ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Розробка алгоритму розпізнавання рукописних символів в системах тестування»

на здобуття освітнього ступеня бакалавра
зі спеціальності 122 Комп'ютерні науки
(код, найменування спеціальності)
освітньо-професійної програми Комп'ютерні науки
(назва)

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Олександр УШАКОВ
(підпис) *(Ім'я, ПРІЗВИЩЕ здобувача)*

Виконав:
здобувач вищої освіти
група КНД-42 Олександр УШАКОВ

Керівник:
науковий ступінь, вчене звання Олег ІЛЬЇН
д.т.н. професор

Рецензент:
науковий ступінь, вчене звання _____
(Ім'я, ПРІЗВИЩЕ)

Київ 2024

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

Навчально-науковий інститут інформаційних технологій

Кафедра Комп'ютерних наук

Ступінь вищої освіти Бакалавр

Спеціальність 122 Комп'ютерні науки

Освітньо-професійна програма Комп'ютерні науки

ЗАТВЕРДЖУЮ
Завідувач кафедрою Віктор ВИШНІВСЬКИЙ
_____ Ім'я, ПРІЗВИЩЕ
«____»_____2024р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Ушаков Олександр Сергійович
(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: Розробка алгоритму розпізнавання рукописних символів в системах тестування

керівник кваліфікаційної роботи Олег ІЛЬЇН професор, д.т.н,
(Ім'я, ПРІЗВИЩЕ, науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологійвід «27» лютого 2024 р. №36

2. Строк подання кваліфікаційної роботи «31» травня 2024 р.

3. Вихідні дані до кваліфікаційної роботи: Збір та підготовка набору рукописних текстових даних є першим кроком у дослідженні методів розпізнавання;
визначення архітектури нейронної мережі для розпізнавання рукописного тексту;
Аналіз набору даних для визначення особливостей, таких як розмір зображень, тощо;
Вибір параметрів нейронної мережі, таких як кількість шарів та швидкість навчання.
Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
1. Вступне пояснення теми та актуальності дослідження.
2. Дослідження концепцій розпізнавання тексту за допомогою штучного інтелекту.
3. Аналіз технологій та алгоритмів, що використовуються в OCR.
4. Нейронні мережі в OCR: ефективність та застосування.

4. Перелік ілюстративного матеріалу: *презентація*

1.Мета розпізнавання рукописного тексту.
2.Проблеми та виклики у зберіганні рукописних текстів.
3.Основні методи розпізнавання тексту.
4.Архітектура моделі OCR для розпізнавання тексту.
5.Висновки.

5. Дата видачі завдання «23» лютого 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Проведення огляду наукових робіт та підготовка необхідних матеріалів для подальшої роботи	15.04.2024 р.	
2	Глибоке вивчення основних концепцій та методів нейронних мереж, а також їх застосування у розпізнаванні тексту.	20.04.2024 р.	
3	Аналіз технологій та методів, використовуваних для розпізнавання тексту з використанням штучного інтелекту.	25.04.2024 р.	
4	Вивчення алгоритмів та технологій, що застосовуються в OCR для ефективного розпізнавання тексту зі зображень.	10.05.2024 р.	
5	Написання тексту розрахунково-пояснювальної записки та підготовка презентації результатів дослідження.	15.05.2024 р.	
6	Оформлення та завершення роботи над кваліфікаційною роботою.	17.05.2024 р.	
7	Підготовка доповіді до захисту.	20.05.2024 р.	

Здобувач(ка) вищої освіти _____ Олександр УШАКОВ
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Керівник
кваліфікаційної роботи _____ Олег ІЛЬЇН
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

РЕФЕРАТ

Текстова частина кваліфікаційної роботи: 60 сторінок, 15 рисунків, 19 джерел.

Наукове завдання – розробка та вдосконалення моделі нейронної мережі для точного та швидкого розпізнавання рукописного тексту з метою покращення якості та ефективності процесу OCR.

Мета роботи – аналіз принципи та методи роботи нейронних мереж та машинного навчання з метою розробки ефективної моделі для розпізнавання рукописного тексту.

Об'єкт дослідження – нейронні мережі та методи машинного навчання для розпізнавання рукописного тексту.

Предмет дослідження – принципи роботи нейронних мереж та алгоритми машинного навчання, які використовуються для ефективного розпізнавання рукописного тексту.

Короткий зміст роботи: Проведено аналіз сучасних технологій розпізнавання символів та методів оцифрування текстових даних. Розроблено модель для оптичного розпізнавання рукописного тексту з використанням мови програмування Python.

Проведено тестування якості розпізнавання та визначено її ефективність.

Розроблена платформа дозволяє ефективно скорочувати час розробки нових веб-платформ для оптичного розпізнавання тексту. Застосування цієї платформи значно зменшує час, необхідний для аналізу збережених рукописних текстів.

КЛЮЧОВІ СЛОВА: АНАЛІЗ, НЕЙРОННІ МЕРЕЖІ, РУКОПИСНИЙ ТЕКСТ, ОПТИЧНЕ РОЗПІЗНАВАННЯ СИМВОЛІВ, ЗГОРТКОВІ МЕРЕЖІ, РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ

ЗМІСТ

ВСТУП

Актуальність дослідження. У сучасному технологічному світі важливість розпізнавання рукописного тексту надзвичайно висока, оскільки воно забезпечує конвертацію рукописних записів в формат, зрозумілий комп'ютерам. Ця технологія має широкі перспективи застосування, включаючи оцифрування рукописних документів, покращення зв'язку, автоматизовану обробку даних та пошук інформації.

Дослідження в цій області розширить наше розуміння методів нейронних мереж для розпізнавання рукописного тексту, що буде корисним ресурсом з потенційно великим впливом на різні галузі промисловості. Отримані знання сприятимуть покращенню обробки документів, транскрипції послуг та систем пошуку інформації.

Значна увага була приділена дослідженню методів розпізнавання рукописного тексту через їх потенцій для покращення ефективності та точності різних процедур. Поліпшення передових алгоритмів машинного навчання, таких як нейронні мережі, сприяло появі нових можливостей для підвищення рівня точності в цій області.

Мета полягає в досягненні підвищеної точності та продуктивності у розпізнаванні рукописних символів, що відкриває шлях до нових технологічних досягнень і використання їх у практичних сферах.

1 ПІДХІД ДО АНАЛІЗУ ТЕКСТУ З ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

1.1 Роль та поняття штучного інтелекту в розпізнаванні тексту

Штучний інтелект відіграє значну роль у розпізнаванні тексту, що дозволяє автоматизувати та полегшити обробку великих обсягів текстової інформації. Однією з ключових задач є автоматичне розпізнавання тексту зі зображень або сканованих документів за допомогою методів машинного навчання. Штучний інтелект дозволяє створювати моделі, які можуть виявляти та класифікувати текстову інформацію з високою точністю.

Однією з технік штучного інтелекту, що використовується в розпізнаванні тексту, є моделі глибокого навчання, такі як рекурентні нейронні мережі (RNN) та трансформери. Ці моделі можуть аналізувати послідовності слів та контекст, розуміти семантику тексту та виявляти патерни у данних. Вони використовуються для завдань розпізнавання іменованих сутностей, аналізу настроїв, автоматичного перекладу мов та інших текстових завдань. Штучний інтелект дозволяє покращити точність та швидкість розпізнавання тексту, що робить його важливим інструментом у сучасній обробці мовленнєвої інформації.

Нейронні мережі є надзвичайно потужними математичними інструментами, які використовуються в усьому: від класифікації даних до автономної навігації транспортних засобів і прогнозів фондового ринку. Нейронні мережі були розроблені для точного моделювання взаємодії між нейронами, що вимагає обговорення основної біології. Тіло людини складається з багатьох клітин, які виконують різні функції, і для цілей цього обговорення ми зосередимося на клітинах нервової системи. У нервовій системі є два основних типи клітин: нейрони та глія. Ненейрональні гліальні клітини відіграють вирішальну роль у підтримці фізіологічного гомеостазу, формуванні мієліну та участі в передачі сигналу. Особливе значення для наших цілей мають нейрони як основні будівельні блоки мозку. Мозок складається з багатьох нейронів, кожен з яких складається з

дендритів, клітинних тіл і аксонів.

Рисунок 1.1 ілюструє анатомію типового нейрона, показуючи його три окремі домени. Дендрити характеризуються розгалуженими та дерев'янистими структурами та виконують важливу функцію передачі електрохімічних сигналів, отриманих від сусідніх нервових клітин, до тіла клітини. Тіло клітини містить ядро та аксональний горбок і відіграє ключову роль у функції нейрона. Аксональний горбок — це спеціалізована ділянка в тілі клітини, яка позначає точку ініціації аксона та містить велику кількість енергозалежних іонних каналів.

