єктів навколишнього світу, як правило, супроводжується значними труднощами, які виникають ще на етапі постановки задачі.

Слід зазначити, що ці труднощі частіше за все є наслідком недосконалості обчислювальних методів та засобів їх реалізації. Вони також стають причиною існування великої кількості нерозв'язаних проблем, що виникають при розв'язанні задач в різних галузях практичної діяльності. Тому особливо актуальними стають задачі, що мають некоректну постановку, тобто такі, для яких відсутні оптимальні алгоритми їх розв'язку [4, с. 38].

Останнім часом проблема розпізнавання символів користується все більшою популярністю, адже розпізнавання зображення, тексту чи мови, а також різноманітних явищ сприяє спрощенню комунікативного зв'язку людини з персональним комп'ютером (ПК), допомагає застосовувати різні системи штучного інтелекту в різних галузях народного господарства. Можливість сприймати зовнішній світ у формі символів сприяє передумовам дослідження властивостей величезної кількості об'єктів завдяки ознайомленню з кінцевою їх кількістю, а об'єктивна ознака засадничої властивості символів допомагає створювати модель їх розпізнавання. Крім того, системи розпізнавання символів, що використовують штучний інтелект (ШІ), наразі отримали активний розвиток та з успіхом впроваджуються в практичному житті. Поліпшується як науковий рівень штучних нейронних мереж (ШНМ), так і математичні потужності обладнання технічного гатунку [8, с. 12].

Варто акцентувати, що наразі існує багато підходів стосовно вирішення проблеми розпізнавання символів, але більшість із них або вузько спрямовані на

певну сферу розпізнавання (вони показують високі результати розпізнавання, але не є універсальними), або якість розпізнавання є дуже низькою, й сам метод працює незадовільно. Тому для вирішення задач оптимізації системи розпізнавання символів (СРС) часто застосовуються високоінтелектуальні системи на основі ШНМ. Однак штучні нейронні мережі не є інструментом для вирішення задач будь-якого типу. Вони є непридатними для виконання таких задач як нарахування зарплатні, проте вони мають перевагу під час задач розпізнавання символів, з якими погано або взагалі не справляються звичайні ПК.

Однак, незважаючи на активні дослідження в цій сфері, все ще залишаються не вирішеними в повному обсязі багато проблем, пов'язаних з розробкою методів і алгоритмів синтезу нейронної мережі та інтерпретації результатів її роботи для конкретних умов, недостатньо точно описані питання створення комбінованих моделей для розпізнавання символів та особливості їх реалізації. Виходячи з вищенаведеного, наше дослідження особливостей розробки практичного застосування штучних нейронних мережв СРС є актуальним.

Мета та завдання дослідження. Метою роботи було проаналізувати особливості розробки та навчання нейронної мережі для розпізнавання символів. Для досягнення мети були поставлені наступні завдання:

- розглянути теоретичні засади навчання нейронних мереж;
- дослідити класифікацію нейронних мереж;
- визначити алгоритми навчання нейронних мереж;
- проаналізувати особливості навчання нейронних мереж для СРС;
- розробити та застосувати алгоритм ШНМ для СРС;

Об'єкт дослідження – процес проектування та навчання нейронної мережі для розпізнавання символів.

Предмет дослідження – специфіка розробки та навчання нейронної мережі для розпізнавання символів.

- У роботі застосовуються такі загальнонаукові **методи дослідження** як
- методи системного аналізу,

- аналіз наукової літератури,
- спостереження,
- абстрагування,
- узагальнення.

Джерела дослідження представлені науковими працями таких науковців, як Г. Аваліані, Ю. Борисов, Е. Вентцель, Г. Вороновський, О. Галушкін, Б. Головкін, В. Комашинський, В. Круглов, В. Месюра, Д. Поліщук, Б. Советов, Ф. Уосерман, Л. Ясницький та інших.

Наукова новизна одержаних результатів. Результати дослідження пропонують альтернативний метод застосування штучного інтелекту під час розпізнавання символів.

Практичне значення одержаних результатів полягає в тому, що дослідження ґрунтується на результатах поглибленого вивчення особливостей застосування перенесення навчання нейронних мереж, а також розроблено модуль на основі запропонованого алгоритму, який має перевагу в швидкості навчання в порівнянні з альтернативними підходами.

Публікації. За результатами наукового дослідження опубліковано наукову статтю.

Структура роботи. Дипломна робота складається зі вступу, трьох розділів, висновків, списку джерел посилання, додатків. Основний зміст роботи викладено на 86 сторінках друкованого тексту. Список використаних джерел складається з 43 найменувань.

1 ОГЛЯД ПРЕДМЕТНОЇ ГАЛУЗІ І ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Визначення поняття «нейронної мережі» та його сутність

Нейронна мережа (НМ) являє собою ланцюг нейронів, іншими словами ШНМ мережа складається з штучних нейронів або вузлів. Таким чином, нейронна мережа - це або біологічна нейронна мережа, що складається з реальних біологічних нейронів, або штучна нейронна мережа, яка створюється для вирішення завдань штучного інтелекту. ШНМ можуть використовуватися для прогнозного моделювання, адаптивного управління і додатків, де їх можна навчати за допомогою набору даних. Самонавчання на основі досвіду може відбуватися в мережах, які можуть робити висновки зі складного і, здавалося б, незв'язаного набору інформації [9, с. 247].

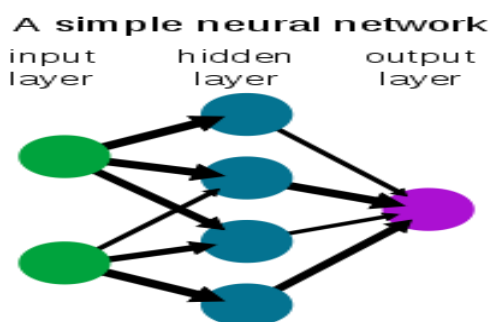


Рисунок 1.1 - Спрощений вид штучної нейронної мережі з прямим зв'язком

НМ складається з груп, хімічно чи функціонально пов'язаних один з одним нейронів. Один нейрон може пов'язуватися з багатьма іншими нейронами, і загальне число нейронів та з'єднань в мережі іноді доволі значне. Зв'язки, що носять назву синапсів, зазвичай утворюються від аксонів до дендритам, хоча можливі і інші зв'язки. Крім електричних сигналів, існують і інші форми передачі

сигналів, що виникають в результаті дифузії нейромедіаторів. Штучний інтелект, когнітивне моделювання та нейронні мережі - це парадигми обробки інформації, натхненні тим, як біологічні нейронні системи обробляють дані. Штучний інтелект і когнітивне моделювання намагаються імітувати деякі властивості біологічних нейронних мереж. В сфері штучного інтелекту ШНМ успішно застосовуються для розпізнавання мови, аналізу зображень та адаптивного управління з метою створення програмних агентів (в комп'ютерних і відеоіграх) або автономних роботів[32, с. 156].

Історично цифрові комп'ютери походять від моделі фон Неймана і працюють за допомогою виконання явних інструкцій через доступ до пам'яті ряду процесорів. З іншого боку, походження нейронних мереж засноване на спробах моделювання обробки інформації в біологічних системах. На відміну від моделі фон Неймана, нейромережеві обчислення не поділяють пам'ять і обробку. Теорія нейронних мереж послужила як для більш точного визначення того, як функціонують нейрони мозку, так і для забезпечення основи для зусиль зі створення штучного інтелекту. В кінці 1940-х років психолог Дональд Хебб створив гіпотезу навчання, засновану на механізмі нейронної пластичності, яка зараз відома як навчання Хебба. Це навчання вважається «типовим» правилом навчання без учителя, а його пізніші варіанти були ранніми моделями для довгострокового потенціювання. Ці ідеї почали застосовуватися до обчислювальних моделей в 1948 році з В-типу машин Тьюринга.

В 1958 році Розенблатт створив перцептрон, алгоритм розпізнавання образів, заснований на дворівневої навчається комп'ютерної мережі з використанням простого додавання і віднімання. У математичній нотації Розенблатт також описав схему, що не входить в базовий перцептрон. Ключовим моментом в більш пізніх досягненнях використання ШНМ був алгоритм зворотного поширення помилки, який ефективно вирішив проблему «АБО» (1975 рік).

На даний час нейронні мережі являють собою системи, що мають архітектурну складову, використання якої допомагає імітувати робочий процес нейронів. Математичний прототип нейрона являє собою деякий універсальний нелінійний елемент із можливістю широкої зміни і настроювання його характеристик (параметрів).

Відомо, що ще в 40-х роках минулого століття досягнення нейробіології дозволили спроектувати першу ШНМ, що імітувала робочий процес мозку людини. Але тільки через десятиліття, разом з виникненням сучасних ПКі відповідного ПЗ стала можливою розробка складних додатків в галузі штучних нейронів. З того часу наукова дисципліна, присвячена дослідженню ШНМ, стала однією із найперспективніших напрямків наукового вивчення. Причиною стала природа паралельних обчислень та здатність до адаптивного навчання нейронних мереж. Намагаючись відтворити роботу мозку людини, науковці спроектували прості апаратні, а згодом також програмні прототипи біологічних нейронів та програми їх з'єднання. Коли нейрофізіологи володіли більш глибоким розумінням функціонування нервової системи людини - такі ранні спроби почали сприймати як помилки. Однак у цій царині було досягнуто вражаючих результатів, які сприяли подальшим дослідженням і привели до проектування більш досконалих мереж [34, с. 3].

Разом із тим, вивченню людського мозку – тисячі років. З появою сучасної електроніки, почалися спроби апаратного відтворення процесу мислення. Першим етапом розробки та аналізу штучних нейронних мереж є наукова праця Дж. Маккалока та У. Питтса під назвою «Логічне обчислення ідей», що відносяться до нервової діяльності. В роботі сформульовано засадничі принципи створення ШНМ. Вагомий внесок у вивчення ШНМ зробила детекторна наука. Її зачинателем вважають Дж. Маккалока, який разом із групою своїх працівників у 1959 році видав статтю під назвою «Про що ока жаби говорять мозкові жаби», де вперше було введено поняття нейрона-детектора.

Водночас великим проривом в області нейроінтелекту стало створення нейрофізіологом одношарової НМ, що отримала назву перцептрон. Вона була застосована для досить широкого класу задач, серед яких: прогноз погоди, аналіз кардіограм і штучний зір. На рис. 1.2 представлено виклад перцептрон.

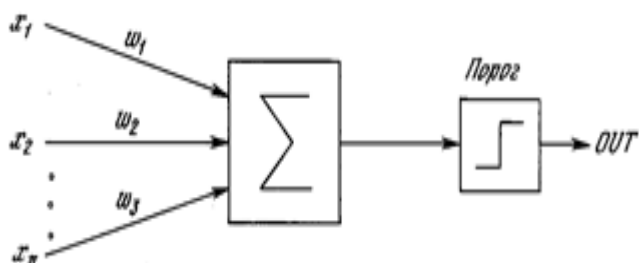


Рисунок 1.2 - Зовнішній вигляд перцептрон

Отже, перший перцептрон був створений Ф. Розенблатом у 60-х роках та користувався величезною популярністю. Згодом інтерес до нього змінився розчаруванням, адже з'ясувалося, що нейрон перцептрон не може навчитися виконувати елементарні задачі. М. Мінський суворо дослідив дану проблематику та довів, що існує ряд жорстких обмежень стосовно того, на що здатні одношарові нейрони перцептрон. Також продемонстрував, чому вони здатні навчитися. Оскільки в той період способи навчання НМ були майже невідомими, науковці перейшли до більш багатообіцяючої галузі, та дослідження в галузі НМ стали користуватися все меншею популярністю. Відкриття способів навчання багатшарових мереж в значній мірі, ніж який-небудь інший фактор, сприяло відродженню інтересу та дослідницьким зусиллям науковців [19, с. 22].

Отже, нейронні мережі являють собою набір пов'язаних між собою нейронних шарів, які отримують вхідну інформацію, оброблюють її та генерують результат, який видають на виході. Між вузлами видимих вхідного і вихідного шарів може бути наявною певна кількість прихованих шарів. Нейронні мережі здебільшого являють собою непрозору діяльність. Це значить, що побудована модель, як правило, не здатна чітко ітерпретуватися. Велика кількість пакетів ПЗ, які реалізують алгоритми НМ, використовуються в галузі оброблення

комерційних даних, в процесі розпізнавання образів, а також під час розшифровки рукописів та з метою читання кардіограм.

Характерно, що апаратні і програмні засоби реалізації алгоритмів НМ мають назву «нейрокомп'ютер». Нейрокомп'ютери є достатньо ефективними там, де потрібен аналог діяльності людського мозку, наприклад, інтуїції. Так, використовується з метою розпізнавання образів, можливості читати рукописні тексти, з метою підготування аналітичного прогнозування, перекладів на іншу мову. Саме для таких завдань досить важким є процес розробки алгоритму. ШНМ допомагають створити ефективне ПЗ для ПК з високим рівнем розпаралелювання оброблення (обчислень). З іншого боку, нейрокомп'ютери є дуже простими, на кшталт текстових процесорів, а через це з ними можна працювати навіть низькокваліфікованим користувачам [20, с. 77].

Крім того, інтелектуальні програми на базі ШНМ допомагають ефективно розв'язувати проблемні питання розпізнавання образів, прогнозування, оптимізаційних рішень, асоціативній пам'яті та управління. Традиційні способи стосовно вирішення даної проблеми не завжди приносить бажаний результат і тому велика кількість додатків виграють через застосування НМ. Штучні НМ - це своєрідні електронні моделі нейронної структури мозку людини, який, здебільшого вчиться на досвіді. Природне порівняння демонструє, що множина проблемних рішень, яка не піддається розв'язці за допомогою традиційних комп'ютерів, може ефективно вирішуватися завдяки нейромережам.

Варто відмітити, що при моделюванні прототипу мозку розглядаються локальні і глобальні засади пізнання його діяльності. Головною глобальною ознакою, що суттєво ускладнює процес проектування, є надзвичайно велике число базових компонентів структури. Людський мозок має в собі близько 100 млрд. нейронів, кожен з яких тисячами поєднань з'єднується з іншими НМ. Застосування навіть найпростішого прототипа нейрону не дає розробити прототип мозку, який би за своїми глобальними характеристиками наближався до

реального об'єкта проектування. До локальних ознак відносяться власне характеристики, за якими проектують модель НМ.

Наразі слід зауважити, що нейробіологія має в своєму арсеналі значні успіхи в питаннях дослідження НМ як елементарних структурних одиниць мозку. Відкрито величезне число закономірностей, які розкривають процеси його функціонування та зв'язки з іншими нейронами. Проте, як і раніше, лишаються проблеми реалізації таких властивостей мозку, як пам'ять і свідомість. Таким чином, до кінця не з'ясовано зв'язки між локальними ознаками нейронів та 12-тьма глобальними функціями людського мозку. Проте саме така взаємопов'язаність і є базою розробки ШНМ, що проектують мозкові функції. Через це головною проблемою концептуального рішення відносно проектування є проблема НМ є питання взаємодії компонентів відносно всіх рівнів функціонування. На переконання дослідників, тільки завдяки вдалій координації дій великого числа структурних компонентів можна досягти якісно нової характеристики моделі в цілому [22].

Варто додати, що успішний розвиток дисципліни ШНМ за останні роки дозволив впровадити комплекс таких глобальних характеристик, як:

1. навчання,
2. узагальнення,
3. абстрагування.

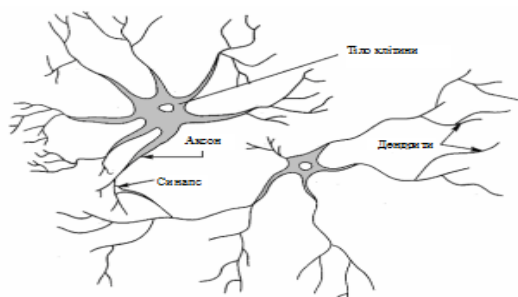


Рисунок 1.3. –Приклад біологічного нейрону

На рис. 1.3 представлено структуру кількох біологічних нейронів. Дендрити нейрона слідує від тіла нервової клітини до інших нейронів. Там нейрони отримують сигнали у відповідних синапсах. Сигнали, що приймаються синапсом, йдуть до тіла нейрону. Вже тут сигнали аналізуються, причому одні входи активують нейрон, а інші, навпаки, - знижують його активність. Коли сумарна активність нейрону переходить певну межу, нейрони переходять до більш активного стану, посилаючи по аксону, іншими словами «виход нейрону», сигнал іншим нейронам. У складі даної функціональної схеми існує багато спрощених процедур і виключень, однак переважна кількість ШНМ проектує виключно прості ознаки. Головними елементами нейронних мереж є нейрони або вузли, що по'єднані між собою певними зв'язками. Сигнали від них здатні передаватися за допомогою зважених зв'язкам, з кожним з яких з'язаний відповідний коефіцієнт чи вага.

Моделі нейронних мереж поділяються на:

- програмні моделі,
- апаратні моделі,
- моделі із застосування розпізнавання образів,
- моделі прогнозування,
- моделі проектування асоціативної пам'яті.

ШНМ імітують в першому наближенні ознаки біологічних нейронів. На вхід ШН поступає множинність сигналів, що є виходами інших нейронів. Кожен такий вхід множиться на значення відповідної ваги, що є аналогічною його синаптичній силі. Потім загальна кількість виходів підсумовується, визначається ступінь активації нейронів [17, с. 33].

Великий час еволюції додав людському мозоку багато властивостей, що відсутні в сучасних ПК з архітектурою фон Неймана. До таких властивостей слід зарахувати наступні [19, с. 8]:

- розподілене подання даних і паралельні обчислення,
- здатність до навчання й узагальнення,

- адаптивність,
- толерантність до похибки,
- низький рівень енергоспоживання.

Системи, які спроектовані на принципі функціонування біологічного нейрону, містять в собі перераховані властивості, що слід вважати істотним проривом в галузі оброблення даних. Досягнення в сфері нейрофізіології допомагають зрозуміти механізми природного мислення, в яких збереження даних відбувається за рахунок складного образу [12, с. 44].

Таким чином, процес збереження даних як образів, застосування образу та вирішення поставлених проблем формують нову галузь відносно оброблення інформації, що, не застосовуючи принципів класичного програмування, забезпечує розробку паралельних мереж та здатне їх навчати.

Отже, аналіз ШНМ перш за все пов'язаний з тим, що вони допомагають наблизитися до здатності оброблювати інформацію людським мозком, що являє собою надскладний, нелінійний, паралельний ПК системи оброблення даних. Мозок може організовувати власні компоненти, які називаються нейронами, таким чином, аби вони змогли виконати конкретні завдання на кшталт розпізнавання образів, оброблення сигналів органів чуття чи моторні функції в рази швидше, ніж це дозволяють зробити найшвидкодійні сучасні ПК. Наразі є багато прикладів застосування ШНМ для галузі прогнозування, класифікацій, оптимізаційних рішень і розпізнавання образів.

1.2 Класифікація нейронних мереж

Штучні нейронні мережі поділяються [17, с. 35]:

1) По типу вхідної інформації:

- аналогові (використання інформації в формі дійсних чисел);
- двійкові (інформація представлена в двійковому коді).