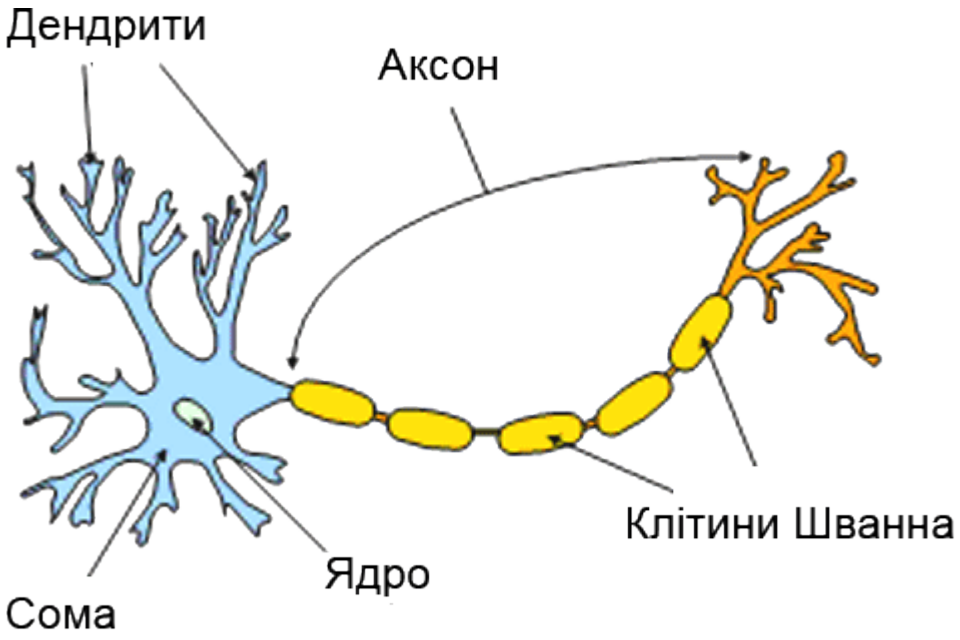


Рис. 1.1 – Анатомія типового нейрона

Ці канали часто є місцем, де ініціюються потенціали дії. Точніше, коли електрохімічні сигнали досягають тіла клітини, вони консолідуються на аксональному горбку. Якщо сума цих сигналів перевищує певний поріг, генерується потенціал дії, який передається по аксону. На рисунку 1.2 показано візуальний поріг для спрацьовування потенціалу дії та результуюча вихідна напруга.

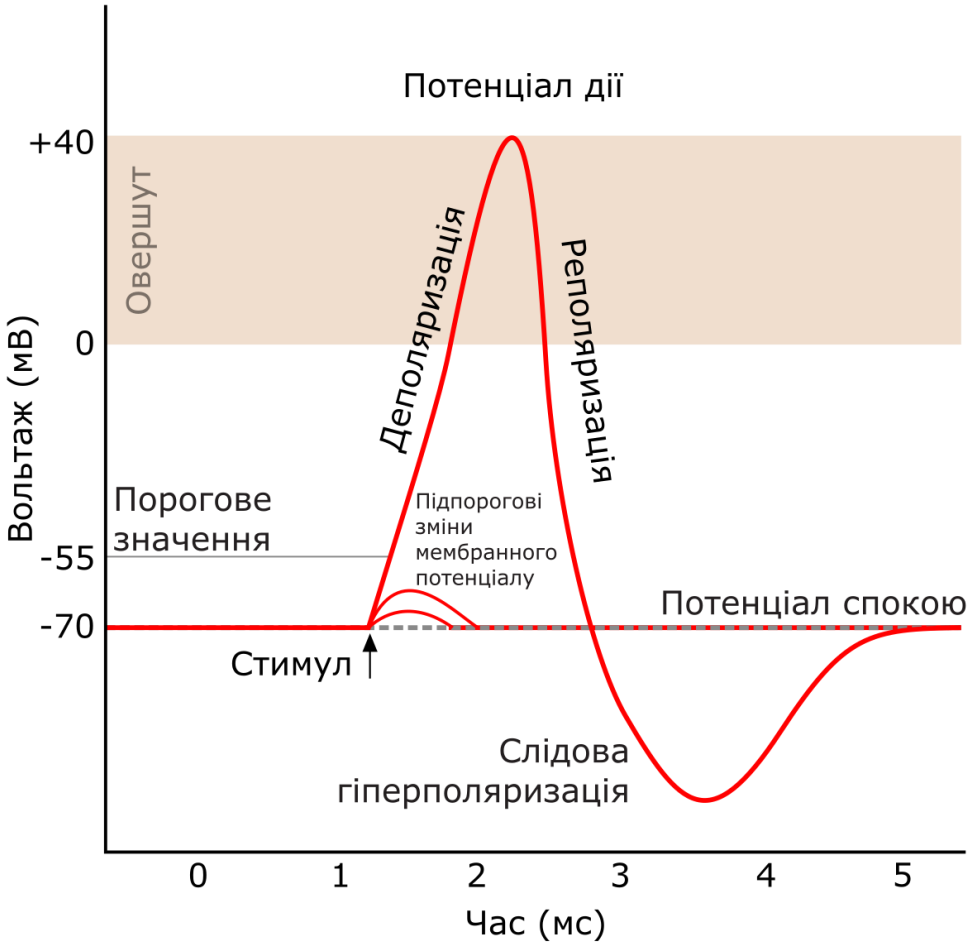


Рис. 1.2 – Візуальний поріг для спрацьовування потенціалу

Щоб отримати глибше розуміння окремого нейрона, давайте подивимося на нього з математичної точки зору, зосередившись на тому, як він представлений як вузол, який обробляє вхідні дані та генерує вихідні дані. Наш аналіз почнеться з вивчення дендритів, які ми називаємо вхідним шаром. Кожен n-дендрит передає електрохімічний сигнал з певною вагою. Символічно ми позначаємо вузли вхідного рівня як x1, x2, ..., xn, а відповідні ваги як w11, w12, ..., w1n, де wij представляє вагу, пов'язану з передачею xj до вузла i.

На рисунку 1.3 показана конфігурація одного нейрона, що складається з трьох вхідних вузлів x1, x2 і x3, і їх відповідних ваг w1, w2 і w3. Зверніть увагу, що нижній індекс i опущено, оскільки є вузол, хоча він буде представлений як w11, w12 і w13. Важливо розуміти, що коли мережа стає складнішою, знаки ваг і вузлів також ставатимуть складнішими.

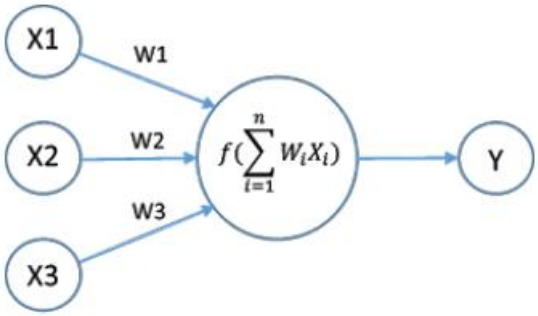


Рис. 1.3 – Математична конфігурація одного нейрона

Нейронна мережа — це обчислювальна модель, яка може ідентифікувати шаблони в наборах даних [11]. Отримавши достатньо навчальних даних, нейронна мережа може навчитися передбачати результати на основі майбутніх даних.

По суті, це можна описати як комп'ютерну систему, розроблену для імітації когнітивних здібностей людського мозку. Незважаючи на те, що концепція може здатися фантастичною, автомобіль дійсно може демонструвати подібні можливості. Чи перевершать машини інтелектуальні здібності людського мозку та досягнуть вищого рівня інтелекту, залишається питанням, яке матимуть можливість досліджувати майбутні покоління.

Звичайна комп'ютерна система використовує двійковий код для виконання операцій, однак нейронна мережа також використовує код, який не є бінарним, але натомість має здатність змінювати напрямок цих операцій. Наприклад, якщо автомобіль заправляється певним кодом і системою чисел, він буде суворо дотримуватися цих правил. Однак цій подвійній системі не вистачає здатності автомобіля сприймати такі перешкоди, як великий камінь, що блокує дорогу. Через нейронну мережу автомобіль здатний розпізнавати потенційні проблеми під час роботи, включаючи наявність каменів, якщо вони є, автомобіль уникне зіткнення, маневруючи навколо них. Нейронні мережі мають здатність розпізнавати помилки в певних кодах, що дозволяє їм самостійно вирішувати проблему або вживати інших дій без негативного впливу на процес.

Штучні нейронні мережі покликані відтворювати складну організацію людського мозку та отримувати інформацію з навколишнього середовища,

обробляти її, а потім реагувати. Фундаментальна концепція штучного інтелекту полягає в тому, що машина може володіти когнітивними здібностями, якщо їй надається багато даних. Важливо визнати, що машини насправді не можуть мислити, так само як у монстра Франкенштейна відсутні людські риси. Однак ключовим аспектом є те, що чим більше інформації зберігається в комп'ютері, тим потужнішим і універсальнішим він стає. Наприклад, програмісти IBM навмисно запрограмували кожен можливий шаховий хід і стратегію в комп'ютері, це створило пристрій, який міг обчислити всі можливі результати і навіть передбачити хід суперника. Це явище відоме як банк пам'яті.

Нейронні мережі служать спрощеним представленням когнітивних процесів, залучених до розпізнавання рукописного тексту. Використовуючи цю передову технологію, машини можуть випередити людську майстерність у цій конкретній справі. Через різноманітний і складний характер індивідуального стилю почерку розшифровка рукописних текстів може бути трудомісткою та виснажливою. Використання нейронної мережі в запропонованій системі є найефективнішим рішенням, ця мережа є винятковою у розпізнаванні значущих закономірностей у складних даних і ідентифікації потенційних альтернативних людям методів.