2) По характеру навчання:

- з «вчителем» (в процесі навчання відомо набір вихідних значень);
- без «вчителя» (формує вихідні значення тільки на основі вхідних впливів),

їх називають самоорганізуючими.

3) По характеру налаштування синапсів:

- з фіксованими зв'язками (вагові коефіцієнти вибираються на початковому етапі виходячи з даної задачі);
- з динамічними зв'язками (в процесі навчання відбувається настройка вагових коефіцієнтів).

4) За методом навчання:

- навчання по алгоритму зворотного розповсюдження помилки;
- з конкурентним навчанням;
- з навчанням по правилу Хебба;
- з гібридним навчанням.

5) По характеру зв'язків:

- з прямими зв'язками (розповсюджується інформація тільки в одному напрямку від рівня до рівня: це різноманітні персептронні мережі);
- із зворотнім поширенням інформації:
 - релаксаційні – циркуляція інформації відбувається доти, поки не перестають змінюватись вихідні значення НМ – це НМ Хопфілда, Хемінга;
 - багаторівневі мережі – в них відсутній процес релаксації. Це: рекурентні – в них існує зворотній зв'язок між входом і виходом, вихідне значення визначається як залежність вхідних і вихідних значень на попередньому кроці; рециркуляційні – характеризуються як прямим, так і зворотнім перетворенням інформації. Навчання відбувається без «вчителя», тобто вони самоорганізуються в процесі роботи[29, с. 1099].

Отже, сьогодні існують сотні нейронних мереж для вирішення задач, специфічних для різних виробничих галузей. Розглянемо різні типи базових нейронних мереж в порядку зростання складності. Ці типи нейронних мереж поділяються на[30, с. 234]:

1. Дрібні нейронні мережі (спільна фільтрація). Такі НМ складаються з груп перцептронів, що імітують нейронну структуру людського мозку. Дрібні НМ мають єдиний прихований шар перцептрону. Одним з поширених прикладів неглибоких НМ є спільна фільтрація. Прихований шар перцептрону буде навчений відображати подібності між сутностями для вироблення рекомендацій. Наприклад, система рекомендацій в Netflix, Amazon, YouTube Використовує версію спільної фільтрації, щоб рекомендувати свої продукти відповідно до інтересів користувачів.

2. Багатошаровий перцептрон (глибокі нейронні мережі). Нейронні мережі з більш ніж одним прихованим шаром називаються глибокими нейронними мережами. Усі наступні нейронні мережі є формою глибокої НМ, налаштованої / покращеною для вирішення проблем, специфічних для предметної галузі. В цілому вони допомагають нам домогтися універсальності. З огляду на достатню кількість прихованих шарів нейрона, глибока НМ може наближатися, тобто вона може вирішити будь-яку складну реальну проблему. Кожна версія глибокої нейронної мережі розробляється за допомогою повністю пов'язаного шару максимального об'єднаного продукту множення матриць, який оптимізується алгоритмами зворотного поширення (Рис. 1.4).

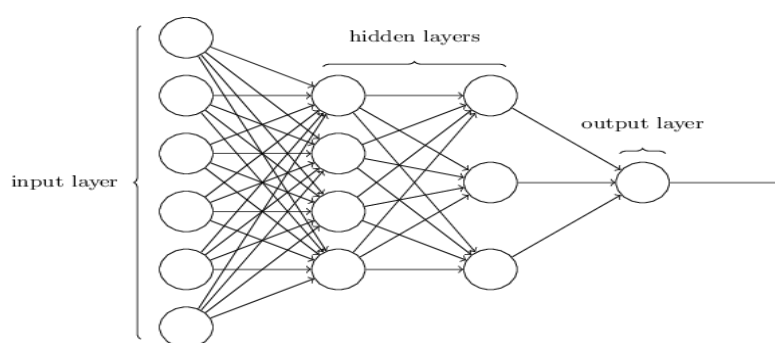


Рисунок 1.4 – Приклад багатошарового перцептрону

3. Згорткові нейронні мережі (ЗНМ) - це найбільш зріла форма глибоких нейронних мереж, що дозволяє отримувати найбільш точні, тобто кращі, ніж

людей, результати комп'ютерного зору. ЗНМ складаються з шарів згорток, створених шляхом сканування кожного пікселя зображень в наборі даних. У міру того, як дані апроксимуються шар за шаром, ЗНМ починає розпізнавати шаблони і тим самим розпізнавати об'єкти на зображеннях. Ці об'єкти широко використовуються в різних додатках для ідентифікації, класифікації. Недавні практики, такі як трансферне навчання в ЗНМ, привели до значних поліпшень в неточності моделей. Google Translator і Google Lens - найсучасніші приклади ЗНМ. Застосування ЗНМ є експоненціальним, оскільки вони навіть використовуються для вирішення проблем, які в першу чергу не пов'язані з комп'ютерним зором (Рис. 1.5).

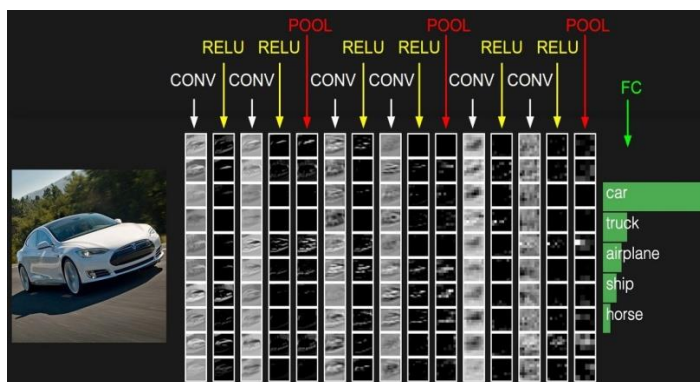


Рисунок 1.5. - Приклад ЗНМ

4. Рекурентна нейронна мережа (РНМ). РНМ - це новітня форма глибоких нейронних мереж для вирішення проблем в нейрон-лінгвістичному програмуванні. Простіше кажучи, РНМ передають вихідні дані кількох прихованих шарів назад на вхідний рівень для агрегування і перенесення наближення на наступну ітерацію (епоху) вхідного набору даних. Це також допомагає моделі самообучатися і в деякій мірі швидше коригувати прогнози. Такі моделі дуже корисні для розуміння семантики тексту в операціях нейрон-лінгвістичного програмування. Існують різні варіанти РНМ, такі як Long Short Term

Memory, Gated Recurrent Unit. На наведеній нижче діаграмі активація з h_1 і h_2 подається з входами x_2 і x_3 відповідно (Рис. 1.6).

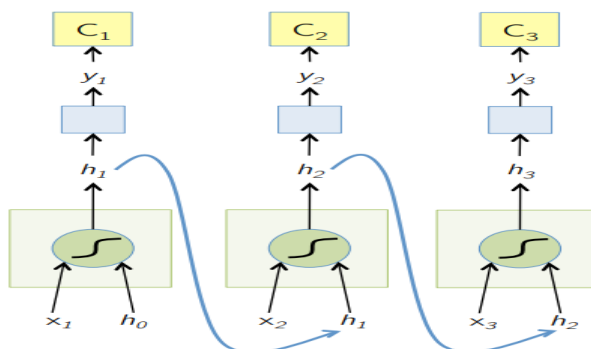


Рисунок 1.6 – Діаграма робочого процесу РНМ

5. Мережі довготривалої короткострокової пам'яті, що розроблені спеціально для вирішення проблеми зникаючих градієнтів в РНМ. Зникаючі градієнти трапляються з великими нейронними мережами, де градієнти функцій втрат мають тенденцію наближатися до нуля, змушуючи нейронні мережі припинитися для навчання. Такі мережі вирішують цю проблему, запобігаючи функції активації в своїх повторюваних компонентах і не змінюючи збережені значення (Рис. 1.7).

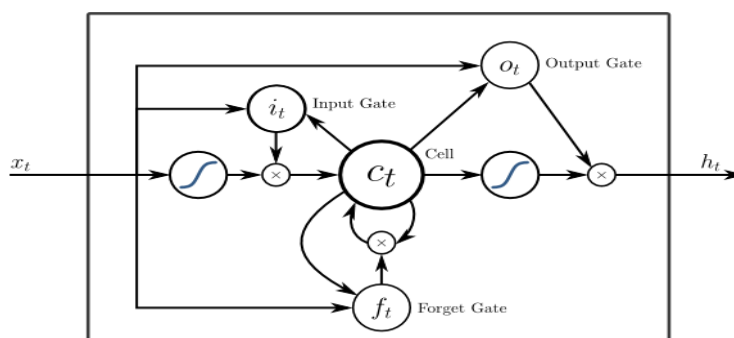


Рисунок 1.7 – Приклад мережі довготривалої короткострокової пам'яті

6. Нейронні мережі на основі уваги. Моделі уваги повільно беруть на себе навіть нові РНМ на практиці. Моделі уваги будуються шляхом зосередження на частини підмножини інформації, яку їм дають, тим самим усуваючи переважна кількість фонові інформації, яка не потрібна для поточного завдання. Моделі уваги побудовані на поєднанні принципів застосування уваги і підгонки шляхом поширення такого принципу. Ієрархічно складена модель множинної уваги називається трансформатором. Ці перетворювачі більш ефективні для паралельного запуску стеків, так що вони дають сучасні результати з порівняно меншими даними і часом для навчання моделі. Розподіл уваги стає дуже потужним при використанні з ЗНМ / РНМ і може створювати текстовий опис зображення, як показано нижче (Рис. 1.8).



Рисунок 1.8 – Приклад текстового опису зображення за допомогою ШНМ

7. Генеративна змагальна мережа. Хоча моделі глибокого навчання забезпечують найсучасніші результати, їх можуть обдурити набагато розумніші люди, додавши шум до реальних даних. Генеративна змагальна мережа - це остання розробка в галузі глибокого навчання для вирішення таких сценаріїв. Такі мережі використовують навчання без учителя, коли глибокі нейронні мережі навчаються з використанням даних, створених за допомогою моделі II, разом з фактичним набором даних, щоб підвищити точність і ефективність моделі. Ці змагальні дані в основному використовуються, щоб обдурити дискримінаційну модель, щоб побудувати оптимальну модель. Результуюча модель має тенденцію бути кращим наближенням, ніж може подолати такий шум (Рис. 1.9).

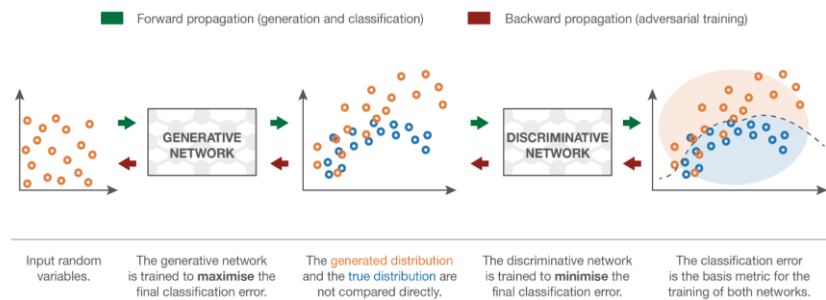


Рисунок 1.9 – Приклад генеративної змагальної мережі

Отже, глибокі нейронні мережі розширюють межі ПК. Вони не обмежуються тільки класифікацією ЗНМ та РНМ або прогнозами та навіть здатні генерувати дані. Ці дані можуть варіюватися від гарної форми мистецтва до спірних глибинних підробок, але кожен день вони перевершують людей по завданню. Отже, важливим завданням є усвідомлення впливу ШНМ та створення ефективної моделі ШНМ для конкретних задач.

1.3 Алгоритми навчання нейронних мереж

Нейронні мережі навчаються (або навчаються) на прикладах обробки, кожен з яких містить відомі «вхідні дані» і «результат», формуючи між ними і з врахуванням ймовірності асоціації, які зберігаються в структурі даних самої мережі. Навчання нейронної мережі з даного прикладу зазвичай проводиться шляхом визначення різниці між обробленим висновком мережі (часто пророкуванням) і цільовим висновком помилка. Потім мережа коригує свої виважені асоціації відповідно до правила навчання і з використанням цього значення помилки [21].

Послідовні настройки приведуть до того, що нейронна мережа буде видавати результат, який стає все більш схожим на цільовий. Після достатньої кількості цих коригувань навчання може бути припинено за певними критеріями.

Такі системи «вчать» виконувати завдання, розглядаючи приклади, як правило, без програмування конкретних правил. Наприклад, при розпізнаванні

зображень вони можуть навчитися ідентифікувати зображення, що містять кішок, шляхом аналізу прикладів зображень, які були вручну позначені як «кішка» або «без кішок», і використання результатів для ідентифікації кішок на інших зображеннях. Вони роблять це, не знаючи заздалегідь про кішок, наприклад, про те, що у них шерсть, хвости, вуса і котячі особи. Замість цього вони автоматично генерують ідентифікують характеристики із прикладів, які вони обробляють. Нейрони зазвичай діляться на кілька рівнів, особливо при глибокому навчанні. Нейрони одного шару з'єднуються тільки з нейронами безпосередньо передус і безпосередньо наступного шарів. Шар, який отримує зовнішні дані, є вхідним шаром. Шар, який дає остаточний результат, - це вихідний шар. Між ними нуль або більше прихованих шарів. Також використовуються одношарові і неслойні мережі [24, с. 21].

Між двома рівнями можливі кілька схем з'єднання. Вони можуть бути повністю пов'язані, при цьому кожен нейрон одного шару з'єднується з кожним нейроном наступного шару. Вони можуть об'єднуватися, де група нейронів в одному шарі з'єднується з одним нейроном в наступному шарі, тим самим зменшуючи кількість нейронів в цьому шарі. [Нейрони тільки з такими зв'язками утворюють орієнтований ациклічний граф і відомі як мережі прямого поширення.

З іншого боку, мережі, які дозволяють з'єднання між нейронами в одному або попередніх шарах, відомі як повторювані мережі.

Навчання - це адаптація мережі для кращого вирішення завдання з урахуванням вибіркового спостережень. Навчання включає в себе настройку ваг (і необов'язкових порогових значень) мережі для підвищення точності результату. Це досягається за рахунок мінімізації спостережуваних помилок. Навчання завершено, коли вивчення додаткових спостережень не знижує частоту помилок. Навіть після навчання частота помилок зазвичай не досягає 0. Якщо після навчання частота помилок занадто висока, мережа зазвичай необхідно перепроєктувати.

Нижче наведено традиційну схему алгоритму навчання ШНМ (Рис. 1.10).

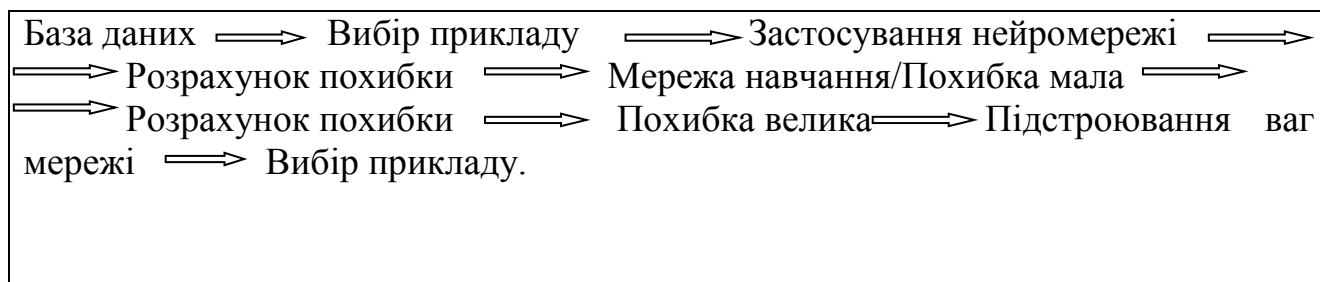


Рисунок 1.10 – Схема навчання НМ

Практично це досягається шляхом визначення функції витрат, яка періодично оцінюється під час навчання. Поки його обсяг виробництва продовжує знижуватися, навчання триває. Вартість часто визначається як статистика значення якого можна тільки приблизно визначити. Насправді вихідними даними є числа, тому, коли помилка мала, різниця між результатом (майже напевно кішка) і правильною відповіддю (кішка) невелика. Навчання намагається зменшити загальну різницю між спостереженнями. Більшість моделей навчання можна розглядати як пряме додаток теорії оптимізації і статистичної оцінки [6, с. 8].

Швидкість навчання визначає розмір коригувальних кроків, які модель робить для коригування помилок в кожному спостереженні. Висока швидкість навчання скорочує час навчання, але з меншою кінцевою точністю, тоді як більш низька швидкість навчання займає більше часу, але з потенціалом для більшої точності. Оптимізація, така як Quicksrpor, в першу чергу націлена на прискорення мінімізації помилок, в той час як інші поліпшення в основному спрямовані на підвищення надійності.

Щоб уникнути коливань всередині мережі, таких як чергування ваг з'єднань, і для підвищення швидкості збіжності, уточнення використовують адаптивну швидкість навчання, яка збільшується або зменшується в міру необхідності.

Концепція імпульсу дозволяє зважувати баланс між градієнтом і попереднім зміною, так що коригування ваги залежить в деякій мірі від попередньої зміни.

Імпульс, близький до 0, підкреслює градієнт, а значення, близьке до 1, підкреслює остання зміна.

У навчанні з учителем використовується набір парних входів і бажаних виходів. Завдання навчання - зробити бажаний результат для кожного входу. У цьому випадку функція вартості пов'язана з усуненням неправильних відрахувань. Зазвичай використовується вартість - це середньоквадратична помилка, яка намагається мінімізувати середньоквадратичнепомилку між виходом мережі і бажаним виходом. Завдання, які підходять для навчання з учителем, - це розпізнавання образів (також відоме як класифікація) і регресія (також відома як апроксимація функцій).

Навчання з вчителем також можна застосувати до послідовним даними (наприклад, для розпізнавання рукописного введення, мови і жестів). Це можна розглядати як навчання з «учителем» у формі функції, яка забезпечує постійний зворотний зв'язок про якість рішень, отриманих на даний момент [36, с. 110].

У таких додатках, як гра у відеоігри, актор виконує ряд дій, отримуючи зазвичай непередбачуваний відповідь від навколишнього середовища після кожного з них. Мета полягає в тому, щоб виграти ГРУ. Отримати найбільш позитивні (з найменшими витратами) відгуки. У навчанні з підкріпленням мета полягає в тому, щоб зважити мережу (розробити політику) для виконання дій, які мінімізують довгострокові (очікувані сукупні) витрати. У кожен момент часу агент виконує дію, а середовище генерує спостереження і миттєву вартість відповідно до деякими (зазвичай невідомими) правилами.

Правила та довгострокову вартість зазвичай можна тільки оцінити. У будь-який момент агент вирішує, чи слід досліджувати нові дії, щоб розкрити свої витрати, або використовувати попереднє навчання, щоб діяти швидше.