Штучна нейронна мережа - це математична модель, яка імітує роботу нейронної системи людини і використовується для вирішення різноманітних завдань у галузі штучного інтелекту. Вона складається зі штучних нейронів, які організовані у велику мережу і спроможні здійснювати обчислення та приймати рішення на основі вхідних даних.

Кожен штучний нейрон приймає вхідні сигнали, обробляє їх за допомогою внутрішніх параметрів (ваги) та функцій активації, і генерує вихідний сигнал. Нейрони організовані у шари, де вхідний шар отримує вхідні дані, а вихідний шар генерує результат. Проміжні шари називаються прихованими і виконують проміжні обчислення.

Штучні нейронні мережі можуть бути використані для різних завдань, таких як класифікація, прогнозування, розпізнавання образів, машинний переклад та багато інших. Вони демонструють великий потенціал у вирішенні складних

завдань у багатьох галузях, включаючи комп'ютерне зорове визначення, медичну діагностику, фінансовий аналіз та автономне керування транспортними засобами.

Штучна нейронна мережа здатна здійснювати аналіз великої кількості даних і виявляти складні залежності між ними. Вона може виконувати різноманітні завдання аналізу, такі як класифікація, прогнозування, впізнавання образів та генерація нових даних.

Штучна нейронна мережа вивчає внутрішні залежності в навчальних даних і використовує цю інформацію для прийняття рішень. Вона може виявляти складні шаблони та невидимі зв'язки між різними ознаками даних, що дозволяє здійснювати точний аналіз та прогнозування майбутніх подій.

Завдяки своїй здатності до аналізу та узагальнення інформації, штучні нейронні мережі застосовуються в різних галузях, включаючи фінанси, медицину, технології, науку про дані та багато інших. Вони є потужним інструментом для виявлення складних закономірностей в даних і вирішення різноманітних завдань аналізу та прогнозування.

Штучні нейронні мережі використовуються у різних галузях та застосунках. Ось декілька прикладів їх використання:

1. Медицина, нейронні мережі використовуються для діагностики хвороб, аналізу медичних зображень (наприклад, рентгенівських знімків, МРТ) для виявлення патологій та лікування різних захворювань.

2. Фінанси, вони застосовуються для прогнозування фінансових ринків, аналізу портфеля інвестицій, виявлення шахрайства та аналізу кредитного ризику.

3. Комерційні додатки, наприклад, у рекомендаційних системах для персоналізованих пропозицій для покупців або у виявленні аномальних звернень клієнтів.

4. Автономні транспортні засоби, штучні нейронні мережі використовуються для автономного керування автомобілями, дронами та іншими транспортними засобами.

5. Комп'ютерне зорове визначення, вони використовуються для розпізнавання об'єктів та розуміння зображень у відеоспостереженні, роботах з реальним часом та інших додатках.

6. Мовна обробка, нейронні мережі застосовуються у системах машинного перекладу, розпізнаванні мови та аналізі настроїв.

Штучний інтелект (ШІ) - це галузь комп'ютерних наук, яка займається створенням систем та програм, які можуть виконувати завдання, які зазвичай потребують людського інтелекту. Штучний інтелект прагне створювати програмне забезпечення або машини, здатні мислити, вчитися, розуміти, приймати рішення та розв'язувати проблеми, схожі на ті, що вирішуються людьми.

Основні складові штучного інтелекту включають в себе машинне навчання, обробку природної мови, комп'ютерне бачення, планування, розпізнавання образів, робототехніку та багато інших. Штучний інтелект може бути використаний для автоматизації повсякденних завдань, покращення процесів виробництва, розробки нових продуктів та послуг, управління складними системами та багато іншого.

Штучний інтелект стає все більш важливим і потужним інструментом у багатьох сферах життя, включаючи медицину, фінанси, транспорт, освіту, науку та бізнес. Він відкриває нові можливості для вирішення складних проблем і покращення якості життя людей.

У результаті термін «ШІ» зазвичай застосовується до будь-якого підходу, який використовується в будь-якому сценарії, незалежно від того, вигаданий він чи ні. Однак, якщо цей термін призначений для демонстрації розумної поведінки, цей термін можна застосовувати до будь-якої художньої літератури. Наразі комп’ютери не можуть точно передбачити вартість властивостей без використання суворих методів машинного навчання або без ретельного обмірковування та кодування вони не можуть негайно повідомити керівнику будівлі про збій системи.

ШІ тепер є даниною інноваційній природі людського духу та постійному прагненню розвиватися. Це симбіотичне поєднання методів, витончена прогресія кодів і кульмінація людського таланту. Історична розповідь про штучний інтелект є даниною безмежним наслідкам, які виникають із сили творчості та людського

інтелекту. Пошуки штучного інтелекту були захоплюючою подорожжю, яка була наповнена численними захоплюючими моментами.

У 1956 році відбулася Дартмутська конференція, що визначила початок розвитку штучного інтелекту як окремої галузі досліджень. Подібно до міфічного фенікса, який воскрес з попелу, ця область стрімко поширювалася завдяки спільним зусиллям людей, що бажали розгадати загадку розумних машин. У 1980-ті роки з'явилася машинне навчання – інноваційна концепція, яка дозволила комп'ютерам самостійно розвиватися та адаптуватися. Світ дивувався, як ці "дива техніки" перевершили своїх творців і взяли кермо свого власного розвитку. Штучний інтелект (ШІ) все більше проникає в різні сфери нашого життя, починаючи від розпізнавання голосу й закінчуючи розпізнаванням облич. Він діє як непомітний помічник у нашій повсякденній рутині. Але наразі комп'ютери не в змозі точно прогнозувати значення властивостей та ефективно сповіщати операторів будівель про збої в системі, якщо не використовувати строгі методи машинного навчання. Існують різні підходи до машинного навчання, такі як контрольоване, неконтрольоване та навчання з підкріпленням. Контрольоване навчання передбачає тренування моделей за допомогою позначених прикладів, тоді як неконтрольоване навчання виявляє закономірності та зв'язки у непозначених даних. Навчання з підкріпленням означає тренування агента через його взаємодію з оточенням, отримуючи винагороди чи покарання за свої дії. Машинне навчання відіграє життєво важливу роль у галузі штучного інтелекту, імітуючи когнітивні здатності людини через обробку та використання сенсорних даних за допомогою різних процесів, таких як перетворення, зменшення, збереження, пошук й застосування. Люди можуть обробляти великі обсяги інформації, використовуючи абстрактні знання для поліпшення свого розуміння вхідних даних.

Присутність штучного інтелекту (AI) в нашому повсякденному житті набуває величезного значення. AI володіє дивовижною здатністю оптимізувати та прискорювати монотонні та трудомісткі завдання, що дає нам можливість приділяти увагу більш послідовним заняттям. Крім того, алгоритми, які

використовує AI, здатні ретельно досліджувати великі обсяги даних, що дозволяє адаптувати продукти, послуги та досвід відповідно до наших індивідуальних уподобань і потреб. Персональні помічники на основі штучного інтелекту, такі як Siri, Google Assistant і Amazon Alexa, інтегровані в смартфони, розумні колонки та інші пристрої та можуть виконувати різноманітні завдання, від встановлення нагадувань і надсилання повідомлень до відтворення музики та керування пристроями розумного будинку. Поява штучного інтелекту спричинила розвиток концепції "розумного будинку". Різноманітні пристрої на базі штучного інтелекту, такі як камери безпеки, терморегулятори та системи освітлення, мають можливість асимілювати інформацію про вподобання людини та згодом змінювати їх налаштування відповідно до їхньої повсякденної поведінки. Наприклад, розумний терморегулятор може ефективно зменшити споживання енергії, отримуючи інформацію про присутність або відсутність користувача в домогосподарстві. Таке підходить не лише для фінансової економії для користувача, а й сприяє більш екологічно свідомому та сталому способу життя. Організації поступово використовують віртуальних помічників та чат-ботів на основі штучного інтелекту для цілодобового обслуговування клієнтів. Ці чат-боти використовують обробку людської мови, щоб зрозуміти запити споживачів і надати необхідні відповіді. Вдалим прикладом є компанія H&M, яка використовує чат-боти на основі штучного інтелекту для надання клієнтської підтримки.

Важливий поворотний момент стався в 1956 році, коли дартмутська конференція виділила ШІ в окрему область досліджень. Як і міфічний Фенікс, що піднімається з попелу, Сфера штучного інтелекту швидко розширюється завдяки блиску здорового глузду, який намагається розгадати таємницю мислення machines.In у 1980-х роках ми стали свідками появи машинного навчання, інноваційної концепції, яка дозволила комп'ютерам розвиватися та адаптуватися самостійно. Світ був здивований, побачивши, як ці машини, подібні геніям, спіткали долю, незалежну від творців.

Штучний інтелект (ШІ) все більше проникає в різні аспекти нашого життя, від розпізнавання мови до розпізнавання облич.він діє як непомітний партнер у

нашому повсякденному житті. Однак в даний час, якщо не застосовуються суворі методи машинного навчання, комп'ютери можуть точно прогнозувати вартість майна і швидко повідомляти будівельних операторів про збої в роботі системи.