ШНМ служать навчальним компонентом в таких додатках. Динамічне програмування в поєднанні з ШНМ (що забезпечує нейродинамічний програмування) застосовувалося до таких проблем, як маршрутизація транспортних засобів, відеоігри, управління природними ресурсами і медицина

через здатність ШНМ знижувати втрати точності навіть при зменшенні щільності дискретизації сітки для чисельної апроксимації рішення задач управління. Завдання, які підпадають під парадигму навчання з підкріпленням, - це завдання управління, ігри та інші завдання послідовного прийняття рішень.

Самонавчання в нейронних мережах було введено в 1982 році разом з нейронною мережею, здатної до самонавчання, під назвою Crossbar Adaptive Array (CAA). Це система тільки з одним входом, ситуацією s і тільки одним виходом, дією (або поведінкою). У нього немає ні зовнішнього ради, ні зовнішнього підкріплення з навколишнього середовища. CAA перехресним чином обчислює як рішення про дії, так і емоції (почуття) з приводу виниклих ситуацій.

Система управляється взаємодією пізнання і емоцій.

У ситуації s скористайтеся засобами a ;

Отримайте наслідок ситуації s' ;

Обчислити емоцію перебування в ситуації наслідки $v(s')$;

Оновити пам'ять поперечини:

$$w'(a, s) = w(a, s) + v(s').$$

Цінність зворотного поширення (вторинне підкріплення) - це емоція по відношенню до ситуації наслідків. CAA існує в двох середовищах: одна - це поведінкова середовище, в якому вона поводить себе, а інша - генетична среда, звідки вона спочатку і тільки один раз отримує початкові емоції, які можуть виникнути в поведінковій середовищі. Отримавши вектор генома (вектор видів) з генетичної середовища, CAA вивчатиме цілеспрямовану поведінку в поведінковій середовищі, яка містить як бажані, так і небажані ситуації.

Доступні два режими навчання: стохастичний і пакетний. У стохастичному навчанні кожен вхід створює коригування ваги. При пакетному навчанні ваги коректуються на основі пакету вхідних даних, накопичуючи помилки по пакету. Стохастичне навчання вносить «шум» в процес, використовуючи локальний

градієнт, розрахований з однієї точки даних; це знижує ймовірність застрявання мережі в локальних мінімумах.

Цінність зворотного поширення (вторинне підкріплення) - це емоція по відношенню до ситуації наслідків. САА існує в двох середовищах: одна - це поведінкова середовище, в якому вона поводить, а інша - генетична среда, звідки вона спочатку і тільки один раз отримує початкові емоції, які можуть виникнути в поведінковій середовищі. Отримавши вектор генома (вектор видів) з генетичної середовища, САА вивчатиме цілеспрямовану поведінку в поведінковій середовищі, яка містить як бажані, так і небажані ситуації.

Доступні два режими навчання: стохастичний і пакетний. У стохастичному навчанні кожен вхід створює коригування ваги. При пакетному навчанні ваги коректуються на основі пакету вхідних даних, накопичуючи помилки по пакету. Стохастичне навчання вносить «шум» в процес, використовуючи локальний

Однак пакетне навчання зазвичай дає більш швидкий і стабільний спуск до локального мінімуму, оскільки кожне оновлення виконується в напрямку середньої помилки пакета. Поширеним компромісом є використання «міні-партій», невеликих партій з вибірками в кожній партії, обраними стохастически з усього набору даних [30, с. 86].

Отже, варто констатувати, що всі дані, яку мережа має в якості задачі, утримується в певній сукупності прикладів. Через це якість навчання НМ безпосередньо залежить від числа прикладів у вибірці навчання, а також від того, наскільки повно дані приклади характеризують конкретне завдання. Так, наприклад, немає сенсу застосовувати НМ для прогнозів фінансової кризи, якщо в навчальній вибірці таких криз не існує. На думку дослідників, для повноцінного тренування потрібно мати біля сотні прикладів. Таким чином, навчання мережі – це досить непростий та наукомісткий процес. Алгоритми навчання здебільшого мають різні властивості та налаштування, для управління якими необхідно розуміти їхній вплив.

1.4 Постановка задачі дослідження

Метою є проектування системи розпізнавання символів за допомогою нейронної мережі.

Ставимо перед собою такі завдання:

1. розглянути теоретичні засади навчання нейронних мереж;
2. дослідити класифікацію нейронних мереж;
3. визначити алгоритми навчання нейронних мереж;
4. проаналізувати особливості навчання нейронних мереж для СРС;
5. розробити та застосувати алгоритм ШНМ для СРС

Висновки до розділу 1

Підсумовуючи перший розділ, можна зробити наступні висновки:

1. Визначено теоретичні засади навчання нейронних мереж, окреслено поняття «нейронна мережа», з'ясовано його сутність.
2. Деталізувано історичний розвиток теорії ШНМ.
3. Проведено аналогію між біологічними та штучними мережами.
4. Окреслили класифікацію НМ.
5. Досліджено алгоритми навчання нейронних мереж.

2 ОСОБЛИВОСТІ НАВЧАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ

2.1 Специфіка застосування нейронних мереж

Нейронні мережі не є універсальними для вирішення всіх обчислювальних проблем. Комп'ютери та обчислювальні способи в більшості випадків є ідеальними для реалізації. Комп'ютерні технології перевершують людину в питаннях математичних обчислень та символічних обчислень. Проте людина здатна без особливих зусиль вирішувати доволі складні завдання сприйняття зовнішньої інформації з такою швидкістю та точністю, з якою жоден ПК у на її фоні буде виглядати тугодумом.

Володіючи людською здатністю вирішувати проблеми і застосовувати цю навичку до величезних наборів даних, нейронні мережі мають такі потужними атрибутами:

- Адаптивне навчання: як і люди, нейронні мережі моделюють нелінійні і складні відносини і спираються на попередні знання. Наприклад, програмне забезпечення використовує адаптивне навчання для навчання математики та мовним мистецтвам.

- Самоорганізація: здатність кластеризувати і класифікувати величезні обсяги даних робить нейронні мережі унікальними для організації складних візуальних проблем, що виникають при аналізі медичних зображень.

- Робота в реальному часі: нейронні мережі можуть (іноді) надавати відповіді в реальному часі, як у випадку з безпілотними автомобілями і навігацією дронів.

- Прогноз: здатність НМ прогнозувати на основі моделей має широкий спектр додатків, в тому числі для погоди і дорожнього руху.

- Відмовостійкість: коли значна частина мережі втрачена або відсутній, нейронні мережі можуть заповнити прогалини. Ця здатність особливо корисна при освоєнні космосу, де завжди можлива відмова електронних пристроїв.

Розглянемо предметніше галузі використання ШНМ для розв'язання емпіричних завдань.

Нейронні мережі дуже цінні, тому що вони можуть виконувати завдання з аналізу даних, зберігаючи при цьому всі свої інші атрибути. Ось важливі завдання, які виконують нейронні мережі [17, с. 33]:

- Класифікація: мережеві мережі організують шаблони або набори даних в заздалегідь визначені класи.

- Прогноз: вони виробляють очікуваний результат на основі заданих вхідних даних.

- Кластеризація: вони ідентифікують унікальну особливість даних і класифікують її, не знаючи попередніх даних.

- Зв'язування: ви можете навчити НМ «запам'ятовувати» шаблони. Коли ви показуєте незнайому версію патерну, мережа пов'язує її з найбільш порівнянної версією в своїй пам'яті і повертається до останньої. Нейронні мережі мають фундаментальне значення для глибокого навчання, надійного набору технік НМ, які дозволяють вирішувати абстрактні проблеми, такі як біоінформатика, розробка ліків, фільтрація соціальних мереж і переклад на природну мову.

- Глибоке навчання - це те місце, де ми вирішуємо найскладніші завдання в науці і техніці, включаючи передову робототехніку. У міру того як нейронні мережі стають розумнішими і швидше, ми щодня вдосконалюємося.

Ось список інших додатків для проектування нейронних мереж, які в даний час використовуються в різних галузях:

- Аерокосмічна промисловість: детектори і моделювання несправностей компонентів літака, системи управління літаком, високопродуктивне автопілотування і моделювання траєкторії польоту.

- Автомобільна промисловість: вдосконалені системи наведення, розробка силових агрегатів, віртуальних датчиків і аналізаторів гарантійного обслуговування. Електроніка: аналіз відмов мікросхеми, схеми мікросхем, машинне зір, нелінійне моделювання, прогнозування кодової послідовності, управління процесом і синтез голосу.

- Виробництво: аналіз конструкції хімічної продукції, динамічне моделювання систем хімічних процесів, управління процесами, діагностика процесів і машин, проектування і аналіз продукції, прогнозування якості паперу, торги по проектам, планування і управління, аналіз якості комп'ютерних мікросхем, системи візуального контролю якості та аналіз якості зварювання.

- Механіка: моніторинг стану, моделювання систем і управління.

- Робототехніка: роботи-навантажувачі, контролери маніпуляторів, системи управління траєкторією і системи технічного зору.

- Телекомунікації: управління мережею банкоматів, автоматизовані інформаційні послуги, системи обробки платежів клієнтів, стиснення даних, еквайзери, усунення несправностей, розпізнавання рукописного введення, проектування мережі, управління, маршрутизація і контроль, моніторинг мережі, переклад розмовної мови в реальному часі і розпізнавання образів (особи, об'єкти, відбитки пальців, семантичний аналіз, перевірка орфографії, обробка сигналів і розпізнавання мови).

Крім того, існують додаткові поточні приклади бізнес-додатків ШНМ. Серед них [28, с. 105]:

- Банківська справа: виснаження кредитних карт, оцінка кредитних і кредитних заявок, оцінка шахрайства і ризиків, а також прострочення по кредитах.

- Бізнес-аналітика: моделювання поведінки клієнтів, сегментація клієнтів, схильність до шахрайства, дослідження ринку, структура ринку, структура ринку і моделі відсіву, дефолту, покупок і продовжень.

- Захист: боротьба з тероризмом, розпізнавання осіб, виділення функцій, придушення шуму, розпізнавання об'єктів, датчики, гідролокатор, радар і обробка сигналів зображення, ідентифікація сигналу / зображення, відстеження цілі і управління зброєю.

- Освіта: програмне забезпечення для адаптивного навчання, динамічне прогнозування, аналіз і прогнозування системи освіти, моделювання успішності учнів і профілювання особистості.

- Фінанси: рейтинги корпоративних облігацій, корпоративний фінансовий аналіз, аналіз використання кредитної лінії, прогнозування курсів валют, консультування по кредитах, перевірка іпотечних кредитів, оцінка нерухомості і торгівля портфелями.

- Медицина: аналіз ракових клітин, аналіз ЕКГ і ЕЕГ, консультації по тестуванню в відділенні невідкладної допомоги, скорочення витрат і підвищення якості лікарняних систем, оптимізація процесу трансплантації і дизайн протеза.

Цінні папери: автоматичний рейтинг облігацій, аналіз ринку і консультаційні системи з торгівлі акціями.

- Транспорт: системи маршрутизації, системи діагностики гальм вантажівок і планування транспортних засобів.

Перспективами розробки нейромережевих технологій можуть бути наступні сфери [37, с. 66]:

- Інтеграція нечіткої логіки: нечітка логіка розпізнає більше, ніж просто справжні і несправжні значення - вона бере до уваги поняття, які є відносними, наприклад, щось, іноді і зазвичай. Нечітка логіка і нейронні мережі інтегровані для використання в самих різних областях, таких як відбір кандидатів на вакансії, автомобільна інженерія, управління будівельними кранами і моніторинг глаукоми. Нечітка логіка стане важливою функцією в майбутніх додатках нейронних мереж.

- Імпульсні нейронні мережі: нещодавно дані нейробиологічних експериментів прояснили, що біологічні нейронні мережі ссавців з'єднуються і

взаємодіють за допомогою пульсації і використовують синхронізацію імпульсів для передачі інформації і виконання обчислень. Це визнання прискорило значні дослідження, включаючи теоретичний аналіз, розробку моделей, Нейробіологічні моделювання та розгортання обладнання, спрямованих на те, щоб зробити обчислення ще більш схожими на те, як функціонує наш мозок.

- Спеціалізоване обладнання: в даний час спостерігається бурхливий ріст розробки обладнання, яке прискорить і в кінцевому підсумку знизить вартість нейронних мереж, машинного навчання та глибокого навчання. Відомі компанії і стартапи прагнуть розробити поліпшені мікросхеми та блоки графічної обробки, але справжня новина - це швидкий розвиток блоків обробки нейронних мереж та іншого обладнання для II, разом іменуються нейросінаптичеські архітектурою. Нейросінаптичні чіпи мають фундаментальне значення для розвитку III, тому що вони більше схожі на біологічний мозок, ніж на ядро традиційного комп'ютера. Зі свого технологією Brain Power компанія IBM є лідером в розробці нейросінаптичеських чіпів. На відміну від стандартних мікросхем, які працюють безперервно, мікросхеми Brain Power управляються подіями і працюють в міру необхідності. Технологія об'єднує пам'ять, обчислення і зв'язок.

- Поліпшення існуючих технологій: завдяки новому програмному забезпеченню та обладнанню, а також поточним технологій нейронних мереж і збільшеною обчислювальної потужності нейросінаптичеських архітектур нейронні мережі тільки почали показувати, на що вони здатні. Безліч бізнес-додатків більш швидких, дешевих і більш людських методів вирішення проблем і поліпшених методів навчання дуже прибуткові.

- Робототехніка: було безліч прогнозів щодо роботів, які зможуть відчувати себе, як ми, бачити як ми, і робити прогнози щодо світу навколо них.

Перш ніж ці системи зможуть по-справжньому мислити рухомим, непорушним чином, має відбутися так багато речей. Один з найважливіших факторів - це здатність встановлювати і діяти відповідно до самовизначення цінностями в режимі реального часу, що ми, люди, робимо тисячі разів в день. Без

цього ці системи будуть давати збій кожен раз, коли умови виходять за межі заздалегідь заданій області.

Отже, нові можливості обчислень потребують нових умінь розробника поза межами класичних обчислень. Спочатку обчислення були виключно апаратними та програмісти зробили їх функціональними. Згодом, були фахівці з ПЗ: системні інженери, фахівці з баз даних та проектувальники. Зараз вже існують нейронні архітектори. Новий кадровий склад спеціалістів повинен мати відповідні кваліфікаційні вміння, яка має бути вищою за їх попередників. Наприклад, вони мають володіти знаннями відносно статистики для вибору та оцінки навчальної та тестової множини. Зазвичай, при проектуванні ефективних ШНМ, важливим для сучасних програмістів є логічне мислення та досвід.

2.2 Огляд програмних засобів, що використовують нейронні мережі

На сьогоднішній день існує понад двохсот нейропакетів, які випускаються багатьма компаніями та окремими науковцями та допомагають конструювати, навчати і застосовувати ШНМ для вирішення практичних задач. Трудомісткість проектування НМ скорочується при використанні готового нейромережевого ПЗ. Нижче перерахуємо найбільш поширені програми для ШНМ [18, с. 123]:

1. Neural Designer - це настільний додаток для інтелектуального аналізу даних, в якому використовуються нейронні мережі, які є парадигмою машинного навчання. Нейронні мережі в нейронних дизайні - це математичні моделі функцій мозку, обчислювальні моделі, натхненні центральними нервовими системами мозку, які можна навчити для виконання певних завдань.

Neural Designer володіє більшістю передових методів підготовки даних, машинного навчання та розгортання моделей. Його візуальний графічний користувацький інтерфейс забезпечує вичерпні та наочні результати без необхідності написання коду або збірки блоків. Програмне забезпечення реалізує багатоядерну обробку для аналізу великих обсягів даних за менший час.

2. Neuroph - це проект з відкритим вихідним кодом, розміщений на SourceForge під ліцензією Apache. Це бібліотека для створення нейронних мереж і використання машинного навчання. Neuroph - це полегшена структура нейронної мережі Java для розробки загальних архітектур нейронних мереж. Користувачі можуть взаємодіяти з Neuroph, використовуючи:

- Інструмент на основі графічного інтерфейсу користувача
- Бібліотека Java.

Обидва підходи засновані на ієрархії класів, яка будує штучні нейронні мережі з шарів нейронів.

Neuroph містить красивий графічний редактор нейронної мережі для швидкого створення компонентів нейронної мережі Java. Програмне забезпечення спрощує розробку нейронної мережі, надаючи бібліотеку нейронної мережі Java і інструмент з графічним інтерфейсом користувача, який підтримує створення, навчання і збереження нейронних мереж.

3. Darknet - це фреймворк нейронної мережі з відкритим вихідним кодом, написаний на C і CUDA і підтримує обчислення CPU і GPU. Це сверточное нейронна мережа, що складається з дев'ятнадцяти шарів. Попередньо навчена мережа може класифікувати зображення по 1000 категоріям об'єктів, таким як клавіатура, миша, олівець і багато тварин. У результаті мережа навчилася відображати багаті можливості для широкого діапазону зображень. Darknet встановлюється тільки з двома обов'язковими залежностями, такими як OpenCV, якщо користувачам потрібно ширший спектр підтримуваних типів зображень, або CUDA, якщо їм потрібні обчислення на GPU. Користувачі можуть почати з установки базової системи, яка була протестована тільки на комп'ютерах Linux і Mac.

4. Keras - це бібліотека глибокого навчання для Theano і TensorFlow. Бібліотека нейронних мереж високого рівня написана на Python і здатна працювати поверх обох додатків. Keras - це API, розроблений для людей, а не для машин. Програмне забезпечення слід передовим методам зниження когнітивної

навантаження. Він пропонує послідовні і прості API-інтерфейси і зводить до мінімуму кількість дій користувача, необхідних для типових випадків використання.

Keras надає чіткі і корисні повідомлення про помилки, а також велику документацію і керівництва для розробників. Бібліотека глибокого навчання Keras дозволяє легко і швидко створювати прототипи завдяки повній модульності, мінімалізму і розширюваності. Він підтримує згорткові нейронні мережі та рекурентні мережі, а також їх комбінації.

5. NeuroSolutions - це середовище розробки програмного забезпечення для нейронних мереж, розроблена NeuroDimension. Він поєднує в собі модульний інтерфейс мережевого дизайну на основі значків з реалізацією розширених процедур навчання, таких як пов'язані градієнти, Levenberg Marquardt і зворотне поширення в часі. Сімейство продуктів NeuroSolutions - це передове програмне забезпечення нейронних мереж для інтелектуального аналізу даних для створення високоточних і прогнозних моделей з використанням передових методів обробки, інтелектуального автоматичного пошуку топології нейронної мережі за допомогою передових розподілених обчислень.

Це дизайнерський інтерфейс з передовим штучним інтелектом і алгоритмами навчання з використанням інтуїтивно зрозумілих майстрів або простого у використанні інтерфейсу Excel. Програмне забезпечення надає три окремих майстра для автоматичної побудови моделей нейронних мереж:

- Менеджер даних,
- Neural Builder,
- Neural Expert.

6. Tflearn - це модульна і прозора бібліотека глибокого навчання, побудована на основі Tensorflow. Програмне забезпечення призначене для надання високорівневого API для TensorFlow, щоб полегшити і прискорити експерименти, залишаючись при цьому повністю прозорим і сумісним з ним. API високого рівня в даний час підтримує більшість останніх моделей глибокого

навчання, таких як Convolutions, LSTM, BiRNN, BatchNorm, PreLU, залишкові мережі і генеративні мережі. Заглядаючи в майбутнє, Tflern також призначений для того, щоб залишатися в курсі останніх технологій глибокого навчання, і в даний час він знаходиться на ранній стадії розробки.

7. ConvNetJS Бібліотека ConvNetJS дозволяє користувачам створювати і вирішувати нейронні мережі на Javascript. Спочатку він був написаний аспірантом зі Стенфорда. З тих пір бібліотека була розширена за рахунок внесків спільноти. Користувачі відчувають себе комфортно, оскільки мережа навчається, просто відкриваючи її без вимог до програмного забезпечення, компіляторів, установок, графічного процесора і поту. Код доступний на Github під ліцензією MIT.

8. Torch - це наукова обчислювальне середовище з широкою підтримкою алгоритмів машинного навчання, які ставлять графічні процесори на перше місце. Простий і швидкий мова сценаріїв LuaJIT і лежить в його основі реалізація C/CUDA пропонують користувачам просту у використанні і ефективну програму. Користувачі Torch зможуть скористатися перевагами його основних функцій, таких як потужний N-мірний масив, безліч підпрограм для індексації, нарізки, транспортування, приголомшливий інтерфейс для C через LuaJIT, підпрограми лінійної алгебри, нейронні мережі і енергоспоживання. моделі.

9. ЦИФРИ NVIDIA. Система NVIDIA Deep Learning GPU Training (ЦИФРИ) поставити владу глибокого вивчення в руки інженерів і вчених даних. Програмне забезпечення використовує високоточні глибокі нейронні мережі для швидкої класифікації зображень, сегментації і виявлення об'єктів. DIGITS спрощує спільні завдання глибокого навчання, такі як управління даними, проектування і навчання нейронних мереж в системах з декількома графічними процесорами, моніторинг продуктивності в режимі реального часу за допомогою розширених візуалізацій і вибір найбільш ефективною моделі з браузера результатів для розгортання. Він повністю інтерактивний, тому фахівці з даними можуть зосередитися на проектуванні та навчанні мереж, а не на програмуванні та налагодженні.

10. Симулятор нейронної мережі Штутгарта Stuttgart Neural Network Simulator (SNNS) - це нейронний симулятор, спочатку розроблений в Штутгарту університеті. Спочатку він був побудований для X11 під Unix, пізніше - JavaNNS. Симулятор SNNS складається з двох основних компонентів:

- Ядро симулятора написано на C.
- Графічний користувальницький інтерфейс під X11R4 або X11R5.

Проект має на меті створити ефективну і гнучку середу моделювання для дослідження та застосування нейронних мереж. Ядро симулятора оперує внутрішніми мережевими структурами даних нейронних мереж і виконує всі операції навчання і виклику. Його також можна використовувати без інших частин як програму C, вбудовану в призначені для користувача програми. Він підтримує довільні мережеві топології і, як RCS, підтримує концепцію сайтів. SNNS може бути розширений користувачем за допомогою визначених користувачем функцій активації, функцій виведення, функцій сайту і процедур навчання, які написані як прості програми на мові C і пов'язані з ядром симулятора.

Отже, перевагою вищеперерахованого програмного забезпечення є простота розробки та зрозумілість процесу виведення, відсутність проблемних рішень відносно редагування. В якості недоліку можна відзначити високу вартість програмного забезпечення.

2.3 Формалізація кластерного аналізу та адаптація методів сегментації до умов задачі

Кластерний аналіз спрямований на виявлення природного поділу об'єктів. Іншими словами, він групує спостереження, які подібні до підмножин. Такі підкласи виявляють закономірності, що безпосередньо мають відношення досліджуваного явища. Функціональна особливість виміру відстані

використовується в якості аналізу, якщо схожість між об'єктами і широким спектром кластеризації алгоритмів на основі по різних поняттям в наявності.

Заходи подібності спочатку обчислюються між спостереженнями і між кластерами, коли спостереження починають групуватися в кластери. Кілька метрик, таких як евклидово і Манхеттенський відстань, кореляція або взаємна інформація, можуть використовуватися для обчислення подібності. Крім того, можливі кілька стратегій злиття, які призводять до різних шаблонів кластеризації. Тому результати кластеризації дещо суб'єктивні, оскільки вони сильно залежать від вибору користувачів [8, с. 56].

Традиційний кластерний аналіз зазвичай проводиться для групування спостережень або змінних окремо, але одночасна кластеризація (або бікластеризація) рядків і стовпців матриці даних також є підходящою альтернативою для пошуку біомаркерів. Оскільки він використовує ієрархічну конфігурацію - дерево, зване Дендрограммою. У цьому випадку функція зв'язку визначає критерії оцінки відстаней між спостереженнями і кластерами. На кожній ітерації найближчі об'єкти групуються для формування нового кластера.

Детальніше розглянемо формалізацію процесу кластерного аналізу. Існує X - об'єктів, та Y кластерів, а також з'ясування відстані між об'єктами. Вибірка для навчання об'єктів БД. Мета - поділ заданої вибірки на своєрідні підмножини, які не мають змоги пересікатися. Такі підмножини являють собою кластери, де кожен кластер складається з об'єктів при застосуванні найменшої відстані між собою \min . Об'єкти таких груп мають відрізнятися одне від одного. Далі кожному об'єкту надається власний номер групи.

Потрібно відмітити, що кластеризація являє собою функцію $X \rightarrow Y$, в якій для кожного об'єкта x відповідає номер групи y , що відомий наперед.

Процес кластеризації відносять до способів навчання без вчителя. Кластеризація відрізняється від класифікації, наприклад, тим, що вихідні об'єкти в цьому процесі можуть не задаватися, тому має існувати велика кількість Y .

Рішення задач методу кластеризації є складною задачею. На це існують такі причини [10, с. 56]:

1. не існує чіткого критерію якості кластеризації, однак відома величезна кількість критеріїв, способів, що роблять кластеризацію ефективною. Цікавим фактом також є те, що різні способи дають також різні показники;

2. на показники кластеризації впливає сильно метричність.

Процес роботи кластеризації представлено на рисунку 2.1.

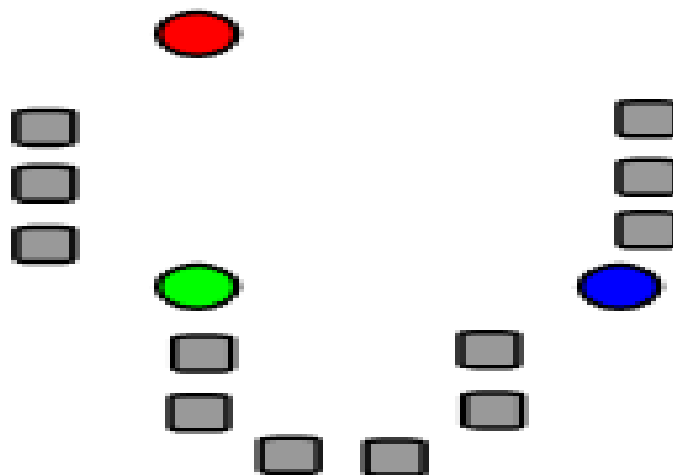


Рисунок 2.1 - Вихідні точки і випадково вибрані початкові точки

На рисунку 2.2. Представлено першу ітерацію та віднесено кожний об'єкт до відповідної групи, за одним із способів визначення відстані.

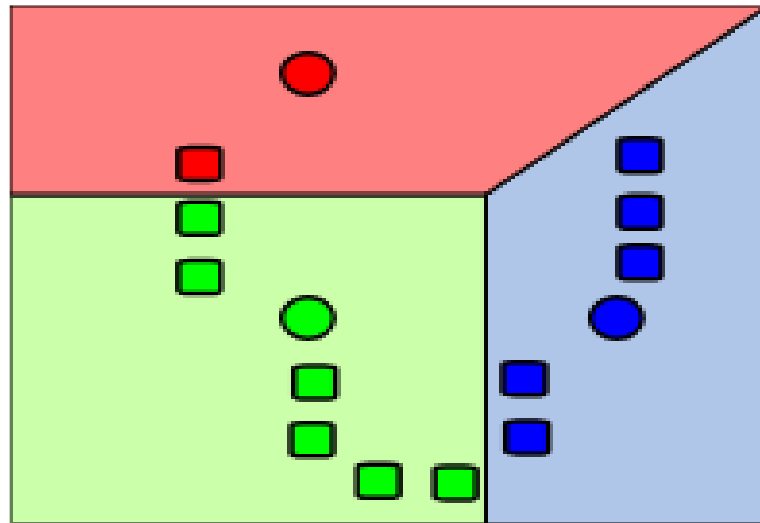


Рисунок 2.2 - Точки, віднесені до початкових центрів

На рисунку 2.3. представлено перерахунок метричності кожного об'єкту до відповідного центру кластерів та перевизначення його центрів.

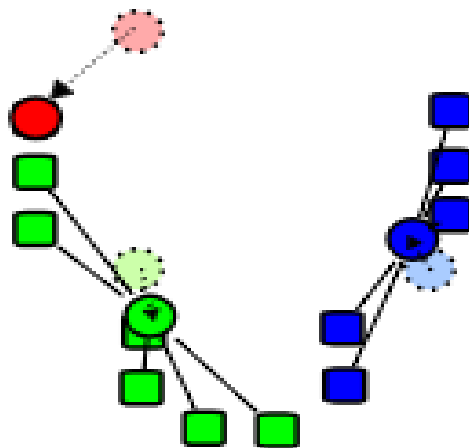


Рисунок 2.3 - Обчислення нових центрів кластерів

На рисунку 2.4 представлено останню ітерацію, щоподібна до попередньої та є обов'язковою умовою для зупинення ітераційного циклу.

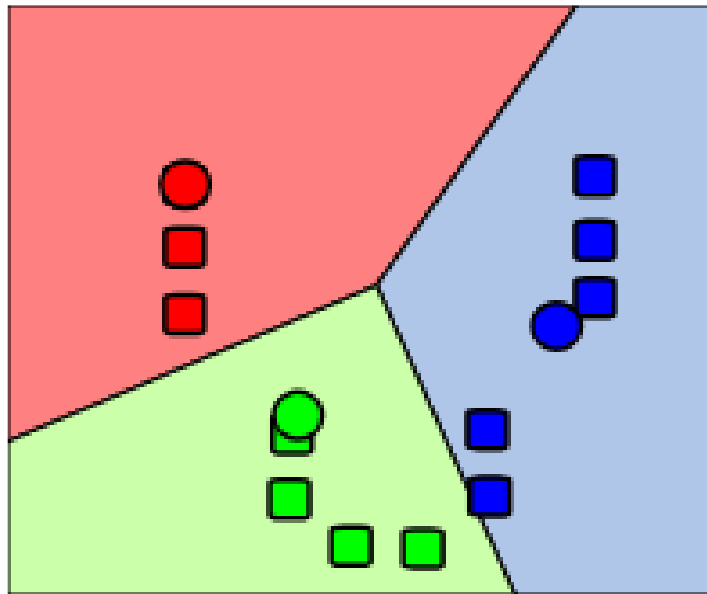


Рисунок 2.4 - Кінцевий результат кластерного алгоритму

Переважно показники, які представлено на рисунках 2.1-2.3 повторюються, доки не досягнуть результату, що зображений на рисунку 2.4.

Однак здебільшого в алгоритмах величезне число ітерацій відсутнє, тут існує залежність від вхідних даних.

Таким чином, процес розпізнання символів включає в себе порядок кластерного методу аналізування, який складається з таких взаємопов'язаних елементів[11, с. 215]:

1. початку;
2. зчитування вхідних даних;
3. створення матриці координат;
4. обчислення значень міри схожості;
5. відокремлення ліній на зображенні;
6. кінця.

Застосування кластерного аналізу означає такі дії:

1. можливість обрати об'єкт для кластеризації;
2. здатність аналізу критеріїв, завдяки яким аналізуватимуться об'єкти;
3. спроможність розрахувати значення відстані між об'єктами;

4. проектування подібних кластерів об'єктів;

5. результати аналізу.

Програми розпізнавання символів, що моделюється людиною, завжди мають спеціалізовану направленість на відміну від її власних фізичних можливостей. Відносно загальних підходів до створення будь-якої програми, то тепер у разі, якщо є деяка система об'єктів чи явищ, які потребують процесу розпізнання або класифікації, на базі узагальнення алгоритмів при моделюванні систем розпізнавання етапність вирішення конкретних задач виглядає наступним чином[37, с. 66]:

- відноснообраного принципу набір об'єктів чи явищ поділяється на класи, іншими словами, йому надаєтьсяябетка класів;
- проектуютьсяалгоритмвластивостей (словник);
- за словникомвластивостей характеризується кожен клас;
- моделюютьсямеханізми визначення властивостей;
- на механізмах обчислень впроваджується процес зіставлення апостеріорної та апріорної складових та визначаються результати загального процесу розпізнавання.

Проте хоча послідовність дій визначена, залишаються проблеми [38, с. 145]:

- поділ об'єктів на різноманітні класи;
- способи накопичення та обробки інформації;
- причини вибору ознаки;
- характеристика класів на мові ознак;
- вибір методів для порівняння апріорної та апостеріорної інформації;
- поява системи розпізнавання.

В процесі розпізнання символів груповий аналіз застосовується так, що під час розпізнання символу, його зображення переходить до піксельного, а це, в свою чергу, створює матрицю координат. Далі обирається міра відстані пікселів та формується певна лінія символу.

Порядок дій кластерного аналізу виглядає наступним чином:

1. зчитування і обробка вхідних даних;

Такими даними можуть бути зображення символів.

2. проектування матриці координат та виокремлення головних значень. На цьому етапі присвоюється кожному пікселю зображення номер кольорової гама, - потім він відшукується серед інших точок;

3. обчислення ступеню подібності. Виокремлення пікселю в групі з найближчою відстанню, це в свою чергу робить зображення ліній зображення;

4. кінцеве формування ліній символу та редакція для більшої якості, засобами видалення шумів;

5. фініш роботи алгоритму кластеризації і представлення даних для подальшого виконання програми.

Реалізувати подібний комплексний алгоритм на практиці досить важко, оскільки однією з причин є складність формалізації і обчислень ряду частинних критеріїв ефективності на різних етапах обчислень, на основі яких приймаються рішення про подальші дії з обробки і аналізу зображень.

У деяких випадках рішенням даної проблеми є формування моделей опису процесів і управління цими процесами на основі методів теорії інформації. До основних переваг даного підходу слід віднести [41, с. 719]:

- можливість оцінки стану процесу з безліччю результатів;
- універсальність опису різних процесів, у тому числі одно - та багато-етапних;
- можливість оцінки обсягів інформації, необхідних для реалізації процесу спостереження, незалежно від конкретних результатів процесу.

Виявлення об'єктів. Знаходження символу об'єкту в шумах та перешкодах зображення може визначатися у вигляді процесу перевірки певної умови порівняння кількох значень – результату трансформації досліджуваного зображення та певного порогового значення:

$$L [A (x, y)] \geq \Pi [A (x, y)],$$

де $L [.]$ – Оператор перетворення вихідного зображення,

$\Pi [.]$ – Оператор формування порогового значення.

Рішення про наявність об'єкта приймається у разі виконання умови. Якість виявлення характеризується ймовірністю виконання умови за наявності об'єкта в аналізованому зображенні.

Певний вид оператора $L [.]$, $\Pi [.]$ і якість розпізнавання символів залежать від наявності апріорних даних про очікувані об'єкту, шуму, перешкоди та видозмін. Підґрунтям для визначення оптимальних властивостей операторів є теорія рішень статистики.

Так, наприклад, під часозначення частин об'єктів схема зображення об'єкта, як правило, досліджується функціями плями розсіювань об'єктиву, яку вважають пріоритетною, при цьому завдання полягає в тому, щоб виявити класичну процедуру символу існуючого типу у площині адитивно- нормального шуму з нульовими середніми значеннями. За таких обставин операторами трансформації є кореляційні інтеграли, які обчислюються за допомогою заданих описів існуючих символів об'єктів та типових шумів. Визначення полягає в процесі порівняння зображень об'єктів з взірцями заданих класів. Результат про об'єкт визначається згідно найкращих співпадінь.

Метод прямого порівняння об'єктів з еталонними достатньо простий та здатний легко втілюватися. Проте, якщо наявні в реальних умовах дестабілізуючі фактори, надійність методу стає автоматично невисокою.

При більшому значенні порогів, можуть задовольнятися різні об'єкти та виникатимуть помилки, що зв'язані з невірним трактуванням визначення об'єкта, тобто під час помилок першого порядку. В процесі регулювання величин, видається можливим втілювати співвідношення між ймовірностями того, що помилки виникнуть як в першому, так і в другому порядку помилок згідно заданим критеріям оптимізації.

Кореляційний спосіб створений на обрахуванні взаємної кореляції об'єктівта еталонів. З множини k можливих варіантів обирається той об'єкт чи еталон, при якому досягається максимальна сума взаємних кореляцій:

$$K(k) = \sum_{x,y} S_k(x,y) \cdot F_k(x,y),$$

При визначенні об'єктів зручно мати справу з коефіцієнтами кореляцій, які в першому наближенні оцінюють ймовірності віднесення об'єктів до того чи іншого еталону:

$$R(k) = K(k) / K_{\max}(k),$$

де $K_{\max}(k)$ – значення автокореляції еталонів.

Кореляційний спосіб є надійнішим, проте потребує більшої кількості різноманітних обчислень. Однак при обробленні бінарних зображень це не є суттєвим, оскільки перемножування одnobітових цифр зводиться до одної логічної операції «І».

Охарактеризовані вище способипотребують однакового орієнтування зображень об'єктівта еталонів, з'єднання їх у площині координат і дотримання однакового масштабування.

Способи розпізнавання на базі системи властивостей також мають на увазі використання еталонів об'єкту, але в якості компонентів для порівняльного аналізузастосовуються властивості об'єктівта еталонів, що допомагає скороченню обсягу еталонних даних та часу на обробку інформаційних потоків.

Протеслід враховувати, що на практиці акцентування властивостей об'єктів відбувається з певною похибкою, а це означає, що потрібно виявляти та аналізувати параметри можливих розсіювань оцінки властивостей, які використовуються, для кожного із змодельованих об'єктів, іншими словами слід користуватися гістограмами розподілу значення властивостей.

При великій кількості можливих альтернатив об'єктів бажано в якості рекомендації використовувати ієрархічну структурованість.

При цьому на кожному рівні процесу розпізнавання об'єкту задіяна одна з певних властивостей об'єкта, наприклад, площа або периметри, радіус вписаного та описаного кола, момент інерційності, числа та знаходження кутів.

На нижніх ступенях задіяні властивості, які не вимагають великих обчислень витрат наприклад, площачи периметр об'єкта, а найінформативніші, такі як моменти інерції мають використовуватися на верхніх рівнях, де кількість варіантів є мінімальною [29, с. 1099].