Існують різні категорії підходів машинного навчання, які охоплюють контрольоване навчання, неконтрольоване навчання та розвиток. Контрольоване навчання включає модель навчання з використанням позначених зразків, а неконтрольоване навчання виявляє закономірності та взаємозв'язки в немаркованих даних. Навчання підкріплення передбачає навчання агентів взаємодії з навколишнім середовищем та отримання винагород та покарань у відповідь на їхні дії. Машинне навчання відіграє важливу роль у галузі штучного інтелекту, який обробляє та використовує сенсорні дані за допомогою різних процесів, таких як перетворення, декомунізація, пам'ять, пошук та програми.Люди мають здатність обробляти великі обсяги інформації, використовуючи абстрактні знання, тим самим покращуючи розуміння вхідних даних. Подібним чином моделі машинного навчання, завдяки своїй адаптивності, можуть імітувати певні аспекти людського пізнання.

Присутність штучного інтелекту (ШІ) у нашому повсякденному житті дуже важлива. ШІ має неймовірну здатність оптимізувати та прискорювати монотонні та трудомісткі завдання, що дає нам можливість звернути увагу на більш послідовні зусилля. Крім того, алгоритми, що використовуються в штучному інтелекті, можуть всебічно досліджувати великі обсяги даних і адаптувати продукти, послуги та досвід до індивідуальних переваг і потреб.

Персональні помічники на основі штучного інтелекту, такі як Siri, Google Assistant та Amazon Alexa, інтегровані в такі пристрої, як смартфони та розумні динаміки, і можуть виконувати різні завдання, такі як встановлення нагадувань, надсилання повідомлень, відтворення музики та керування пристроями розумного дому.

З появою штучного інтелекту виникла концепція "розумного будинку". Різні пристрої на основі штучного інтелекту, такі як камери відеоспостереження, Термостати і системи освітлення, здатні засвоювати інформацію про переваги

людини і змінювати налаштування відповідно до повсякденного життя. Наприклад, розумний термостат може ефективно збільшити споживання енергії, отримуючи інформацію про присутність або відсутність користувача в будинку. Таким чином, це не тільки економить фінансові кошти користувачів, але й сприяє більш екологічно свідомому та стійкому способу життя.

Організація поступово забезпечує цілодобове обслуговування клієнтів за допомогою віртуальних помічників та чат-ботів на основі штучного інтелекту. Ці чат-боти використовують обробку людської мови для розуміння вимог споживачів та надання їм необхідних відповідей. Хорошим прикладом цього є H & M, яка використовує чат-ботів на основі штучного інтелекту для забезпечення підтримки клієнтів. Ці чат-боти можуть відповідати на різні запити, включаючи відстеження замовлень та обробку повернення.

Гіганти електронної комерції, такі як Amazon, використовували вдосконалені алгоритми штучного інтелекту, щоб кардинально змінити спосіб взаємодії з клієнтами. Ці передові системи ретельно аналізують пошукові запити клієнтів, зберігають вашу історію перегляду приватною та перевіряють іншу відповідну інформацію. декомунізують пошукові запити клієнтів, зберігаючи конфіденційність вашої історії переглядів і перевіряючи іншу відповідну інформацію. Ефективно поєднуючи ці великі дані, ми вміло підбираємо індивідуальні рекомендації щодо продуктів відповідно до уподобань та потреб кожної людини.

Одним із прикладів штучного інтелекту є знаменитий Google Translate, який ефективно долає історично важливі бар'єри на шляху мовного бар'єру. Ці чудові інструменти призначені для швидкого розшифрування та інтерпретації як письмових, так і усних виразів, щоб вони могли зрозуміти, як відбуваються глобальні взаємодії.

Штучний інтелект також використовується в самохідних транспортних засобах, таких як автомобілі, вантажівки та автобуси, для ретельного вивчення та розуміння навколишнього середовища, ретельного визначення найбільш прибуткових маршрутів та прийняття обґрунтованих та важливих рішень щодо

водіння. Ця передова технологія може значно знизити ризик нещасних випадків, зменшити затори і зменшити погіршення стану навколишнього середовища. Цей типовий приклад-чудовий спосіб орієнтуватися в хаотичній дорожній мережі, основних магістралях і навіть обмежених місцях паркування без втручання людини.

1.2 Принципи та методи розпізнавання тексту за допомогою ШІ

Розпізнавання тексту за допомогою штучного інтелекту (ШІ) включає кілька принципів та методів, які використовуються для перетворення зображень тексту в машинозчитувані дані.

Принципи розпізнавання тексту

1. Збирання та підготовка даних, процес розпізнавання тексту починається зі збирання високоякісних зображень тексту. Ці зображення повинні бути чіткими та добре освітленими. Крім того, важливим етапом є анотація зображень, що може виконуватися вручну або автоматично, для точного визначення розташування тексту на зображенні.

2. Попередня обробка зображень, цей етап включає нормалізацію зображення для вирівнювання освітлення та контрасту, фільтрацію шуму для видалення зайвих артефактів та бінаризацію, що перетворює зображення в чорно-біле для полегшення процесу розпізнавання.

3. Сегментація тексту на етапі зображення розбивається на окремі символи або слова. Виділяються рядки тексту, ідентифікуються окремі символи, що значно полегшує подальший процес розпізнавання.

Методи розпізнавання тексту

1. Оптичне розпізнавання символів (OCR), цей метод використовує машинне навчання та комп'ютерний зір для розпізнавання окремих символів з зображень. OCR є основою більшості сучасних систем розпізнавання тексту.

2. Глибинне навчання, використовується для виділення ознак з зображень тексту. Рекурентні нейронні мережі (RNN) обробляють послідовності символів і слів, а архітектури типу Transformer, завдяки механізмам уваги, забезпечують високу ефективність обробки тексту.

3. Трансферне навчання, використання попередньо навчених моделей, таких як Tesseract, дозволяє прискорити процес розпізнавання та поліпшити точність результатів. Це досягається завдяки вже накопиченим знанням і здатності адаптуватися до нових даних.

4. Натуральна обробка мови (NLP), після розпізнавання тексту застосовуються методи натуральної обробки мови для корекції можливих помилок розпізнавання та аналізу контексту. Це допомагає краще зрозуміти і інтерпретувати розпізнаний текст.

5. Постобробка результатів, завершальним етапом є видалення артефактів, усунення неправильних розпізнань та форматування тексту. Це включає відновлення структури тексту, такої як абзаци, заголовки.

Дослідники вважають, що Машинне навчання (ДЕКА) є ще одним компонентом штучного інтелекту (ШІ). У більш широкому сенсі процес навчання відіграє важливу роль у пізнанні людини. Люди мають здатність обробляти значну кількість інформації, застосовуючи абстрактну інформацію, яка допомагає їм зрозуміти нові дані. Враховуючи адаптивність, моделі ОД мають здатність імітувати когнітивні здібності людей.

Концепція машинного навчання стала популярною в 1959 році, коли Артур Семюель, відомий у галузі штучного інтелекту та комп'ютерних ігор, популяризував цей термін. Визначення Семюеля машинного навчання підкреслює його сутність як сферу, в якій комп'ютери можуть здобувати знання та навички без необхідності явного програмування.

Машинне навчання (ML) включає різні методи, які часто використовуються для вирішення різних практичних завдань, використовуючи комп'ютерну систему, яка дозволяє навчитися вирішувати проблеми, а не покладатися на явне програмування. Щоб проілюструвати це, замість того, щоб чітко інструктувати

комп'ютерну систему щодо конкретних слів у запиті, які вказують на наявність запитів клієнтів, система з великою колекцією навчальних шаблонів отримує інформацію про загальні шаблони слів та їх комбінації.Це відповідає класифікації потреб.

У 1959 році в журналі IBM Research & Development Journal була опублікована стаття Артура Семюеля з IBM. У цій статті досліджується використання методів машинного навчання в іграх в шашки та демонструється ідея про те, що комп'ютерні системи можна ефективно запрограмувати, щоб отримати більш високий рівень здатності грати в шашки порівняно з людськими програмістами.

Хоча термін "Машинне навчання" вже використовується, Артур Самуель широко відомий як перша людина, яка представила та описала Машинне навчання в його нинішньому вигляді. Впливова стаття Семюеля в журналі "вивчення машинного навчання за допомогою гри в шашки" свідчить про те, що ми можемо перенести власні здібності до навчання в штучні системи.

Це визначення інформатики, яке дозволяє комп'ютерам здобувати знання та навички без необхідності прямого програмування, все ще широко поширене навіть через майже 60 років у його роботі Артура Самуеля.

Це явно не зазначено у визначенні Артура Самуеля, але ключовою характеристикою машинного навчання є концепція автономного навчання. Це передбачає виявлення закономірностей за допомогою статистичного моделювання та використання даних та емпіричної інформації без необхідності явних інструкцій з програмування. Виконавець Самуель назвав цю здатність навчанням без явного програмування, що означає, що машини створюють рішення без початкового програмування. Це дуже залежить від програмування. Навпаки, Самуель зрозумів, що машині не потрібні прямі команди введення для виконання певних завдань, а натомість потрібні вхідні дані.

Область машинного навчання обертається навколо комп'ютерів, які витягують інформацію про шаблони з існуючих даних і використовують цю інформацію для прогнозування нових даних. Ці завдання прогнозування та класифікації, широко відомі як контрольоване навчання, є важливою частиною

машинного навчання. Термін "контрольоване навчання" використовується, коли прогнозується бажаний результат, наприклад, для визначення того, чи є електронний лист спамом, а також для управління та контролю процесу навчання.

У контексті сучасного комп'ютерного програмування існує декомунікативна різниця між програмістом і машиною, яка відрізняється від традиційних методів комп'ютерного програмування. Цей поділ відбувається через здатність машини приймати рішення на основі власного досвіду та моделювати процес прийняття рішень людиною. Щоб проілюструвати це, машина YouTube, яка аналізує переваги для вчених, які вивчають дані, не помітила суттєвої декомунізації між науковцями, які вивчають дані, та їх інтересом до відео про котів. Крім того, машина також визначає фізичні характеристики бейсболіста і показує ймовірність отримання нагороди найціннішого гравця (MVP) за даний сезон.

У першому випадку машина перевіряла улюблені відео дослідника даних на YouTube, беручи до уваги такі показники взаємодії з користувачем, як лайки, підписки і повторювані перегляди. У наступних випадках машина оцінювала фізичні характеристики колишнього найціннішого бейсболіста (Mvp), а також інші фактори, такі як вік та освіта. Однак у цих сценаріях машина явно не запрограмована на отримання певних результатів. Натомість він створив спеціальний алгоритм, який забезпечує вхідні дані, а остаточне прогнозування визначається самою машиною шляхом самонавчання та моделювання даних.

Процес створення моделі даних можна порівняти з навчанням собак-поводирів. Завдяки спеціальній дресируванню собаки-поводирі набувають здатність правильно реагувати в різних ситуаціях. Їх вчать повертати на червоне світло, долати перешкоди і забезпечувати безпеку своїх власників. Поступово, оскільки собаки добре навчені, собак-поводирів можна навчати самостійно в різних сценаріях без нагляду, а моделі машинного навчання можуть навчити їх приймати рішення на основі минулого досвіду, тим самим усуваючи необхідність у присутності дресирувальників.

Приклади включають розробку моделі виявлення спаму як приклад. Модель навчена забороняти отримувати електронні листи по всьому тілу з підозрілими

історіями і 3 або більше ідентифікованими ключовими словами: дорогі друзі, безкоштовно, виставлення рахунків, PayPal, Віагра, казино, Оплата, банкрутство, переможець. Однак важливо зазначити, що нинішній етап не включає впровадження машинного навчання. Повертаючись до візуального представлення команд введення і вхідних даних, стає зрозуміло, що цей процес складається всього з 2 етапів.За командою слідує дія.

Процес декомунікаційної машини складається з 3 послідовних етапів: збір даних, побудова моделі і побудова дій. Щоб інтегрувати Машинне навчання в систему виявлення спаму, вам потрібно замінити термін "команда" на "дані" та ввести поняття "модель" для отримання результатів. У цьому конкретному сценарії дані стосуються колекції зразків електронних листів, але модель базується на статистичних правилах. Параметри моделі охоплюють ті самі ключові слова, що і перший список непотрібних елементів. Надані дані потім використовуються для навчання та оцінки моделі.

Після введення даних у Модель припущення моделі можуть призвести до неправильних прогнозів. Приклад цього можна побачити, переглянувши рядок теми "PayPal отримав платіж за Casino Royale, придбаний на ebay" в електронному листі, який автоматично класифікується як спам відповідно до інструкцій моделі.

Оскільки цей електронний лист надійшов із справжньою автовідповідачем PayPal, він особливо вразливий до таких сценаріїв через шахрайські системи виявлення спаму, наявність негативних слів, перелічених у шаблоні, та відсутність властивої традиційним програмам здатності оцінювати припущення та коригувати Типові правила. І навпаки, Машинне навчання може адаптувати та коригувати припущення, використовуючи 3-ступінчасту обробку та реагування на помилки.

Жовтень дек. 10/12 спочатку різниця між машинним навчанням та традиційним програмуванням може здатися незначною, але ми розглянемо додаткові приклади та розглянемо більш складні сценарії. Витоки машинного навчання, видобутку даних, комп'ютерного програмування та інших суміжних дисциплін (за винятком класичної статистики) можна простежити до інформатики, галузі, що охоплює всі аспекти проектування та використання комп'ютерів (рис.

1.4). Існує більш конкретна область, яка називається наукою про дані, і існують способи та системи використання платформ обробки даних для отримання цінної інформації та розуміння даних.

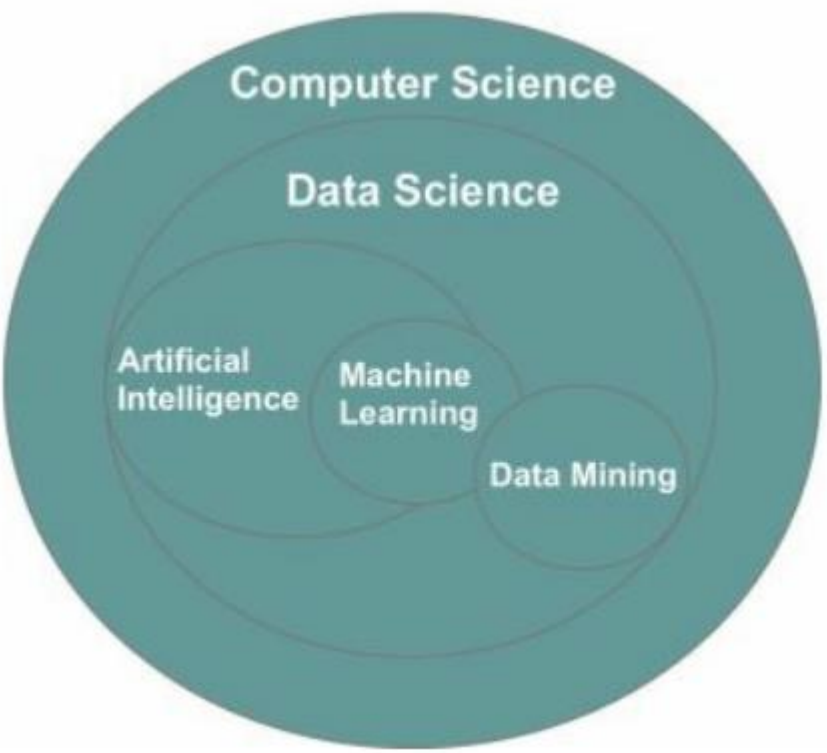


Рис. 1.4 – Сфери які взаємопов’язані з обробкою даних

Штучний інтелект, який часто називають штучним інтелектом, відноситься до здатності машин виконувати інтелектуальну та когнітивну діяльність. Подібно до того, як промислова революція призвела до створення машин, здатних імітувати фізичну роботу, штучний інтелект підтримує розробку машин, здатних відтворювати когнітивні здібності.

Штучний інтелект охоплює набагато більш конкретні підсектори, ніж Інформатика та наука про дані, які зараз користуються великим попитом. Декомунізація включає пошук та планування, логічні міркування та представлення інформації, сприйняття, обробку природної мови (НЛП) та машинне навчання. Зокрема, Машинне навчання узгоджується з іншими сферами штучного інтелекту, такими як НЛП та сприйняття, оскільки воно використовує популярні алгоритми Самонавчання.

Крім жовтня декомунікаційні алгоритми, що використовуються на машинах, мають потенціал для міждисциплінарних додатків, що охоплюють такі області, як сприйняття і обробка природної мови, оскільки Машинне навчання забезпечує більш цілеспрямовану і життєздатну дослідницьку основу в порівнянні з концептуальною двозначністю, властивою штучному інтелекту. Крім того, хоча ступінь магістра може бути достатньою для певної міри здібностей до машинного навчання, докторська ступінь необхідна для досягнення реального прогресу в галузі штучного інтелекту.

Як згадувалося раніше, Машинне навчання та видобуток даних - це тісно пов'язані сфери, які мають спільні методи та цілі. Видобуток даних, спрямований на виявлення закономірностей у великих наборах даних, узгоджується з машинним навчанням. В обох областях для аналізу даних використовуються популярні алгоритми, такі як кластеризація k-середнього, аналіз взаємозв'язків та регресійний аналіз. Однак Машинне навчання зосереджується на ітераційному процесі самонавчання та моделювання даних для прогнозування майбутніх подій, тоді як видобуток даних зосереджується насамперед на минулих подіях.

Важливо підкреслити, що ефективність сучасного машинного навчання багато в чому залежить від наявності високоякісних даних для вивчення алгоритмів, які часто називають "навчальними даними"."Ці дані навчання повинні бути належним чином позначені. Наприклад, щоб класифікувати спам, вам потрібна значна кількість електронних листів, позначених як спам або не позначених як спам. Так само в області діагностичної візуалізації потрібні тисячі зображень з відповідними характеристиками.

Як правило, виділяють наступні методи машинного навчання: неконтрольоване, контрольоване, навчання з підкріпленням. Неконтрольоване Машинне навчання включає підхід, який ідентифікує невизначені шаблони в наборі даних. Таким чином, завдання неконтрольованого навчання не мають абсолютного "правильного" рішення, оскільки немає кінцевої точки відліку.

Контрольоване Машинне навчання (ML) відноситься до того, як ви можете дізнатися про конкретне завдання на основі набору прикладів, що представляють

минулий досвід. У процесі навчання модель здатна до автономного навчання, тому немає необхідності вручну коригувати або програмувати правила або стратегії для вирішення проблем. Точніше, метод відстеження ML застосовує алгоритм до відомого набору даних, щоб отримати уявлення про невідомий набір даних. Цим відомим точкам даних присвоюються семантичні мітки, які служать цілями для моделі ML. Крім того, жовтневі тренування тривалістю від 10 до 9 місяців об'єднують елементи як контрольованих, так і неконтрольованих підрозділів, використовуючи як відмічені, так і непомітні дані.