Так, наприклад, під часозначення частин об'єктів схема зображення об'єкта, як правило, досліджується функціями плями розсіювань об'єктиву, яку вважають пріоритетною, при цьому завдання полягає в тому, щоб виявити класичну процедурисимволуіснуючого типуу площині адитивно- нормального шуму з нульовими середніми значеннями. За таких обставин операторамитрансформаціїє кореляційні інтеграли, які обчислюютьсяза допомогою заданих описівіснуючихсимволів об'єктів та типових шумів. Визначення полягає в процесі порівняння зображень об'єктів з взірцями заданих класів. Результат про об'єкт визначається згідно найкращих співпадінь.

Отже, завдання, які розв'язуються моделювання систем розпізнавання, є достатньо інваріантними по відношенню до предметної галузі, вони мають спільні задачі та ґрунтується на методі методологічного підходу. Водночас кожна програма розпізнавання є унікальною та потрібнатільки для одного певного виду об'єкта чи явища. У разі знаходження галузі застосування розпізнавання програма має проектуватися знову та враховувати нові специфічні властивості об'єкту або явища, що розуміють як сукупність вимірів певних ознак, так і в якості словника властивостей, абетки класів та алгоритмів прийняття рішень. Програма розпізнавання має моделюватися способом послідовних метричних наближень

внутрішніх компонентів на її математичних моделяху процесі додавання нових інформаційних потоків.

Висновки до другого розділу

Підсумовуючи другий розділ, можемо зробити такі висновки:

1. Розглянуто галузі застосування нейронних мереж для розв'язання практичних задач, а також деталізовано специфіку застосування нейронних мереж під час:

- Розпізнавання образів і класифікація.
- Кластеризація.
- Прогнозування і апроксимація.
- Рішення завдань при невідомих закономірностях.
- Прийняття рішень і керування.
- Стиск даних і асоціативна пам'ять.
- Оптимізація.

2. Проведено системний огляд програмних засобів, що використовують нейронні мережі.

3. Здійснено формалізацію кластерного аналізу та адаптацію методів сегментації до умов задачі.

3 ПРОЕКТ РІШЕННЯ З РЕАЛІЗАЦІЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ ЗАДАНИХ УМОВ

3.1 Методика синтезу штучних нейронних мереж для розпізнавання символів

В межах теорії штучного інтелекту розпізнавання символів включається в більш широку наукову дисципліну – теорію машинного навчання (machinelearning), метою якої є розробка методів побудови алгоритмів, що здатні навчатися.

На сьогоднішній день існують два підходи до машинного навчання: індуктивне і дедуктивне.

Індуктивне навчання, або навчання за прецедентами, засноване на виявленні загальних властивостей об'єктів на підставі неповної інформації, отриманої емпіричним шляхом.

Дедуктивне навчання передбачає формалізацію знань експертів у вигляді баз знань (експертних систем).

Кожний образ являє собою набір чисел, що описують його властивості і називаються ознаками (feature). Впорядкований набір ознак об'єкта називається вектором ознак (feature vector). Вектор ознак – це точка в просторі ознак (feature space). Класифікатор, або вирішальне правило (decision rule) – це функція, яка ставить у відповідність вектору ознак образу клас, до якого він належить.

Задачу розпізнавання символів можна розділити на ряд підзадач.

1. Генерування ознак (featuregeneration) – вимірювання або обчислення числових ознак, що характеризують об'єкт.

2. Вибір ознак (featureselection) – визначення найбільш інформативних ознак для класифікації (в цей набір можуть входити не лише первинні ознаки, але й функції від них).

3. Побудова класифікатора (classifier construction) – конструювання вирішального правила, на підставі якого здійснюється класифікація.

4. Оцінка якості класифікації (classifier estimation) – обчислення показників правильності класифікації (точність, чутливість, специфічність, помилки першого та другого роду).

Слід зазначити, що будь-яка нейрона мережа приймає на вході числові значення і видає на виході також числові значення. Передатна функція для кожного елемента мережі зазвичай вибирається в обмеженій області (наприклад, S - функції).

У цьому випадку вихідне значення завжди буде лежати в інтервалі (0,1), а область чутливості для входів ледь ширше інтервалу (-1,+1).

Дана функція є гладкою, а її похідна легко обчислюється – ця обставина важлива для роботи алгоритму навчання мережі (у цьому також криється причина того, що гранична функція для цієї мети практично не використовується).

При використанні нейронних мереж можуть виникати деякі проблеми, зокрема:

- дані мають нестандартний масштаб,
- дані є нечисловими.
- у даних є пропущене або недостовірне значення.

Числові дані масштабуються в прийнятий для мережі діапазон.

Зазвичай дані масштабуються по лінійній шкалі. У пакетах програмних нейромереж реалізовані алгоритми, що автоматично знаходять масштабуючі параметри для перетворення числових значень у потрібний діапазон.

Мережі, що навчені за допомогою алгоритмів глибокого навчання, не просто перевершили за точністю кращі альтернативні підходи, але й у ряді завдань проявили зачатки розуміння сенсу поданої інформації (наприклад, при розпізнаванні зображень, аналізі текстової інформації і так далі).

Найбільш успішні сучасні промислові методи комп'ютерного зору і розпізнавання мови побудовані на використанні глибоких мереж, а гіганти ІТ-

індустрії скуповують колективи дослідників, що займаються глибокими нейронними мережами.

Глибока нейронна мережа (DNN – DeepNeuralNetwork) - це штучна нейронна мережа з декількома прихованими шарами. Подібно до звичайних нейронних мереж, глибокі нейронні мережі можуть моделювати складні нелінійні відносини між елементами.

У процесі навчання глибокої нейронної мережі отримувана модель намагається представити об'єкт у вигляді комбінації простих примітивів. Додаткові шари дозволяють будувати абстракції все більш високих рівнів, що і дозволяє будувати моделі для розпізнавання складних об'єктів реального світу.

Як правило, глибинні мережі будуються як мережі прямого поширення. Однак останні дослідження показали, як можна застосувати техніку глибинного навчання для рекурентних нейронних мереж.

В якості головного компонента системи розпізнавання символів виступає видозмінений (багатошаровий) перцептрон (Рис. 3.1). Кількість нейронів перцептрона визначається кількістю вивчених ним образів (m), кількість синапсів (n) визначається розширенням рисунків, які розпізнаються.

Розглянемо алгоритм зворотного поширення для перцептрона. Символ подається рисунком, розміри якого $79,38 \times 52,92$ мм, тип - бітовий, розширення *.bmp.

Мережею рисунок розбивається на 600 частин (по горизонталі на 20, по вертикалі на 30). Отже, кількість синапсів для одного нейрона буде незмінним числом і визначатися як:

$$n = 20 * 30 = 600 .$$

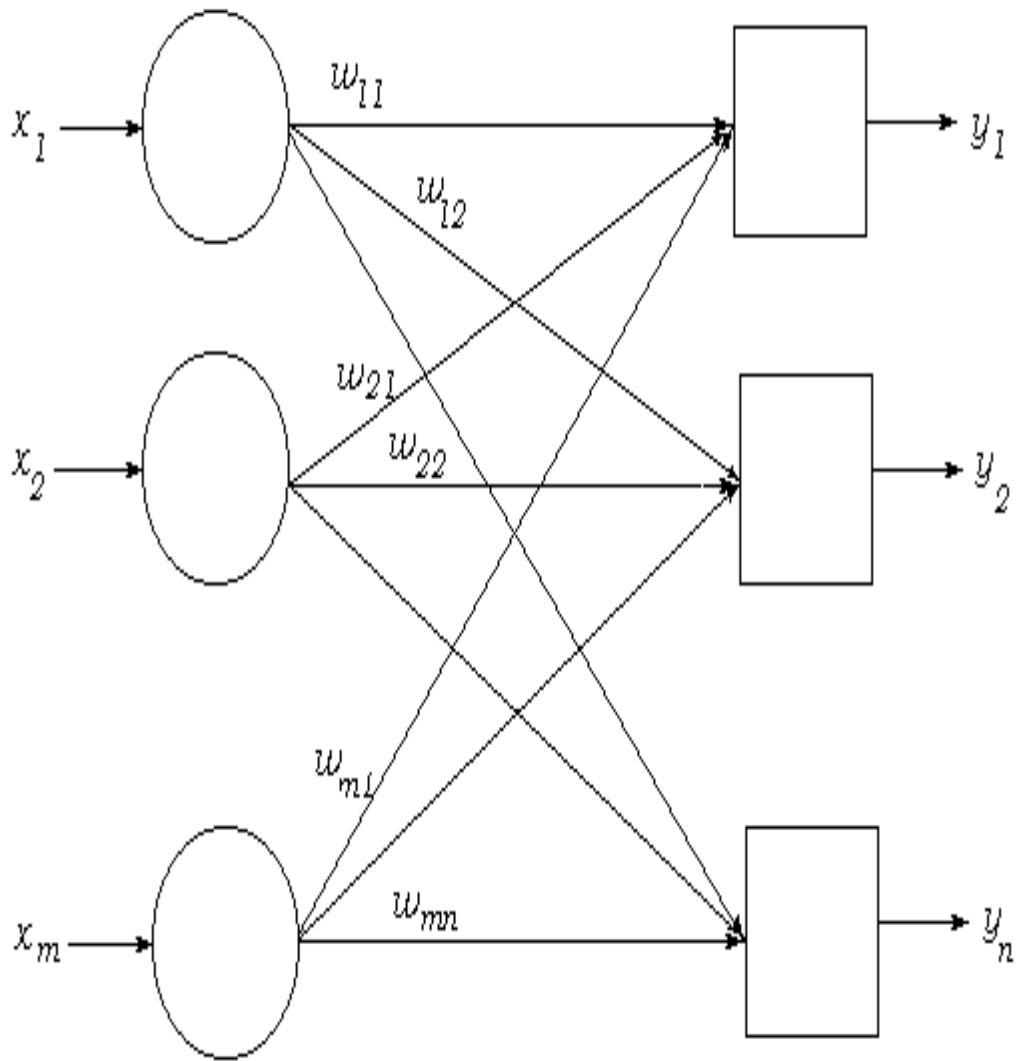


Рисунок 3.1 - Структура видозміненого перцептрона

Навчання перцептрона з урахуванням запропонованого підходу можна подати наступним чином.

1. Ініціалізуються всі ваги мережі в малі ненульові величини.
2. На вхід мережі подається вхідний навчальний вектор X вхідних символів і обчислюється сигнал NET вихідної функції від кожного нейрона, використовуючи формулу:

$$NET_j = \sum_i x_i w_{ij} ,$$

де x_i -й елемент вхідного вектора X ;

w_{ij} - вага входу i нейрона j ;

NET_j -рівень збудження післясинаптичного нейрона.

3. Обчислюється значення граничної функції активації для сигналу NET від кожного нейрона в такий спосіб:

$$OUT_j = \begin{cases} 1, & \text{якщо } NET_j \text{ більше ніж поріг } i_j ; \\ 0, & \text{у протилежному випадку,} \end{cases}$$

де i_j являє собою поріг, що відповідає нейрону j (у найпростішому випадку всі нейрони мають той самий поріг).

4. Обчислюється помилка для кожного нейрона:

$$error_j = target_j - OUT_j ,$$

де $error_j$ - помилка для j -го нейрона;

$target_j$ - необхідний вихід j -го нейрона;

OUT_j - отриманий вихід j -го нейрона.

5. Кожна вага модифікується в такий спосіб:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha \cdot error_j ,$$

де $w_{ij}(t+1)$ - вага входу i нейрона j в момент часу $(t+1)$;

$w_{ij}(t)$ - вага входу i нейрона j в момент часу t ;

α - нормуючий коефіцієнт навчання, початкове значення якого $0,1$, і який зменшується в процесі навчання.

Повторюються кроки з другого до п'ятого, доки помилка не стане досить малою.

Робота інтелектуального модуля з розпізнавання символів полягає в підрахуванні виходів перцептрона й, якщо два або більше нейронів на виході отримали максимальне значення, то переглядається який нейрон має більшу суму на вихідного суматора, і вносяться зміни до результату.

Активаційна функція нейронів має вигляд лінійного порога :

$$Y = S \text{ при } S < T ;$$

$$Y = S \text{ при } S \geq T ,$$

де Y - значення активаційної функції;

S - аргумент активаційної функції;

T - величина порога.

Подальше навчання перцептрона здійснюється під час роботи інтелектуального модуля штучних нейронних мереж із розпізнавання символів. Ваги синапсів корегуються відповідно Дельта-правилу:

$$\delta = (D - Y) ,$$

де δ - рівна різниці між необхідним або цільовим виходом D_i реальним виходом Y .

Дельта-правило модифікує ваги відповідно до необхідного й дійсного значень виходу кожної полярності як для безперервних, так і для бінарних входів і виходів.

Алгоритм навчання зберігається якщо δ збільшується на величину кожного входу x_i , і цей добуток додається до відповідної ваги.

З метою організації управління середньою величиною зміни ваг уведемо коефіцієнт «швидкості навчання» η , що збільшується на δx_i при навчанні:

$$D = \eta \delta x_i ,$$

$$i w(n + 1) = w(n) + D ,$$

де D_i - корекція, пов'язана з i -м входом x_i ;

η - коефіцієнт швидкості навчання;

$w_i(n + 1)$ - i -е значення ваги після корекції;

$w_i(n)$ - i -е значення ваги до корекції.

Дельта-правило модифікує ваги відповідно до i -го необхідного дійсного значення виходу.

Таким чином, використання зворотного поширення для перцептрона, дозволяє досягти ефективного розпізнавання навіть зашумлених символів.

Розпізнавання символів повинно проводитись за таким алгоритмом:

- 1) завантаження параметрів по замовчуванню під час запуску програми;
- 2) створення еталонних зображень (створивши еталони, користувач може перенавчити мережу так, щоб створений ним образ символу був еталоном);
- 3) завантаження чи створення зображення символу для розпізнавання;
- 4) розпізнавання символу (зашумленість якого може бути до 40 %) системою;
- 5) виведення результату розпізнавання.

Зазначений алгоритм може бути реалізований у середовищі Borland Delphi 7, на мові Object Pascal, яка дозволяє використовувати бібліотеки, що реалізують нейронні мережі.

Схему алгоритму функціонування модуля розпізнавання символів з урахуванням обраної мови програмування наведено на рисунку 3.2.

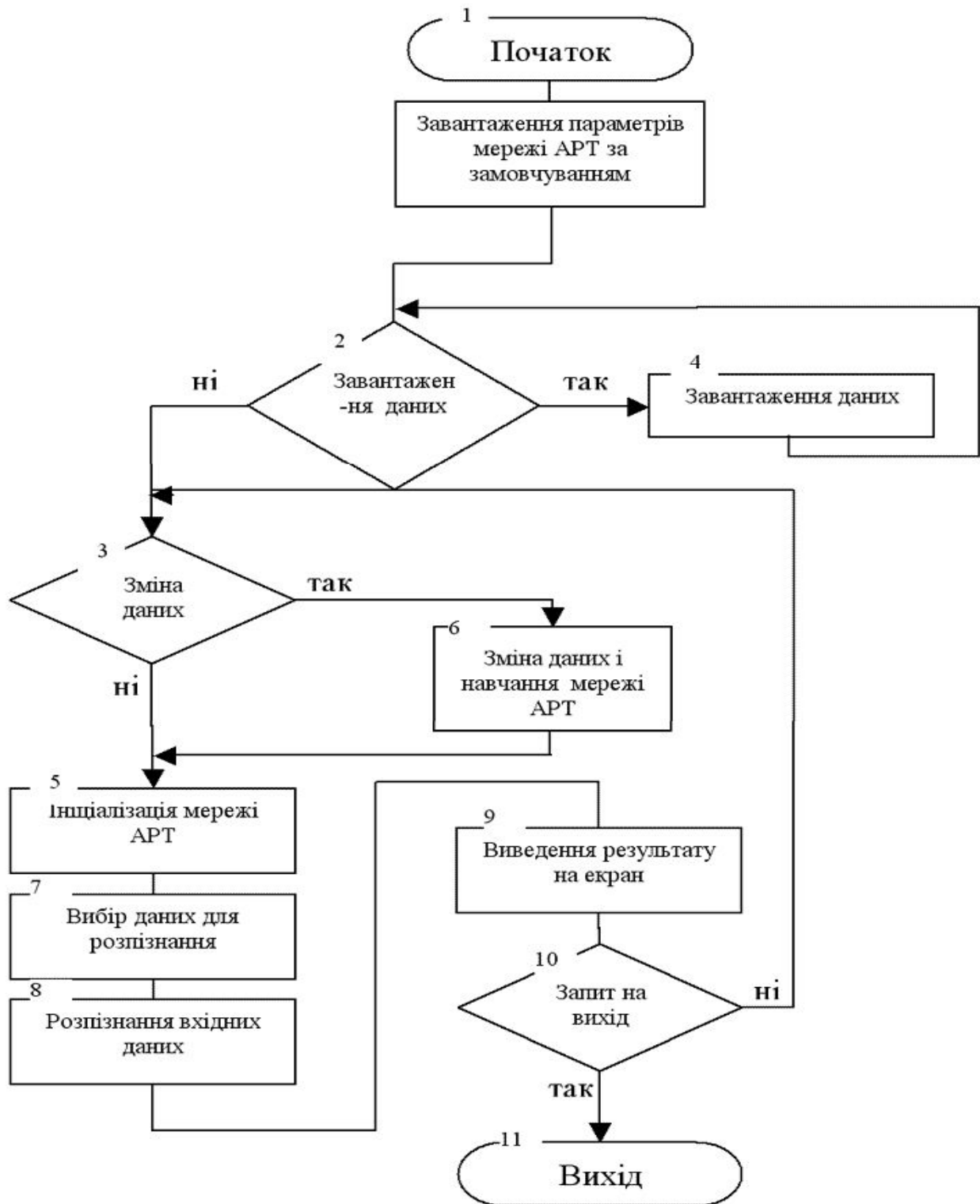


Рисунок 3.2 - Схема алгоритму НМ для розпізнавання символів

Результат реалізації приведенного алгоритму в середовищі програмування

BorlandDelphi 7 надав можливість реалізувати інтелектуальний модуль ШНМ із розпізнавання символів.

Результати роботи інтелектуального модуля можна відобразити на рисунку, на якому зображені еталони образів і зашумлені копії, при яких програмний комплекс може правильно визначити образ (рисунок 3.3). Зображення поділяються за відсотком зашумленості.

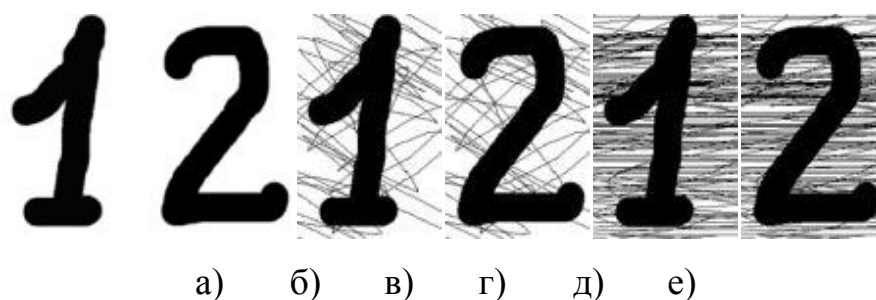


Рисунок 3. 3 - Символи для роботи інтелектуального модуля штучних нейронних мережіз розпізнавання символів: а), б) - еталонні символи; в), г) - символи із зашумленістю до 30 %; д), е) - символи із зашумленістю до 40 %

Таким чином, у результаті реалізації інтелектуального модуля, що функціонує у відповідності із запропонованим алгоритмом, розпізнаються символи, які належать до будь-якого алфавіту, цифри й знаки в залежності від об'ємів пам'яті інтелектуальної системи, що містить відповідні еталони із зашумленістю до 40 %.

3.2 Комбінування двох мереж з однаковою структурою

Архітектури сучасних нейронних мереж найчастіше поділяють на три категорії:

- мережі з повним набором міжнейронних зв'язків;
- мережі з фіксованим індексом оточення;
- мережі з пошаровою структурою.

У ШНМ із повним набором міжнейронних зв'язків забезпечується можливість взаємодії кожного нейрона мережі з будь-яким іншим.

Структури з повним з'єднанням є узагальненими структурами, оскільки всі інші довільні об'єднання нейронів можуть розглядатися як підмножини даної структури. Тому ШНМ із повним з'єднанням є універсальним середовищем для реалізації мережних алгоритмів. Широке використання таких структур обмежується недоліком, який полягає в значному зростанні кількості міжнейронних зв'язків при збільшенні кількості нейронів.

У випадку, коли необхідно використовувати структури з великою кількістю нейронів, застосовують кліткові структури з фіксованим індексом оточення.

Ще однією важливою характеристикою нейронних мереж із постійним індексом оточення є модульність.

Потужність структури у цьому випадку може нарощуватись простим додаванням елементів без зміни ідеології алгоритму, що на ній працює.

Найбільш поширеними є структуровані за шарами нейронні мережі, які, в залежності від свого функціонального призначення, можуть містити однотипні або різнотипні нейрони.

Виходячи з шарової структури ШНМ, характер міжнейронних зв'язків має свої міжшарові та внутрішньошарові особливості.

Нейронні структури з повним з'єднанням можуть бути як одношаровими, так і багатошаровими. У одношаровій структурі з повним з'єднанням всі вхідні сигнали можуть поступати на всі нейрони.

Класичною структурою даного типу є перцептрон Розенблатта. Багатошарові мережі з повним з'єднанням забезпечують можливість передачі інформації з кожного нейрона попереднього шару на будь-який нейрон наступного. Найчастіше це - багатошарові перцептрони.

У випадку прямого міжшарового поширення інформація передається однонаправлено у напрямку зростання номера шару. Пряме поширення в межах

одного шару використовують увипадку, коли група нейронів даного шару з'єднана з попереднім шаром опосередковано через виділені нейрони.

Двонаправлене поширення допускає також зворотну передачу, що дозволяє створювати алгоритми, за якими враховувався взаємний міжнейронний обмін (Рис. 3.4).

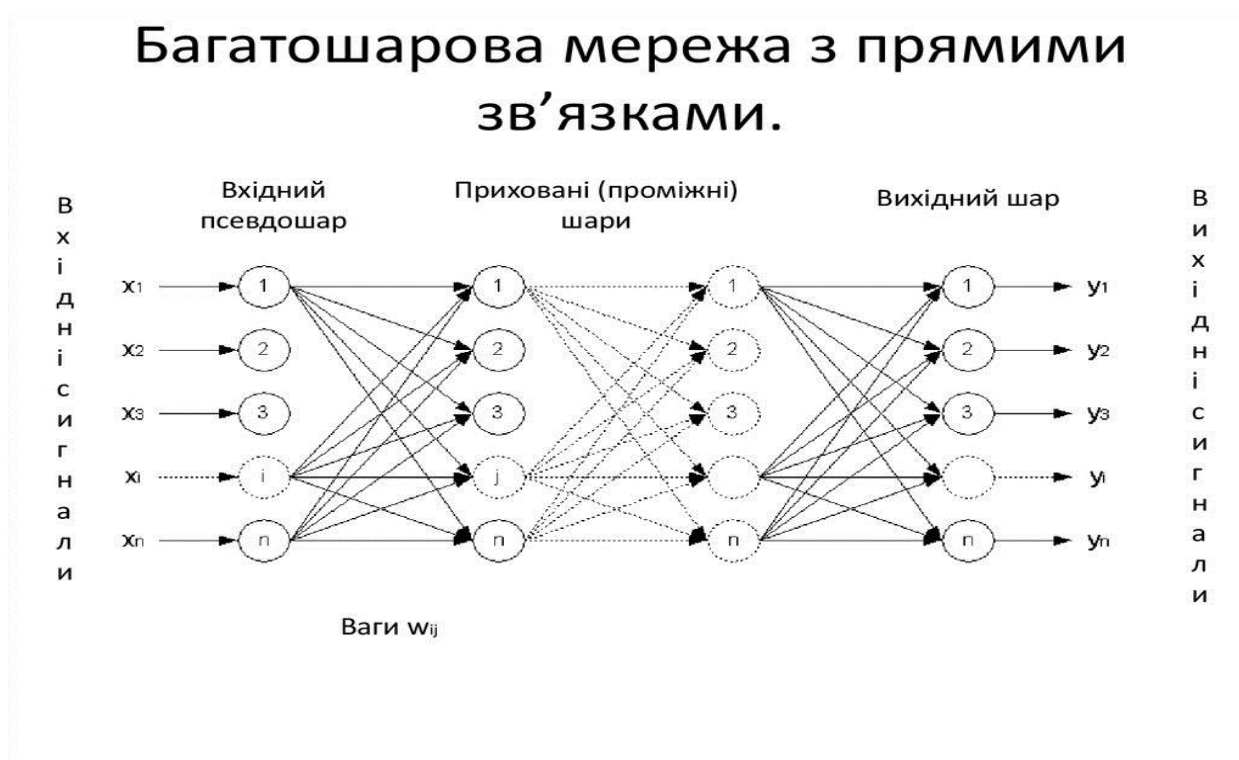


Рисунок 3.4 – Приклад багатошарової мережі з прямими зв'язками

Таким чином, однонаправленість зв'язків призводить до побудови виключно ієрархічних структур, у яких обробка інформації розподіляється по рівнях. За кожний рівень ієрархічної обробки інформації відповідає свій шар нейронів. Вихідна інформація більш високого рівня попереднього шару є вхідною для нейронів наступного шару, який забезпечує глибший рівень обробки.

З іншого боку, 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.

У процесі навчання глибокої нейронної мережі отримувана модель намагається представити об'єкт у вигляді комбінації простих примітивів (наприклад, у задачі розпізнавання осіб такими примітивами можуть бути частини обличчя: ніс, очі, рот і так далі). Додаткові шари дозволяють будувати абстракції все більш високих рівнів, що і дозволяє будувати моделі для розпізнавання складних об'єктів реального світу.

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.

Мережі прямого поширення:

1. Перцептрони,
2. Мережа Back Propagation,
3. Мережа зустрічного поширення,
4. Карта Кохонена.

Зокрема, у навчанні системи розпізнавання піксельних образів для десяти цифр, які представлені двадцятьма прикладами кожної цифри, всі приклади цифри «сім» не доцільно представляти послідовно.

Краще надати мережі спочатку один тип представлення всіх цифр, потім другий тип і так далі.

Головною компонентою для успішної роботи мережі є представлення і кодування вхідних і вихідних даних. Штучні мережі працюють лише з числовими вхідними даними, отже, необроблені дані, що надходять із зовнішнього середовища повинні перетворюватись.

Важливою є нормалізація даних, тобто приведення всіх значень даних до єдиного діапазону.

Нормалізація виконується шляхом ділення кожної компоненти вхідного вектора на довжину вектора, що перетворює вхідний вектор в одиничний.

Попередня обробка зовнішніх даних, отриманих за допомогою сенсорів, у машинний формат є спільною і легко доступною для стандартних ПК.

На відміну від задачі визначення структури, яка є дискретною оптимізаційною (комбінаторною), пошук оптимальних параметрів здійснюється в безперервному просторі за допомогою класичних методів оптимізації.

Для навчання мереж прямого розповсюдження з учителем застосовуються, як правило, алгоритми, що оптимізують деяку цільову функцію. Однак традиційно до уваги береться лише одна ціль як вартісна функція або кілька цілей об'єднуються в одну скалярну функцію виду,

$$\Omega\lambda + F(e) = f$$

де F – функція помилки;

Ω – міра складності моделі, наприклад, кількість вільних параметрів моделі;

$\lambda > 0$ – деякий вільно обраний параметр.

Одержуваний при цьому алгоритм навчання здатний оптимізувати дві цілі, хоча функція цілі є скалярною. Однак використанню скаляризованих цільових функцій для багатоцільової оптимізації (БО) притаманні два основних недоліки:

- по-перше, нетривіальною є задача визначення оптимального параметра λ ;
- по-друге, при цьому може бути отримано тільки одне рішення, яке в ряді випадків може бути неефективним.

Більш потужним порівняно з навчанням на основі скалярної вартісної функції є багатоцільове навчання на основі підходу Парето, коли мінімізується векторна цільова функція, що забезпечує отримання певної кількості Парето-оптимальних рішень. Так, скаляризована двоцільова проблема навчання може

бути сформульована як багатокритеріальна оптимізація на основі Парето таким чином:

$$f\{\min 1, f\}2; (2) () 1 Fe=f; (3) f\Omega=2$$

Найбільш часто в якості f_1 вибирається квадратичний функціонал, а в якості f_2 , що служить для оцінювання складності нейромережевої моделі, сума квадратів вагових параметрів: $=\sum=\Omega M$, або сума їх абсолютних значень $=\sum=\Omega M_i w_i 1$, відомі як регуляризатори Гауса і Лапласа відповідно.

Тут $w_i 1, \dots,$

M – вагові-параметри нейромережевої моделі;

M – загальна кількість нейронів мережі (для радіально-базисної мережі – РБМ) або загальна кількість зв'язків (для багатошарового персептрону – БП).

У зв'язку з тим, що нейромережева модель, з одного боку, повинна бути досить простою і зручною для використання її в прикладних задачах, а з іншого – найбільш повно відображати властивості досліджуваного об'єкта, її якість визначається деяким набором критеріїв, тобто задача побудови нейромоделі є багатокритеріальною.

Дослідження в галузі можливостей комбінації більшої кількості глибоких мереж та розробка самих класифікаторів є дуже важливими, оскільки класи в просторі ознак можуть мати нелінійне поділ.

Нейронні мережі типу персептрона здатні розділяти тільки лінійно розділені класи. А нейронні мережі типу радіально-базисних мереж, здатні розділяти тільки класи зі сферичним розподілом. Часто ці два типи нейронних мереж комбінують. Дуже важливо правильно вибрати простір ознак, так як якщо буде зайва інформація про сигнал, вона ускладнить процедуру розпізнавання з-за того, що класи будуть складно віддільні: вони можуть бути як нелінійно віддільні, так і невіддільні взагалі.

Якщо даних буде недостатньо, то буде проблематично розпізнати об'єкт, з-за того що одного набору даних буде відповідати кілька сигналів. Для розпізнавання, в якості ознак можна ввести статистичні параметри сигналів.

Отже, логічним видається ввести наступні критерії, за якими визначати кількість ознак:

- 1) Ознаки повинні відрізнятися для об'єктів, що належать різним класам.
- 2) Вони повинні збігатися за своїм значенням для об'єктів одного класу.

Повний збіг статистичних параметрів сигналів можливий за умови, якщо сигнал ергодичний і час спостереження прагне до нескінченності.

Неергодичний сигнал або сигнал, в якому спостерігається кінцевий час, утворює деяку область в багатовимірному просторі ознак. Сигнали відносяться до різних класів та утворюють свої області.

Завдання класифікатора полягає в тому, щоб відокремити одну область від іншої. Завдання первинної обробки зводиться до опису сигналу кінцевим вектором, таким чином, щоб сигнали відносилися до одного класу та перебували близько один від одного в багатовимірному евклідовому просторі, а сигнали, що відносяться до різних класів - далеко, іншими словами щоб виконувалася гіпотеза компактності.

Слід зазначити, що сигнали майже завжди записані на тлі шуму, і методи розпізнавання на основі Фур'є-перетворень або розпізнавання по відлікам у часовій області передбачають в якості первинної обробки фільтрацію сигналу.

Метод розпізнавання на основі статистик не потребує фільтрації, якщо перешкода є ергодичним випадковим процесом, оскільки завжди можна «відняти» з моделі сигналу з завадою модель перешкоди.

Таким методом можна розпізнавати біологічні сигнали. Оскільки біологічні сигнали, є дуже специфічними, для їх розпізнавання потрібно враховувати механіку процесів, що породжують ці сигнали.

Скажімо, сигнал ЕКГ сигнал ЕЕГ, досліджуються різними методами. На практиці виявилось, що для вирішення більшості задач розпізнавання досить використовувати всього 4 параметра для розпізнавання, такі як:

- математичні сподівання,
- СКВ,
- ексцес
- асиметрія.

Для прикладу розглянемо розпізнавання сигналу при наступних варіантах:

- 1) немає сигналу (тільки перешкода),
- 2) синусоїда + перешкода,
- 3) прямокутний + перешкода
- 4) радіоімпульс з прямокутною обвідною + перешкода.

Відношення сигнал/шум ($q = \frac{A}{\sigma}$) в експерименті дорівнює 0.2. Амплітуду сигналу візьмемо рівною 1 Вольт(щоб не нормувати), і перешкода у вигляді білого шуму з нормальним розподілом. Для розпізнавання будемо використовувати двошаровий перцептрон, з 4 входами, 4 виходами і 9 нейронами на прихованому шарі(теорема Колмогорова достатня умова).

Первинна обробка.

Складемо вектор у просторі ознак, для початку визначимо модель шуму:

$$\begin{aligned} m &= 0; \\ \sigma &= 5; \\ \gamma_2 &= 0; \\ \gamma_1 &= 0; \end{aligned}$$

Тепер складемо вектор для сигналу з перешкодою:

$$\begin{aligned}
 m &= \int_{-\infty}^{\infty} x * f(x) dx; \\
 \sigma &= \sqrt{D}; \\
 D &= \int_{-\infty}^{\infty} (x - m)^2 * f(x) dx; \\
 \gamma_2 &= \frac{\mu_4}{\sigma^4} - 3; \\
 \gamma_1 &= \frac{\mu_3}{\sigma^3}; \\
 \mu_n &= \sum_{i=0}^n [(-1)^i * C_n^i u_{n-i} * u_1^i]; \\
 u_k &= \int_{-\infty}^{\infty} x^k * f(x) dx;
 \end{aligned}$$

де m - математичне очікування,

D – дисперсія,

σ — СКО,

γ_2 — ексцес,

γ_1 - асиметрія.

Але це було у випадку з неперервними сигналами, і коли відомо аналітичний вираз щільності ймовірності $f(x)$.

З дискретним сигналом інтеграл замінюється на суму.

І ми вже говоримо не про статистичних параметрах, а про їх оцінках.

А там набирають чинності похибки визначення оцінок.

$$\begin{aligned}
 m^* &= \frac{1}{N} \sum_{i=0}^n x_i; \\
 \sigma &= \sqrt{D}; \\
 D^* &= \frac{1}{N-1} \sum_{i=0}^n (x_i - m)^2; \\
 \gamma_2 &= \frac{\mu_4}{\sigma^4} - 3; \\
 \gamma_1 &= \frac{\mu_3}{\sigma^3}; \\
 \mu_n &= \sum_{i=0}^n [(-1)^i * C_n^i u_{n-i} * u_1^i]; \\
 u_k &= \frac{1}{N} \sum_{i=0}^n x_i^k;
 \end{aligned}$$

де N- кількість відліків.

Після цього можемо сформувати вхідний вектор для нейронної мережі.

$$\begin{aligned}
 X &= [m_x, \sigma_x, \gamma_{2x}, \gamma_{1x}]; \\
 m_x &= m - m_\xi; \\
 \sigma_x &= \sqrt{D - D_\xi}; \\
 \gamma_{2x} &= \gamma_2; \\
 \gamma_{1x} &= \gamma_1;
 \end{aligned}$$

Змінні з індексом «x»- компоненти вхідного вектора, без індексу - характеристики сигналу з шумом, а з індексом "ξ" модель шуму.

Для нормування СКО ділиться ще.

Практична реалізація.

Нижче представимо вигляд інтерфейсу тестової програми.

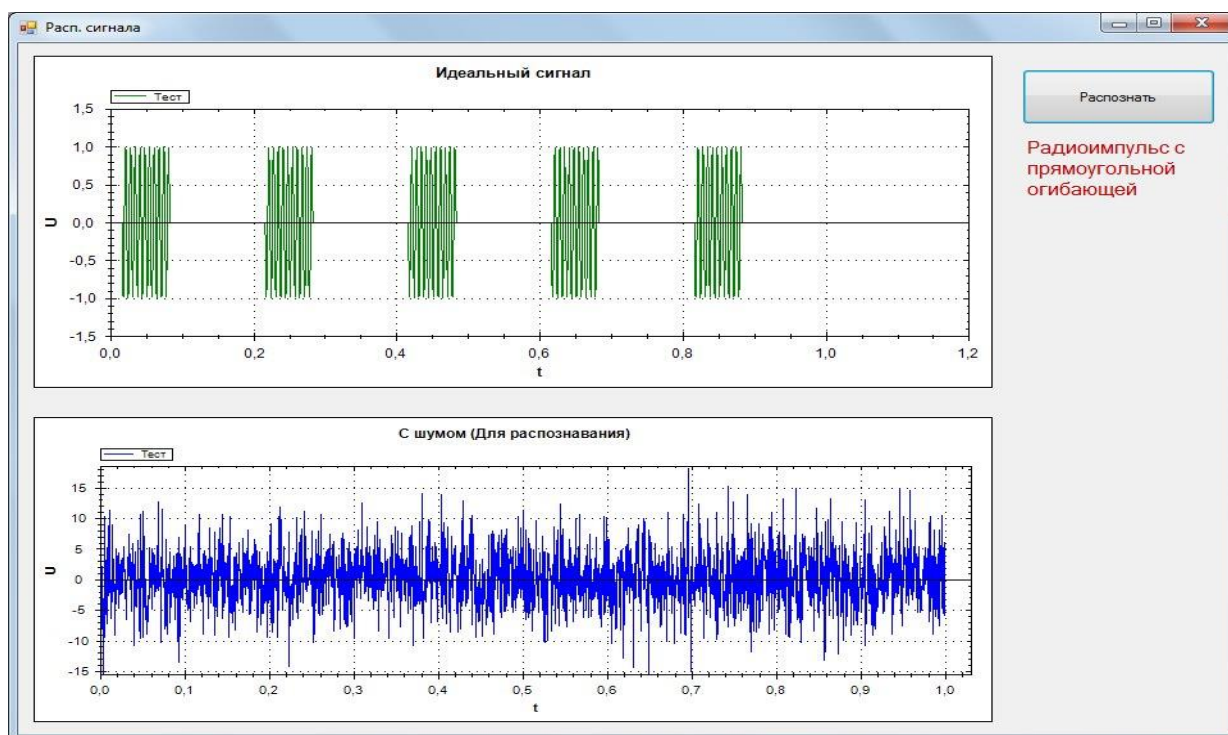


Рисунок 3.5 – Інтерфейс тестової програми розпізнавання

В даному експерименті ймовірність вірного розпізнавання склала 94.6%. Нейронна мережа навчалася методом градієнтного спуску на вибірці 50 реалізацій на сигнал, в сигналі 2001 відлік.