Навчання підкріплення відноситься до підходу, спрямованого на те, щоб навчити розумних представників вибирати поведінку, яка максимізує загальну винагороду. На відміну від контрольованого навчання, навчання підкріплення не повинно забезпечувати суворо Уніфіковані функції та цілі навчання. Натомість модель здобуває знання за допомогою безперервного навчання за рахунок винагород та покарань. Справжня декомунізація-це пошук балансу між дослідженням невідомого середовища та використанням наявних знань.

Навчання підкріплення може бути складною концепцією, яку можна ефективно зрозуміти, використовуючи аналогії відеоігор. Просуваючись по віртуальному середовищі гри, гравці дізнаються про ефективність різних дій в різних ситуаціях і вивчають динаміку гри.Ці набуті знання визначають і формують майбутню поведінку гравця і миттєво підвищують продуктивність на основі тренувань і попереднього досвіду.

Навчання підкріплення передбачає використання алгоритмів для полегшення постійного вивчення моделі. Звичайна модель навчання з підкріпленням включає кількісні показники ефективності, які не позначені результатами, а натомість оцінюються. Наприклад, успіх у запобіганні зіткнень в контексті автономного транспортного засобу отримує позитивний зворотний зв'язок, а уникнення поразки в шаховій партії також винагороджується позитивним зворотним зв'язком.

У контексті неконтрольованого навчання не всі змінні та шаблони даних підлягають класифікації. Натомість машина повинна виявляти приховані шаблони та встановлювати мітки за допомогою алгоритмів неконтрольованого навчання.

Алгоритм кластеризації K-means є яскравим прикладом неконтрольованого навчання. Цей простий алгоритм групує точки даних на основі порівнянних характеристик.

Однією з переваг неконтрольованого навчання є те, що воно може виявляти раніше невідомі закономірності в даних, наприклад, для ідентифікації окремих сегментів клієнтів. Використання методів кластеризації, таких як кластеризація з K-середнім значенням, може полегшити подальший аналіз після ідентифікації дискретних груп.

1.3 Поточні тенденції та напрями розвитку в області розпізнавання тексту з використанням штучного інтелекту

Розпізнавання рукописних цифр і символів стає все більш важливим в сучасну епоху оцифрування для практичного застосування в різних повсякденних справах. Про це свідчить поява в останні роки різних систем розпізнавання, розроблених або запропонованих для використання в різних галузях промисловості, де важлива висока ефективність класифікації. Ці системи, здатні розпізнавати рукописні літери, символи та цифри, дозволяють людям вирішувати більш складні завдання, які в іншому випадку були б трудомісткими та дорогими. Наприклад, використання автоматизованих систем обробки в банках для обробки чеків є прикладом переваг таких технологій. Без цих автоматизованих систем банкам довелося б наймати занадто багато працівників, ефективність яких була б незрівнянна з комп'ютеризованими системами обробки.

Почерк є серйозною проблемою для людей, які володіють навичками грамотності. Він включає складні сенсомоторні механізми управління, включаючи різні мозкові процеси, пов'язані з емоціями, логічними міркуваннями та спілкуванням. відстань. відстань. Дек. дек. Таким чином, дослідження почерку охоплює дуже широку сферу та сприяє міждисциплінарній співпраці та взаємодії між дослідниками з різними знаннями та інтересами, що дозволяє кожному працювати в різних сферах.

Щоб отримати повне уявлення про розпізнавання рукописного вводу, важливо вивчити захоплюючу область того, як люди сприймають і розуміють почерк. Нюанси систем розпізнавання рукописного вводу залежать від чудових можливостей біологічних нейронних мереж, виявлених у людей і тварин. Ці мережі дозволяють розуміти і вирішувати складні нелінійні з'єднання. Використовуючи штучні нейронні мережі, дослідники можуть створити систему розпізнавання рукописного вводу, яка імітує складність нашого власного мозку. Наш людський мозок має особливу здатність ідентифікувати та розрізняти різні елементи почерку, такі як цифри, букви та символи. Але цей процес не безмежний. Люди мають вроджену схильність до декомунізації, що призводить до різних інтерпретацій почерку серед людей. Навпаки, комп'ютеризовані системи позбавлені таких упереджень і можуть ефективно вирішувати дуже складні завдання, які вимагають від людей значного часу та енергії.

Коли люди беруть участь у процесі сприйняття почерку, візуальна система активується, коли вони взаємодіють із символами, літерами, словами та цифрами. Незважаючи на уявну простоту, процес читання рукописного тексту насправді досить складний. Людський мозок використовує свої великі знання, отримані за роки підсвідомого навчання, для інтерпретації візуальних стимулів. Однак розробка комп'ютерних систем, здатних читати рукописний ввід, спричинила серйозні проблеми через складний характер візуального розпізнавання зображень. Декомунікативні нейронні мережі вважаються найефективнішим способом створення систем розпізнавання рукописного вводу з багатьох використовуваних підходів.

Розпізнавання рукописного тексту є важливою сферою досліджень комп'ютерного зору, штучного інтелекту та розпізнавання образів. Можна стверджувати, що комп'ютерні програми, що виконують розпізнавання рукописного вводу, можуть ідентифікувати та витягувати символи з різних джерел, таких як зображення, фізичні документи та інші носії, та перетворювати їх у електронні або машиночитані формати.

Система може витягувати рукописний вміст із фізичних документів за допомогою технології оптичного сканування або інтелектуального розпізнавання слів. Крім того, система може бути запрограмована на виявлення руху наконечника на екрані і розпізнавання вхідних символів. Насправді розпізнавання рукописного тексту передбачає здатність системи виявляти та інтерпретувати рух наконечника на екрані.

Розпізнавання рукописного вводу можна розділити на 2 різні категорії: автономне розпізнавання та онлайн-розпізнавання. Автономне розпізнавання передбачає вилучення текстового вмісту або символів із зображень та перетворення їх у машиночитані коди символів. Процес передбачає отримання цифрових даних із статичного подання рукописного тексту за допомогою системи, оснащеної рукописними документами. З іншого боку, онлайн-розпізнавання рукописного тексту передбачає виявлення та перетворення символів, введених на спеціальному екрані, у режимі реального часу. У цьому контексті система використовує жести чіпів для ідентифікації та інтерпретації символів та слів.

Спочатку технологія розпізнавання символів була в першу чергу зосереджена на визначенні шаблонів, базових ліній, форм і, зокрема, подібності між геометричними об'єктами на деці. Ці методи включали виявлення інсультів та вивчення їх варіацій. Однак цих методологій часто недостатньо для точного розпізнавання рукописного вмісту у формах та документах, оскільки їм не вистачає необхідної складності та складності.

У 1970-х і 1980-х роках комітети зі стандартів у Сполучених Штатах, Канаді, Японії та деяких європейських країнах розробили різні стандартизовані формати друку на долоні, щоб люди могли друкувати в певних місцях. Таким чином, символи, написані в цих заздалегідь визначених формах, могли бути легше ідентифіковані машинами розпізнавання символів з мінімальними стилістичними змінами, особливо коли дані вводилися групою контрольованих осіб, таких як співробітники організації, відповідно до запропонованої схеми.

Висновок до розділу 1

Розглянуто основні концепції штучного інтелекту (ШІ) та його значення у розпізнаванні тексту. ШІ відіграє ключову роль у сучасних системах обробки тексту, завдяки своїй здатності аналізувати та обробляти великі обсяги даних. Застосування ШІ дозволяє досягти високої точності та ефективності у розпізнаванні тексту, що є важливим для багатьох галузей, таких як оцифровка документів, автоматизація роботи з текстами, створення систем перекладу та інше. Використання нейронних мереж, алгоритмів машинного навчання та інших технологій ШІ значно підвищує можливості розпізнавання тексту, роблячи цей процес більш надійним і точним.

Розглянуто принципи та методи, які використовуються для розпізнавання тексту за допомогою ШІ. Було проаналізовано такі методи, як оптичне розпізнавання символів (OCR), глибинне навчання, трансферне навчання та натуральна обробка мови (NLP). Крім того, розглянуто етапи збирання та підготовки даних, попередньої обробки зображень, сегментації тексту та постобробки результатів. Поєднання цих методів дозволяє створювати високоефективні системи розпізнавання тексту, які здатні працювати з різними типами текстових даних та забезпечувати високу точність результатів.

Розглянуто поточні тенденції та напрями розвитку в області розпізнавання тексту з використанням ШІ. Однією з ключових тенденцій є зростання використання глибинного навчання та архітектур Transformer, що дозволяє досягти високих результатів у розпізнаванні тексту. Іншою важливою тенденцією є інтеграція технологій розпізнавання тексту в різні прикладні системи, такі як мобільні додатки, системи автоматизації бізнес-процесів та інше.

2 АНАЛІЗ ТА ВПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ: МЕТОДИ ТА ПРИЙОМИ

Аналіз та впізнавання рукописного тексту є однією з найбільш складних і водночас важливих задач у галузі обробки інформації. На відміну від друкованого тексту, рукописний текст має значну варіативність у формі літер, стилі написання та якості зображення, що створює додаткові труднощі для автоматизованих систем розпізнавання. У цьому контексті використання методів і прийомів штучного інтелекту, зокрема, нейронних мереж та глибинного навчання, відкриває нові можливості для покращення точності та ефективності впізнавання рукописного тексту. Ця робота присвячена аналізу сучасних методів та прийомів, що використовуються для розпізнавання рукописного тексту, а також обговоренню їхніх переваг, недоліків та перспектив розвитку.