Ключовим моментом в розумінні ЗНМ є поняття так званих спільних ваг, тобто частина нейронів деякого шару нейронної мережі може використовувати одні і ті самі вагові коефіцієнти. Нейрони, що використовують одні й ті самі ваги, об'єднуються в карти ознак, а кожен нейрон карти ознак пов'язаний з частиною нейронів попереднього шару.

При обчисленні мережі виходить, що кожен нейрон виконує згортку деякої області попереднього шару (яка визначається множиною нейронів, пов'язаних з даними нейроном). Шари нейронної мережі, побудовані описаним чином, називаються згортковими шарами.

Крім згорткових шарів, в згортковій нейронній мережі можуть бути шари агрегації (субдискретизації), що виконують функції зменшення розмірності карти ознак, і повнозв'язані шари (класифікатор, який знаходиться на виході мережі).

Згорткові шари та шари агрегації можуть чергуватися, найчастіше шари агрегації розміщують за шарами згортки.

Одна з головних проблем використання глибоких нейронних мереж, чи то багат шарових звичайних, чи то згорткових – це перенавчання (англ. *overfitting*). Перенавчання проявляється у тому, що створена модель починає пояснювати тільки приклади з навчальної вибірки, адаптуючись до навчальних прикладів, замість того щоб вчитися класифікувати приклади, які не брали участі в навчанні (втрачаючи здатність до узагальнення). За останні роки було запропоновано безліч рішень проблеми перенавчання, але одне з них перевершило інші, завдяки своїй простоті і досягненням у практичних результатах.

Модель згорткової мережі складається з трьох типів шарів: згорткові (*convolutional*) шари, субдискретизуючі (*subsampling*, підвибірка) верстви і прошарки «звичайної» нейронної мережі - персептрона.

Архітектура згорткових нейронних мереж реалізує три ідеї, які забезпечують інваріантність мережі до невеликих зрушень, змін масштабу і спотворень: кожен нейрон отримує вхідний сигнал від

- локального рецептивного поля (*local receptive fields*) у попередньому шарі, що забезпечує локальну двовимірну зв'язність нейронів; кожен прихований шар мережі складається з безлічі,

- карт ознак, на яких всі нейрони мають загальні ваги (*shared weights*), що забезпечує інваріантність до зміщення і скорочення загальної кількості вагових коефіцієнтів мережі; за кожним шаром згортки слідує обчислювальний шар, який здійснює локальне усереднення та підвибірку, що забезпечує зменшення розширення для карт ознак.

Загальний вигляд згорткової нейронної мережі представлено нижче (Рис.3.6).

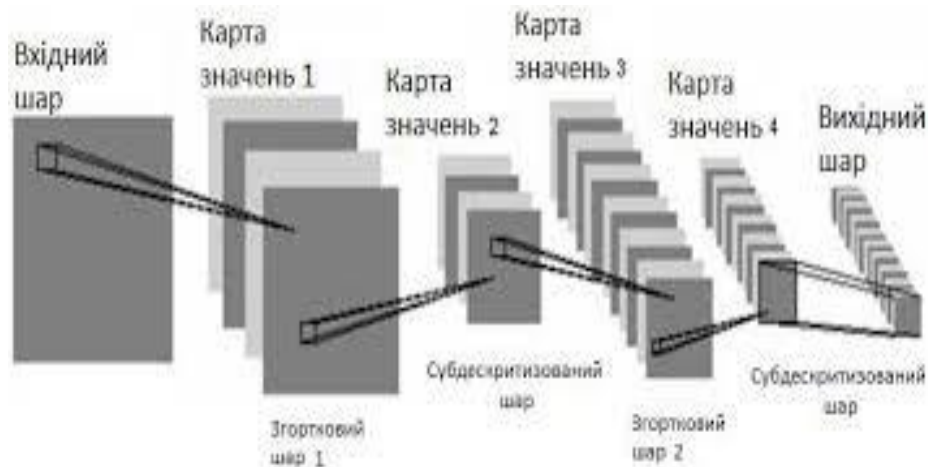


Рисунок 3.6 – Загальний вигляд ЗНМ

Робота згорткової нейронної мережі забезпечується двома основними елементами:

1. Фільтри (filters) (визначники ознак).
2. Карти ознак (featuremaps).

Фільтр – це невелика матриця, що представляє ознаку, яку необхідно знайти на вихідному зображенні. За допомогою верхнього фільтра визначаються частини вихідного зображення з вертикальними лініями, нижній фільтр служить для визначення частин зображення з горизонтальними лініями.

Безпосередньо процес визначення заснований на операції згортки фільтром оригінального зображення. Результати згортки, які визначають місце розташування ознак вихідного зображення, називаються картами ознак.

Мета процесу згортки – зменшити розмірність карти ознак до такої міри, щоб з повним набором ознак могла працювати мережа прямого поширення (в більшості випадків багат шаровий персептрон).

Згортковий шар реалізує ідею локальних рецептивних полів, тобто кожен вихідний нейрон з'єднаний тільки з певною (невеликою) областю вхідної матриці і таким чином моделює деякі особливості людського зору.

Недоліками згорткових нейронних мереж (ЗНМ) є:

висока складність архітектури;

- ✓ повнозв'язаність;
- ✓ фіксована площа вікна шару згортки.

З метою підвищення ефективності роботи ЗНМ необхідно знайти оптимальні значення наступних параметрів:

- ✓ кількість карт ознак;
- ✓ щільність зв'язків між картами ознак;
- ✓ розмір вікна;
- ✓ площа перекриття;
- ✓ початкова ініціалізація ваг.

Основна ідея ЗНМ полягає в тому, що обробка ділянки зображення дуже часто має відбуватися незалежно від конкретного розташування цієї ділянки. Це локальна задача, яку можна вирішити локальними засобами.

Звичайно, взаємне розташування об'єктів грає важливу роль, але спочатку їх потрібно в будь-якому випадку розпізнати, і це розпознавання - локально і незалежно від конкретного положення ділянки з об'єктом всередині великої картинки.

«Згортка» - це всього-на-всього лінійне перетворення вхідних даних особливого виду.

Якщо x_l - карта ознак в шарі під номером l , то результат двовимірної згортки з ядром розміру $2d+1$ і матрицею ваг W розміру $(2d+1) \times (2d+1)$ на наступному шарі буде таким:

$$y_{i,j}^l = \sum_{-d \leq a,b} w_{a,b} x_{i+a,j+b}^{l-1}$$

де $y_{i,j}^l$ - результат згортки на рівні l , $x_{i,j}^{l-1}$ - її вхід, тобто вихід попереднього шару.

Інакше кажучи, щоб отримати компоненту (i, j) наступного рівня, ми застосуємо лінійне перетворення до квадратному вікна попереднього рівня,

тобто скалярно множимо пікселі з вікна на вектор згортки. Це перетворення має ті властивості, про які ми говорили вище:

1. згортка зберігає структуру входу (порядок в одновимірному випадку, взаємне розташування пікселів в двовимірному), так як застосовується до кожної ділянки вхідних даних окремо;

2. операція згортки має властивість розрідженості, так як значення кожного нейрона чергового шару залежить тільки від невеликої частки вхідних нейронів (а, наприклад, в повно-нейронній мережі кожен нейрон залежав би від усіх нейронів попереднього шару);

3. згортка багаторазово перевикористовує одні і ті ж ваги, так як вони повторно застосовуються до різних ділянок входу. Головною особливістю згорткових нейромереж є те, що вони не потребують попередньої обробки даних (особливо на фоні більш класичних методів машинного навчання).

Це означає, що мережа самостійно навчається фільтрам, які в більш традиційних, класичних алгоритмах комп'юторного зору розроблялися за допомоги експертів. Такі мережі незалежні від апріорних знань про предметну область, що є важливим плюсом.

ЗНМ має шари входу та виходу та приховані шари між ними, чим нагадує перцептрон. Різниця полягає в тому, які саме шари використовуються в нейронній мережі. Приховані шари ЗНМ зазвичай складаються зі згорткових шарів (convolutional layers), агрегувальних шарів (pooling layers), повноз'єднаних шарів (dense layers, чи full connection layers) та шарів нормалізації.

Згорткові шари за допомогою згортки імітують реакцію біологічних нейронів на візуальний сигнал. Кожен згортковий нейрон може обробляти інформацію лише в межах свого рецептивного поля. Повноз'єднані нейронні мережі прямого поширення можуть бути використаними як для вивчення ознак, так і для розв'язання задачі класифікації, проте використання цієї архітектури до зображень є непрактичним.

При використанні багат шарового перцептронного нейронного мережі необхідно значно більше число нейронів через дуже великі розміри входу, пов'язані з зображеннями, де кожен піксель є відповідною змінною.

Операція згортки дає змогу розв'язати цю задачу, оскільки вона зменшує кількість вільних параметрів, дозволяючи мережі бути глибшою за меншої кількості параметрів. Це розв'язує проблему зникання або вибуху градієнтів у тренуванні традиційних багат шарових нейронних мереж з багатьма шарами за допомогою зворотнього поширення.

Розглянемо предметніше архітектуру ЗНМ:

ЗНМ складаються з сукупності різних шарів, які трансформують вміст входу до виходу за допомогою певної функції, що є диференційованою. Найчастіше використовуються кілька видів шарів. Найчастіше це одні з наступних шарів:

Згортковий шар: Згортковий шар (англ. convolutional layer) є найбільш важливим (і основним) шаром згорткової нейромережі. Його параметрами є сукупність фільтрів, кожен з яких має рецептивне поле невеликого розміру, яке простирається по всьому обсягу вхідного вмісту.

Під час етапу прямого проходу кожен фільтр обчислює результат скалярного добутку входу та значень фільтру і формує n -вимірну (де n - розмірність даних, з якими працює нейромережа, наприклад для зображень це буде 2) матрицю активації фільтру.

Внаслідок цього мережа вивчає, які фільтри реагують на окремо взяті ознаки, що розташовані певним чином на вхідних даних. Сукупність усіх матриць активації для всіх фільтрів упродовж всього обсягу даних і є результатом виходу згорткового шару.

Також є справедливою така точка зору, коли кожен блок в цій сукупності може розглядатись як результат роботи нейрону, який має невелике рецептивне поле на вході та однакові параметри з іншими нейронами тої ж карти активації.

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.

ЗНМ використовують просторово локальну кореляцію шляхом забезпечення схеми локальної з'єднаності між нейронами сусідніх шарів: кожен нейрон з'єднано лише з невеликою областю вхідної ємності. Обшир цієї з'єднаності є гіперпараметром, що називається рецептивним полем нейрона. 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.

3.3 Рекомендації з використання нейронних мереж для розпізнавання символів

В даний час існує багато систем, які показують можливості ШНМ: мережі можуть подавати текст фонетично, розпізнавати рукописи та зображення. Більше число потужних мереж, які спроможні розпізнавати символи і звуки, застосовують як базу принцип зворотнього розповсюдження. Це є систематичним способом для навчання багат шарових НМ, в якому існують свої обмеження.

Нейронні мережі, які працюють за вищевизначеною технологією прийшли на зміну програмам, які склались лише з одного шару ШНМ і застосовувались для широкого класу задач, у тому числі для штучного зору.

Проте такі програми недосконалі. По-перше, не існує гарантії, що НМ навчиться за визначений проміжок часу. Кожен із спроектованих алгоритмів навчання НМ має свої особливості, проте спільним недоліком є обмеження можливостей навчатися та запам'ятовувати.

При навчанні ШНМ кожна вхідна чи вихідна множина сигналів класифікується в якості вектору.

Навчання відбувається за допомогою послідовного пред'явлення вхідних векторів з одночасною настройкою ваг у відповідності до заданого алгоритму. Під час навчання ваги НМ поступово стають такими, яби кожний вхідний вектор певним чином виробляв вихідний.

Навчальні алгоритми можна класифікувати як алгоритми навчання без учителя та з учителем.

У першій ситуації, при пред'явленні вхідних символів ШНМ самоорганізовується за рахунок налаштування власних ваг у відповідності до визначеного алгоритму. Внаслідок відсутності вказівки наперед окресленого виходу, під час навчання результати є непрогнозованими з погляду визначення збудливих символів для певного набору нейронів.

При цьому, НМ організовується у формі, яка відбиває істотні властивості навчальної сукупності. Наприклад, вхідні символи можуть класифікуватися у відповідності до рівня їхньої схожості так, що символи одного класу будуть активізувати той самий вихідний нейрон.

Окреслені способи навчання НМ передбачають, що відомі тільки вхідні вектори, а на їх базі мережа може вчитися більш кращим значенням вихідної функції.

У другій ситуації існує вчитель, який надає вхідні символи НМ, спів ставляє результати виходу з потрібними, а потім настроює ваги мережі таким, аби зменшувалося розходження та збільшувалася подібність.

Способи навчання ШНМ з учителем передбачають наявність пари «вхід-вихід», інакше кажучи є відомими значення вхідних векторів, та значення вихідних векторів, які їм відповідають.

Отже, ШНМ, яка вчиться з учителем, є надійнішою, адже при певному вхідному сигналі на виході відбувається відповідне вихідне значення.

Навчаючи НМ розпізнавати символи, часто знищуються чи трансформуються результати навчання, проведеного раніше. В разі, коли є лише фіксована сукупність навчальних векторів, - вектори здатні пред'являтися при

навчанні циклічним методом. У НМ зі зворотнім розповсюдженням, навчальні вектори подаються на вхід ШНМ послідовно, доки мережа не стане «навченою» у відповідності до загальної вхідної сукупності.

Однак повністю навчена НМ має запам'ятати новий вектор навчання, і він спроможний поміняти ваги аж настільки, що буде потрібно повністю перенавчати мережу.

АРТ є методом вирішення такої проблеми. Мережі та алгоритми АРТ зберігають пластичність, яка є необхідною для вивчення нових образів, у той же час, не дозволяючи змінюватися раніше запам'ятованим образам, і через те вони бувають ефективнішими під час розпізнавання символів.

Мережа АРТ являє собою векторний класифікатор. Вхідний вектор класифікується у відповідності до того, на яку із множин раніше запам'ятованих образів він є подібним.

Своє класифікаційне рішення така ШНМ демонструє у формі збудження одного з нейронів розпізнавального шару. Якщо вхідний вектор не відповідає ні одному з запам'ятованих образів, проектується нова категорія за рахунок запам'ятовування образів, подібних до нового вхідного вектору. Якщо з'ясовано, що вхідний вектор подібний до одного із раніше запам'ятованих векторів точки зору визначеного критерію схожості, запам'ятований вектор буде мінятися (навчатися) під впливом нового вхідного вектора, аби стати більш подібним до цього вхідного вектору.

Запам'ятований символ не буде мінятися, якщо поточний вхідний вектор не буде достатньо подібним до нього.

Таким чином, вирішується проблемне питання «стабільність-пластичність» функціонування НМ при виконанні алгоритмів із розпізнавання символів. Новий образ здатний проектувати додаткові класифікаційні властивості, проте новий вхідний образ не може примусити змінюватися наявну пам'ять.

Мережа АРТ навчається без учителя за рахунок зміни ваг таким чином, аби пред'явлення НМ вхідного вектора примушувало мережу активувати нейрони в

шарі розпізнавання, які зв'язані з подібним запам'ятованим вектором. Крім того, навчання відбувається у формі, яка не руйнує запам'ятовані раніше образи, упереджуючи тимчасову нестабільність, що регулюється вибором властивості схожості.

АРТ є методом вирішення такої проблеми. Мережі та алгоритми АРТ зберігають пластичність, яка є необхідною для вивчення нових образів, у той же час, не дозволяючи змінюватися раніше запам'ятованим образам, і через те вони бувають ефективнішими під час розпізнавання символів.

Мережа АРТ являє собою векторний класифікатор. Вхідний вектор класифікується у відповідності до того, на яку із множин раніше запам'ятованих образів він є подібним.

Своє класифікаційне рішення така ШНМ демонструє у формі збудження одного з нейронів розпізнавального шару. Якщо вхідний вектор не відповідає ні одному з запам'ятованих образів, проектується нова категорія за рахунок запам'ятовування образів, подібних до нового вхідного вектору. Якщо з'ясовано, що вхідний вектор подібний до одного із раніше запам'ятованих векторів точки зору визначеного критерію схожості, запам'ятований вектор буде мінятися (навчатися) під впливом нового вхідного вектора, аби стати більш подібним до цього вхідного вектору.

Новий вхідний образ, мається на увазі той, що НМ «не бачила» раніше, не буде відповідати запам'ятованим образам з точки зору властивості ідентичності, тим самим проектуючи новий образ, який запам'ятовується.

Вхідний образ, у достатній мірі є відповідним одному із образів, що запам'ятовуються, не буде створювати новий екземпляр, а він буде тільки модифікувати той, до якого він подібний.

Отже, при певному виборі ознаки схожості упереджується запам'ятовування раніше вивчених образів і певна структура нестабільності.

Таким чином, для навчання НМ пропонується спосіб, за яким особа, яка управляє мережею, приймає безпосередню участь у процесі навчання НМ, вона сама визначає еталонні зображення загальної кількості символів, а також перекручені зображення еталонних зразків (копій). Раніше такий спосіб використовувався до ймовірнісних НМ, однак застосування цього способу для мереж багат шарових нейронів суттєво підвищить його продуктивність під час розпізнавання символів.

Висновки до розділу 3

Підсумовуючи третій розділ, можемо зробити такі висновки:

1. Охарактеризовано методику синтезу ШНМ для розпізнавання символів та побудовано модель ШНМ для поставленого завдання розпізнавання символів за рахунок використання багат шарового перцептронів. У результаті впровадження інтелектуального прототипу, який функціонує згідно із пропонованим алгоритмом, можуть розпізнаватися символи, які належать до різних алфавітів, цифрові символи та знаки в залежності від об'єму пам'яті ІС, яка включає в себе певні еталони із зашумленістю до 40 %.

2. Проаналізовано принцип використання алгоритму комбінації двох мереж з однаковою структурою. Об'єднуючись у мережі, ШНМ формують систему обробки інформації, яка забезпечує ефективне пристосування моделі до постійних трансформацій з боку оточуючого середовища. Під час роботи НМ відбувається видозміна вхідного вектора сигналів у вихідний. Конкретний вид трансформації задається архітектурою НМ, властивостями нейронних компонентів, інструментами управління та синхронізації інформаційних даних між нейронами.

3. Розроблено методичні рекомендації щодо використання нейронних мереж для розпізнавання символів. Пропонується реалізація навчання мережі за допомогою адаптивної резонансної теорії, що має ряд переваг.