2.1 Дослідження стилів почерку: Виявлення унікальних характеристик рукопису особи

Процес створення рукописного тексту за допомогою письмових інструментів, таких як ручки та олівці, часто називають рукописним введенням. Спочатку ця практика розглядалася як засіб поліпшення людської пам'яті та полегшення спілкування, оскільки інформація передавалася переважно усно з покоління в покоління. Використання символів і застосування лінгвістичних принципів, що регулюють їх розміщення, призводять до створення письмових символів, які сприяють ефективному спілкуванню і в яких вони сприймаються як слова.

Усні традиції, історія, звичаї, Ритуальні процедури тощо До появи письмової мови. Культура служила основним засобом опису культурних аспектів, включаючи культуру та культуру. culture.culture.In наступні покоління. З розвитком і диверсифікацією людської цивілізації стала очевидною необхідність створення

єдиного методу спілкування. Ця нагальна потреба була задоволена використанням смайлів, які представляли собою вдосконалену ітерацію примітивних малюнків і, таким чином, були піонерами в написанні.

Початковою системою письмового спілкування був шумерський піктографічний метод із використанням глиняних табличок. Пізніше цей підхід був вдосконалений і перетворений на більш складну систему під назвою клинопис, яка розпочалася близько 3200 р.До н. е. фінікійці 11. Вони встановили найдавнішу систему алфавіту в першому столітті, але голосних, що складалися лише з 22 букв алфавіту, не було. Ця система відіграла важливу роль у формуванні івриту та арамейської писемності.

Застосовність почерку як засобу письмового спілкування залежить від використання зошитів та педагогічних підходів, таких як метод Палмера. На відміну від цифрових пристроїв, постійну досконалість почерку можна пояснити його зручністю та простотою використання, на відміну від клавіатури. Почерк служить відображенням особистості і дотримання встановлених норм і стає цінним інструментом для самооцінки.

Представлене візуальне уявлення (рис. 2.1) описує різні методи правопису, що використовуються в англійській мові. Перші 3 рядки - це стиль письма, відомий як друкований або дискретний почерк, при якому кожна буква складається в певну область або чітко розділяє кожну букву 4. Цей рядок є прикладом того, як писати, який часто називають чистим курсивом або пов'язаним почерком, і більшість людей пишуть 5 рядків. Буквально так, щоб всі малі літери були з'єднані разом. Варто відзначити, що в ньому використовується змішаний шрифт, що поєднує елементи стилю друку і почерку, а також Символи, доступні для очей.

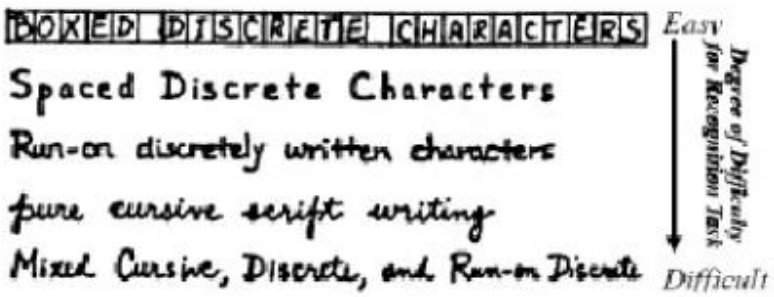


Рис. 2.1 – Стилы написання English

Як друкований почерк, так і почерк важко розпізнати через велику мінливість, властиву онлайн-сигналам почерку. Ця мінливість проявляється як у часі, так і в області сигналів. Часові варіації включають варіації швидкості запису, а просторові варіації сигналу стосуються змін форми окремих символів. Рідко можна знайти 2 однакових письмових знака. Складність розпізнавання рукописного вводу полягає в розробці точних і стабільних моделей, які можуть враховувати мінливість аспектів часової області і області ознак.

жовтні жовтня жовтня, на додаток до тимчасової та просторової мінливості вищевказаних сигналів, рукописний ввід є додатковою формою мінливості, яка ще більше ускладнює завдання. Мінливість декомунізації цієї конкретної форми обумовлена відсутністю чітких меж між символами, що позначають початок або кінець символу. На відміну від друкованого почерку, де ці межі позначені ручкою, в рукописному почерку таких відміток немає. Як результат, розпізнавання рукописного тексту є складним, оскільки розпізнавач повинен явно або опосередковано виконувати операції розділення символів, схильні до помилок.

Графічний аналіз-це вивчення фізичних характеристик і моделей почерку, щоб ідентифікувати автора і зрозуміти його психологічний стан під час написання. Хоча в англійській мові символи представлені символічно за допомогою 8-бітового представлення ASCII, більшість письмових мов можуть бути представлені у 16-бітному форматі Unicode. Ідентифікація почерку означає ідентифікацію автора рукописного тексту з групи авторів з урахуванням унікальності кожного почерку. Перевірка підпису включає визначення того, чи належить конкретний підпис конкретній особі. Ці методи ідентифікації та перевірки широко використовуються в судовій медицині. Розпізнавання та інтерпретація рукописного тексту також можуть використовуватися в повсякденних ситуаціях, наприклад, коли фармацевт розшифровує рецепт. При застосуванні цих методів важливо знати предметну область, щоб виключити варіації і можливості.

Використання нейронних мереж може забезпечити виконання завдання. Нейронна мережа має можливість витягувати інформацію з набору даних, а потім класифікувати невідомі зображення на основі зважених параметрів. Нейронна

мережа-це набір алгоритмів, призначених для розуміння внутрішніх зв'язків у наборі даних, що відображає функції людського мозку. У цьому контексті нейронні мережі взаємодіють з нервовими системами незалежно від того, створені вони органічними чи штучними. Оскільки нейронні мережі мають здатність адаптуватися до змінних вхідних даних, вони можуть генерувати оптимальні результати без зміни початкових умов.

Для підвищення точності та ефективності систем розпізнавання рукописного вводу важливо брати участь у дослідженнях та розробках методів аналізу та ідентифікації рукописного вводу. Відмінною рисою незалежної від набору тексту (W I) системи є чудова здатність розрізняти різні стилі почерку незнайомих користувачів і відрізняти їх від системи, що залежить від набору тексту (W d). На відміну від системи WD, спеціально розробленої для визначення обмеженого числа стилів почерку, система WI враховує значні зміни в почерку різних людей.Люди також мають здатність розпізнавати символи незалежно від речень, але їх здатність розпізнавати власний почерк перевершує здатність розпізнавати почерк незнайомців. Ця різниця у можливостях розпізнавання пов'язана з тим, що системи WD вимагають обмеженої кількості стилів письма, тоді як системи WI вимагають розуміння основних, послідовних та різноманітних характеристик почерку.

Складність завдань розпізнавання рукописного вводу також визначається використовуваним словниковим запасом. Завдання закритого словника передбачає ідентифікацію слів із заздалегідь визначеного фіксованого словника, при цьому розмір словника змінюється без обмежень. Навпаки, завдання з відкритим словником повинні розпізнавати необмежену кількість слів, не обмежуючи обмеження заздалегідь визначеним словником.

Завдання із закритим словником не вважається складним порівняно із завданням із відкритим словником через обмеження, накладені обмеженим словником. Зокрема, завдання закритого словника з невеликим словником особливо просте з кількох причин. По-перше, чим менше у вас словникового запасу, тим менша ймовірність того, що ви зіткнетеся з незрозумілими парами слів. По-друге, чим менший розмір словника, тим більше слів можна моделювати

безпосередньо, але чим більший розмір словника, тим більше символів потрібно моделювати. В основному це пов'язано з обчислювальною складністю прямого моделювання слів. В дек дек по-третє, використання символів для виконання великих завдань словника значно розширить область пошуку і збільшить кількість вузлів в графі пошуку. Зокрема, кількість вузлів збільшується з м на п, де п означає кількість слів, а м означає середню кількість символів у слові (зазвичай від 3 до 10). Крім того, у міру збільшення словникового запасу слова, що не входять до його словникового запасу, з'являються рідше. Таким чином, виконання завдання з великим словниковим запасом можна порівняти з виконанням завдання з використанням відкритого словника.

2.2 Основні принципи розпізнавання тексту: Вивчення та застосування ключових алгоритмів

Існує 2 підходи до оцифрування почерку. Перший метод, який називається "офлайн", - це сканування рукописного або друкованого тексту та перетворення його в цифровий формат. Другий метод, відомий як" Інтернет", передбачає використання стилуса на електронній поверхні, наприклад, дигітайзера, або персонального цифрового помічника з РК-дисплеєм для фізичного набору тексту. При такому підході штрихи пера аналізуються програмним забезпеченням у формі електронного чорнила.

У контексті онлайн-розпізнавання символів безперервні двовимірні координати точок запису зберігаються в певному порядку, і послідовність штрихів може бути легко ідентифікована. Автономне розпізнавання символів, з іншого боку, вимагає аналізу всього написаного сценарію, представленого як зображення, та аналізу двовимірного зображення.