ВИСНОВКИ

Підсумовуючи загальний зміст наукової роботи, можемо констатувати наступне:

1. Інтелектуальні інформаційні технології наразі є одним з ключових напрямків розвитку інформатики. Дослідження за формою і методам створення ІТ для вирішення різних наукових і практичних завдань активно проводяться як в Україні, так і за кордоном. Особливо важливим в цих дослідженнях є вивчення та використання на практиці нових підходів до реалізації ІТ, оскільки в більшості випадків традиційні методи вимагають великих затрат часу і матеріальних вкладень. Останнім часом для вирішення складних завдань управління, інформаційного моніторингу, діагностики, автоматизованого проектування, розпізнавання символів, класифікації використовують технології штучних нейронних мереж.

2. Проблема розпізнавання символів користується все більшою популярністю, адже розпізнавання зображення, тексту чи мови, а також різноманітних явищ сприяє спрощенню комунікативного зв'язку людини з персональним комп'ютером (ПК), допомагає застосовувати різні системи штучного інтелекту в різних галузях народного господарства. Можливість сприймати зовнішній світ у формі символів сприяє передумовам дослідження властивостей величезної кількості об'єктів завдяки ознайомленню з кінцевою їх кількістю, а об'єктивна ознака засадничої властивості символів допомагає створювати модель їх розпізнавання. Крім того, системи розпізнавання символів, що використовують штучний інтелект (ШІ), наразі отримали активний розвиток та з успіхом впроваджуються в практичному житті. Поліпшується як науковий рівень ШІМ, так і математичні потужності обладнання технічного гатунку.

3. Для рішення завдань оптимізації в СРС часто застосовуються високоінтелектуальні програми на базі ШІМ. Проте такі ШІМ не є засобом для вирішення будь-яких завдань. Вони є непридатними для вирішення таких задач як

нарахування заробітної плати, однак такі системи мають перевагу в процесі розпізнавання символів, з майже не справляються традиційні ПК.

4. Нейронна мережа являє собою ланцюг нейронів, іншими словами ШНМ мережа складається з нейронів штучного характеру або вузлів. Отже, НМ - це або біологічна нейронна мережа, що складається з реальних біологічних нейронів, або ШНМ, яка створюється для вирішення завдань штучного інтелекту. ШНМ можуть використовуватися для прогнозного моделювання, адаптивного управління і додатків, де їх можна навчати за допомогою набору даних. Самонавчання на основі досвіду може відбуватися в мережах, які можуть робити висновки зі складного і, здавалося б, незв'язаного набору інформації. На даний час НМ являють собою програми, які мають таку архітектуру, використання якої сприяє умовній імітації роботі нейронів. Математичний прототип нейрона являє собою певний універсальний нелінійний компонент із здатністю широкої трансформації та налаштування його властивостей (параметрів).

5. Штучні нейронні мережі поділяються:

- По типу вхідної інформації:

- аналогові (використання інформації в формі дійсних чисел);
- двійкові (інформація представлена в двійковому коді).

- По характеру навчання:

- з «вчителем» (під навчання відомо набір вихідних значень);
- без «вчителя» (формує вихідні значення тільки на основі вхідних впливів),

їх називають самоорганізуючими.

- По характеру налаштування синапсів:

- з фіксованими зв'язками (вагові коефіцієнти вибираються на початковому етапі виходячи з даної задачі);
- з динамічними зв'язками (в процесі навчання відбувається настройка вагових коефіцієнтів).

- За способом навчання:

- навчання по алгоритму зворотнього розповсюдження похибки;

- з конкурентним навчанням;
- з навчанням по правилу Хебба;
- з гібридним навчанням.

- По характеру зв'язків:

- з прямими зв'язками (розповсюджується інформація тільки в одному напрямку від рівня до рівня: це різноманітні персептронні мережі);
- із зворотнім поширенням інформації (релаксійні та багаторівневі).

6. Кластерний аналіз спрямований на виявлення природного поділу об'єктів. Тобто, він об'єднує спостереження, які є схожими на підмножини. Такі підкласи мають певну закономірність, яка має безпосереднє відношення до досліджуваного явища. Функціональна особливість виміру відстані використовується в якості аналізу, якщо схожість між об'єктами і широким спектром кластеризації алгоритму на базі різних інтерпретацій. Заходи подібності спершу обраховується між спостереженнями та між кластерами під час групування спостережень в кластери. Кілька метрик, таких як евклидово і Манхеттенська відстань, кореляція, можуть застосовуватися для обчислення схожості. Крім того, видаються можливими кілька стратегій об'єднання, що здатні призвести до різних шаблонів кластеризації.

7. Галузі застосування нейронних мереж для розв'язання практичних задач достатньо широкі. Серед них:

- Розпізнавання образів і класифікація.
- Кластеризація.
- Прогнозування і апроксимація.
- Рішення завдань при невідомих закономірностях.
- Прийняття рішень і керування.
- Стиск даних і асоціативна пам'ять.
- Оптимізація.

8. Навчаючи НМ розпізнавати нові символи, часто втрачаються чи змінюються результати навчання, що було проведене раніше. Якщо є лише

фіксована сукупність навчальних векторів - вони зазвичай пред'являються в процесі навчання циклічно. У мережах зворотнього розповсюдження, навчальні вектори подаються до входу мережі поступово, поки мережа не стане «навченою» по відношенню до загальної вхідної сукупності. Проте повністю навчена НМ має запам'ятати новий вектор навчання, і цей вектор здатний змінити ваги настільки, що буде потрібно заново перенавчати мережу. Адаптивна резонансна теорія (АРТ) є вирішенням даної проблеми. Мережі та алгоритми АРТ зберігають пластичність, яка є конче необхідною для аналізу нових образів, у той же час, запобігаючи змінам образів, що запам'ятовується раніше, і через те вони є ефективнішими при розпізнаванні символів. Мережа АРТ являє собою певний векторний класифікатор. Вхідний вектор класифікується у відповідності до того, на яку із множин раніше запам'ятованих образів він є подібним. Своє класифікаційне тлумачення така НМ демонструє збудженням одного з нейронів розпізнавального шару. В ситуації, коли вхідний вектор не є відповідним жодному із запам'ятованих образів, моделюється нова категорія за рахунок запам'ятовування образів, ідентичних до нового вхідного вектора. Якщо з'ясовано, що вхідний вектор подібний до одного з раніше запам'ятованих векторів з точки зору визначеної властивості схожості, запам'ятований вектор буде мінятися (тобто проходити навчатися) за рахунок нового вхідного вектора, аби стати більш подібним до нього. Запам'ятований символ не буде мінятися, якщо поточний вхідний вектор не буде подібним до нього. Отже, вирішується проблемне питання «стабільність-пластичність» роботи ШНМ при виконанні дій із розпізнавання символів. Новий образ здатен моделювати додаткові класифікаційні параметри, проте новий вхідний образ не спроможний примусити змінитися наявну пам'ять.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Авалиани Г. В. Эвристические методы в распознавании образов / Г. В. Авалиани. – Тбилиси: Мецниереба, 1988. – 78 с.
2. Антонюк В.С. Методологія наукових досліджень: [Текст] : навч. посіб./ В.С. Антонюк, Л.Г. Полонський, В.І. Аверченков, Ю.А. Малахов. – К.: НТУУ «КПІ», 2015. – 286 с.
3. Боголюбов Д. П., Чанкин А. А., Стемиковская К. В. Реализация алгоритма обучения самоорганизующихся карт Кохонена на графических процессорах с // Промышленные АСУ и контроллеры. 2012. - № 10. - С. 30-35.
4. Борисов Ю.И. Нейросетевые методы обработки информации и средства их программно-аппаратной поддержки / Борисов Ю.И, Кашкаров В.М., Сорокин С.А. // Открытые системы. – 1997.– № 4. – С. 38 – 40.
5. Вентцель Е. С. Теорія ймовірностей: Підр. для вузів. - 6-е вид.стер. - М.: Вищ. шк., 1999. – С. 12-54.
6. Воеводин В. Параллельные вычисления / Воеводин В .В., Воеводин В. В . – СПб .: БХВ – Петер - бург , 2002. – 608 с.
7. Вороновський Г. К., Махотило К. В., Петрашев С. Н., Сергеев С. А. Генетичні алгоритми, штучні нейронні мережі і проблеми віртуальної реальності. - Замовне. - Х.: ОСНОВА, 1997. - С. 99-112.
8. Галушкин А.И. Теория нейронных сетей. / А.И. Галушкин - М.: ИПРЖР, 2010. – 348 с.
9. Головач Ю. Складні мережі / Ю. Головач, О. Олемской, К. фон Фербер та ін. // Журнал фізичних досліджень. - 2006. - 10, № 4. - С. 247-289.
10. Головкин Б.А. Машинное распознавание и линейное программирование. - М.: Советское радио. 1973. - 100 с.
11. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. / А.Н. Горбань, Д.А. Россиев – Н.: Наука, 2006. – 276с.

12. Горбань А.Н. Нейронные сети на персональном компьютере / А.Н. Горбань Д.А. Россиев. – Новосибирск: Наука, 2006. – 230 с.
13. Глушков В.М., Амосов Н.М., Артеменко И.А. Энциклопедия кибернетики. Том 2. Киев, - 1974 г. – С. 33-54.
14. Гуревич И.Б. Тезаурус и Онтология Предметной Области «Анализ Изображений» / И.Б. Гуревич, Ю.О. Трусова // II Всероссийская конференция «Знания – Онтологии – Теории». (ЗОНТ-09). 20–22 октября 2009 г. 2009. – С. 2-12.
15. Дейч А. М. Методи ідентифікації динамічних об'єктів. – М: Енергія, 1979 – 240 с.
16. Дремин И. М., Иванов О. В., Нечитайло В. А. Вейвлеты и их использование //Успехи физических наук. - 2001. - Т. 171. - № . 5. - С. 465-501.
17. Комашинский В.И. Смирнов Д.А. Внедрение в нейро-информационные технологии. / В.И. Комашинский, Д.А. Смирнов - СПб, 1999. – С. 33-48.
18. Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. - 1-е. - М.: Горяча лінія - Телекому, 2001. - С. 382.
19. Лисе А.А., Степанов М.В. Нейронные сети и нейрокомпьютеры. / А.А. Лисе, М.В. Степанов // Учеб.пособие. ГЭТУ. - СПб, - 2009. - 64 с.
20. Месюра В.І., Ваховська Л.М. Основи проектування систем штучного інтелекту: Навчальній посібник. - Вінниця: ВДТУ, 2000. - 96 с.
21. Моделі нейронних мереж. – Режим доступу: <https://studme.com.ua/1246122010028/neural/models.htm> .
22. Моделі нейронних мереж. – Режим доступу: <http://techn.sstu.ru/kafedri/подразделения/1/MetMat/Terin/neiro/neiro.htm> .
23. Нейромережеве відображення дійсності. – Режим доступу: http://studies.in.ua/mpd_seminar/1313-neyrmerezh.html .
24. Поліщук Д. О. Комплексне детерміноване оцінювання складних ієрархічно-мережевих систем: I. Опис методики / Д. О. Поліщук, О. Д. Поліщук, М. С. Яджак // Системні дослідження та інформаційні технології. - 2015. - № 1.- С. 21–31.

25. Потапов А.С. Распознавание образов и машинное восприятие: Общий подход на основе принципа минимальной длины описания. – СПб.: Политехника, 2007. – 548 с.
26. Поспелов Г.С. Искусственный интеллект - основа новой информационной технологии - М.: Высшая школа, 1988 – 280 с.
27. Роберт Каллан. Основні концепції нейронних мереж = TheEssenceofNeuralNetworksFirstEdition. - 1-е. - «Вильямс», 2001. - С. 268-288.
28. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский. - М. Горячая линия-Телеком 2004. – 112 с.
29. Саймон Хайкин. Нейронні мережі: повний курс = NeuralNetworks: AComprehensiveFoundation. - 2-е. - М.: «Вильямс», 2006. - С. 1099-1104.
30. Советов Б.Я. Информационные технологии / Б.Я. Советов, В.В. Цехановский - М.: Высшая школа, 2005 - 263 с.
31. Тархов Д.А. Нейронные сети. Модели и алгоритмы. – М.: Радиотехника, 2010. – 82 с.
32. Терехов В. А., Єфімов Д. В., Тюкин И. Ю. Нейромережні системи керування. - 1-е. - Высшая школа, 2002. - С. 156-184.
33. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978 – 412с.
34. Уосермен Ф. Нейрокомп'ютерна техніка: Теорія і практика. Переклад українською І.Ю.Юрчак, 2001 р. – С. 3-15.
35. Фомин Я. А. Распознавание образов: теория и применения / Я. А. Фомин. – М.: ФАЗИС, 2012. – 429 с.
36. Царегородцев В.Г. Перспективы распараллеливания программ нейросетевого анализа и обработки данных / В.Г. Царегородцев // Материалы с Всерос. конф. «Математика, информатика, управление – 2008». – Иркутск, с2014. – С. 110-117.

37. Шапиро Л., Стокман Д. Компьютерное зрение //М.: Бинوم. Лаборатория знаний. - 2006. - Т. 752. – С. 66-69.
38. Ясницкий Л. Н. Введение в искусственный интеллект. - 1-е. -Издательский центр «Академия», 2005. -С. 145-176.
39. Asanovic, Krste et al. (Jan 18, 2016). The Landscape of Parallel Computing Research: A View from Berkeley University of California, Berkeley. Technical Report No. UCB/EECS-2016-183.
40. Barbara Chapman, Gabriele Jost, Ruud van der Pas. Using OpenMP: portable shared memory parallel programming (Scientific and Engineering Computation). Cambridge, Massachusetts: The MIT Press., 2008. - 353 pp.
41. Hong S.G., Kim S.W. and Lee J.J., 2015. The Minimum Cost Path Finding Algorithm Using a Hopfield Type Neural Network, Proceedings IEEE International Conference on Fuzzy Systems 4, 719–726.
42. Salamon J. A Dataset and Taxonomy for Urban Sound Research / J. Salamon, C. Jacoby, J. Bello. // 22nd ACM International Conference on Multimedia, Orlando USA. – 2014.
43. Krizhevsky A. Imagenet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G. Hinton. // Advances in Neural Information Processing Systems. – 2012. – №1. – С. 1097–1105.

ДОДАТКИ



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО
ЗАБЕЗПЕЧЕННЯ



РОЗРОБКА ТА НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ

Студент: Буков Ростислав Дмитрович, ПДМ-61

Науковий керівник: к.т.н., доц., Щербина Ірина
Сергіївна

Київ-2021

АКТУАЛЬНІСТЬ РОБОТИ

2

Об'єктом дослідження є процес проектування та навчання нейронної мережі для розпізнавання символів.

Предметом дослідження виступає специфіка розробки та навчання нейронної мережі для розпізнавання символів.

Мета дослідження полягає в аналізі особливостей розробки та навчання нейронної мережі для розпізнавання символів.

ПОСТАВЛЕНА МЕТА ПЕРЕДБАЧАЄ РОЗВ'ЯЗАННЯ ТАКИХ ЗАВДАНЬ:

- розглянути теоретичні засади навчання нейронних мереж;
- дослідити класифікацію нейронних мереж;
- визначити алгоритми навчання нейронних мереж;
- проаналізувати особливості навчання нейронних мереж для СРС;
- розробити та застосувати алгоритм ШНМ для СРС.

МЕТОДИ ДОСЛІДЖЕННЯ:

- Методи системного аналізу;
- Спостереження;
- Узагальнення.

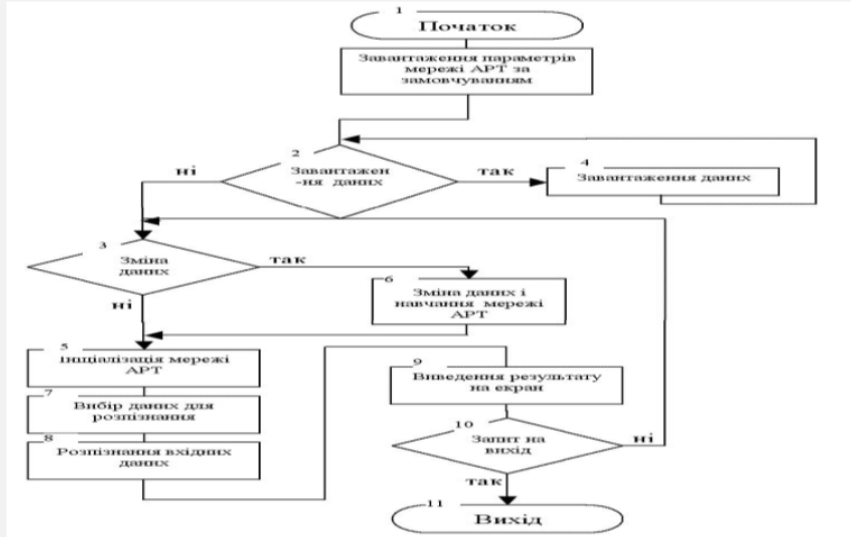
ПРАКТИЧНЕ ЗНАЧЕННЯ ОДЕРЖАНИХ РЕЗУЛЬТАТІВ:

- Полягає в тому, що дослідження ґрунтується на результатах поглибленого вивчення особливостей застосування перенесення навчання нейронних мереж, а також розроблено модуль на основі запропонованого алгоритму, який має перевагу в швидкості навчання в порівнянні з альтернативними підходами.

НАУКОВА НОВИЗНА ОДЕРЖАНИХ РЕЗУЛЬТАТІВ

- Результати дослідження пропонують альтернативний метод застосування штучного інтелекту під час розпізнавання символів.

СХЕМА АЛГОРИТМУ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ СИМВОЛІВ



АЛГОРИТМ РОЗПІЗНАВАННЯ СИМВОЛІВ

- 1) завантаження параметрів по замовчуванню під час запуску програми;
- 2) створення еталонних зображень (створивши еталони, користувач може перенавчити мережу так, щоб створений ним образ символу був еталоном);
- 3) завантаження чи створення зображення символу для розпізнавання;
- 4) розпізнавання символу (зашумленість якого може бути до 40 %) системою;
- 5) виведення результату розпізнавання.

АПРОБАЦІЯ

За матеріалами опубліковані тези:

- Буков Р. Д. Розробка та навчання нейронної мережі для розпізнавання символів. // Загально-університетська, науково-технічна конференція «Проблеми комп'ютерної інженерії»;

та стаття:

- Р. Д. Буков, І. С. Щербина «Дослідження систем розпізнавання тексту та вилучення даних з україномовних документів», Журнал "Зв'язок", №6, 2020, Київ.

ВИСНОВКИ

- Зазначений алгоритм може бути реалізований у середовищі Borland Delphi 7, на мові Object Pascal, яка дозволяє використовувати бібліотеки, що реалізують нейронні мережі. Результат реалізації приведеного алгоритму в середовищі програмування Borland Delphi 7 надав можливість реалізувати інтелектуальний модуль ШНМ із розпізнавання символів.
- В результаті реалізації інтелектуального модуля, що функціонує у відповідності із запропонованим алгоритмом, розпізнаються символи, які належать до будь-якого алфавіту, цифри й знаки в залежності від об'ємів пам'яті інтелектуальної системи, що містить відповідні еталони із зашумленістю до 40 %.