У цих 2 випадках вимоги до зберігання необроблених даних сильно різняться. Автономна система розпізнавання символів обробляє цифрове зображення документа, щоб автоматично інтерпретувати графічну мітку і пов'язати її з відповідним символом. Навпаки, онлайн-система розпізнавання символів працює з

потоком двовимірних координат, що представляють послідовні точки запису, для автоматичного розпізнавання вхідних символів відповідно до траєкторії записуваного пера.

Автономні та онлайн-системи розпізнавання символів мають різні деталі реалізації. Автономні системи розпізнавання в основному використовуються для таких завдань, як обробка банківських чеків, сортування пошти та читання комерційних форм. З іншого боку, системи розпізнавання в Інтернеті в основному використовуються в галузі ручних обчислень та безпеки, особливо для таких завдань, як перевірка підпису та перевірка автора.

Швидкість розпізнавання рукописного вводу в Інтернеті значно вище, ніж в автономному режимі. Позбутися рукописних слів на папері непросто, але розпізнавання рукописного тексту в режимі офлайн складніше через відсутність інформації про час, наприклад, про шум та порядок введення, на відміну від онлайн-рукописного тексту, який може фіксувати додаткову інформацію, таку як напрямок, швидкість та порядок протягом 10 місяців. жовтень. І швидкість в процесі отримання зображень.

2.3 Автоматичний аналіз тексту в реальному часі

Онлайн-розпізнавання рукописного тексту є важливим через кілька ключових аспектів, що мають значний вплив на різні сфери діяльності.

По-перше, це підвищує ефективність роботи. Онлайн-розпізнавання дозволяє швидко перетворювати рукописні записи в цифровий формат, що значно скорочує час на введення даних. Це особливо важливо для бізнесу, освіти та медицини, де велика кількість інформації зберігається в рукописному вигляді. Швидке і точне розпізнавання дозволяє зосередитися на аналізі даних та прийнятті рішень, а не на рутинних завданнях.

По-друге, зручність та мобільність є ще однією значущою перевагою. Можливість розпізнавання рукописного тексту на мобільних пристроях дозволяє користувачам вводити дані будь-де і будь-коли. Це зручно для студентів, лікарів,

бізнесменів та інших професіоналів, які часто працюють поза офісом. Наприклад, лікар може швидко записувати медичні дані пацієнта під час огляду, а бізнесмен може занотовувати ідеї під час зустрічі або подорожі.

Третім важливим аспектом є покращення доступності. Онлайн-розпізнавання рукописного тексту сприяє доступності інформації для людей з обмеженими можливостями, наприклад, для слабозорих або людей з порушеннями моторики, які можуть використовувати голосовий ввід або стилус для написання. Це дозволяє таким користувачам інтегруватися в цифрове середовище і ефективно взаємодіяти з ним.

Четвертим аспектом є автоматизація бізнес-процесів. Автоматичне розпізнавання рукописного тексту допомагає оптимізувати робочі процеси, зменшуючи потребу в ручній обробці документів. Це призводить до зниження витрат та підвищення продуктивності. Бізнес може значно зекономити час і ресурси, автоматизуючи обробку таких документів, як форми, анкети, угоди та інші.

П'ятим важливим аспектом є збереження культурної спадщини. Онлайн-розпізнавання рукописного тексту є важливим інструментом для оцифровки історичних документів, рукописів та інших артефактів. Це дозволяє зберегти і зробити доступними для дослідників та громадськості цінні історичні дані. Такі проекти допомагають зберегти історичну пам'ять та культурні надбання для майбутніх поколінь.

Нарешті, підвищення точності та надійності є значущим аспектом. Сучасні алгоритми розпізнавання рукописного тексту, що використовують штучний інтелект та машинне навчання, досягають високої точності. Це забезпечує надійність даних та мінімізує кількість помилок, які можуть виникати при ручному введенні. Таким чином, розпізнавання рукописного тексту стає не лише швидким, але й максимально точним, що є критично важливим у багатьох сферах, від медицини до юридичної документації.

Традиційно він складається з декількох компонентів, включаючи онлайн-розпізнавач, препроцесор, класифікатор, який присвоює ймовірності різним

категоріям символів або токенів, і постпроцесор динамічного програмування, який часто реалізується у вигляді прихованої Марківської моделі. Крім того, система зазвичай включає мовну модель. жовтень. 10 січня. Під час навчання параметри системи коригуються та коригуються.

Перед лікуванням. Методи попередньої обробки зазвичай застосовуються до попередньо розпізнаного рукописного тексту, включаючи шумозаглушення, нормалізацію та сегментацію. Навпаки, вектори ознак засновані на попередньо обробленому почерку і використовуються в поєднанні з шаблонами символів і мов для полегшення розпізнавання.

Наявність шуму в даних часто викликано обмеженнями в процесі оцифрування, непередбачуваними рухами рук і неточностями в маркуванні олівцем. Для зниження рівня шуму використовується певний набір методів:

1. Згладжування - це загальновживаний метод усереднення кожної точки даних за суміжними точками. Ще однією поширеною стратегією є оцінка основного сліду чорнила за допомогою стандартної кривої.

2. Фільтрація - це процес усунення повторюваних точок даних та зменшення загальної кількості точок. Конкретний використовуваний метод фільтрації залежить від використовуваної технології розпізнавання. 1.1. Один із підходів до фільтрації передбачає встановлення мінімальної відстані декомунікації або фіксованої відстані між послідовними точками. Однак для швидкого запису відстань між послідовними точками може перевищувати цей мінімум. У цих випадках ви можете використовувати інтерполяцію (також відому як передискретизація) для отримання рівномірно розташованих точок. Інший метод фільтрації створює додаткові точки в області критичної кривизни.

3. Виправлення диких точок-передбачає заміну або видалення нормальних аномальних точок даних, які зазвичай виникають внаслідок збою обладнання.

4. Процес зачеплення включає відпускання гачка, яке відбувається на початку і в кінці ходу пера і викликано неправильним положенням при опусканні або підйомі пера.

Нормалізація спрямована на зменшення значних змін розміру символу, і для досягнення цієї мети використовується наступний алгоритм:

1. Корекція спотворень-коригує нахил тексту. Цей прийом застосуємо до всіх слів, а не тільки до окремих буквах.

2. Корекція зміщення базової лінії-вирівнює слова по базовій лінії.

3. Нормалізація розміру-змінює розмір символу відповідно до певного стандартного розміру.

4. Нормалізація довжини рядка-відноситься до процесу налаштування кількості точок у рядку для досягнення певного вирівнювання.

Сегментація відноситься до процедурного поділу рукописних вхідних даних на менші узгоджені одиниці. Ступінь сегментації залежить від конкретних характеристик використовуваного алгоритму розпізнавання:

1. Розділення на рівні рядків відноситься до процесу розділення одного текстового обведення на окремі блоки, що полегшує збільшення або зменшення стилуса.

2. Сегментація на рівні символів відноситься до процесу ідентифікації окремих символів у письмовому тексті. Це може бути досягнуто кількома способами. Курсив друкованого тексту використовує евристичний або Статистичний аналіз на основі навчальних даних, але користувач вводить текст у декомунікативний простір або використовує евристичний метод для аналізу відстані між штрихами. В обох випадках зазвичай створюється кілька потенційних сегментів, і вибираються сегменти, що забезпечують найвищу точність розпізнавання.

3. Сегментація слів-може виконуватися як у часі, так і в просторі або з використанням комбінації обох методів. Обмеження за часом для слова встановлюється, коли час між подіями "скидання декомунізації" і "старіння" перевищує вказаний поріг. З іншого боку, просторовий межа слова встановлюється, коли відстань між осьовими лініями послідовних ударів деки перевищує певний поріг.

4. Сегментація на рівні рядків-використовує як часову, так і просторову інформацію для визначення групи контурів, що відображаються в даному рядку. Якщо розпізнавання не вимагає аналізу в реальному часі, зазвичай спочатку виконується розділення рядків, а потім виконується обведення, пов'язане з ідентифікованим рядком, на рівні слова.

5. Розділення на рівні абзацу-коли користувач поєднує графічну інформацію, таку як діаграма або таблиця, з текстом, сторінка спочатку класифікується за типом інформації. Поля, визначені як текст, розбиваються на рядки за допомогою модуля сегментації на рівні рядків.

Висновок до розділу 2

Досліджено стилі почерку та виявлено унікальні характеристики, які дозволяють ідентифікувати особу за її рукописом. Аналіз почерку враховує різні параметри, такі як форма, нахил, розмір символів, інтервали між літерами та словами. Виявлення цих характеристик є критично важливим для розпізнавання рукописного тексту, оскільки кожна людина має свій унікальний стиль написання. Результати дослідження показали, що використання моделей машинного навчання та нейронних мереж значно покращує точність ідентифікації стилів почерку, що є корисним у різних прикладних сферах, включаючи безпеку та автентифікацію документів.

Розглянуті основні принципи розпізнавання тексту та ключові алгоритми, що використовуються у цій галузі. Було проаналізовано алгоритми оптичного розпізнавання символів (OCR), нейронних мереж, зокрема конволюційних та рекурентних нейронних мереж, а також методи глибинного навчання. Дослідження показало, що поєднання цих алгоритмів дозволяє досягти високої точності розпізнавання тексту, навіть у випадках зі складним почерком або низькою якістю зображення. Важливим аспектом є також попередня обробка даних, яка включає зменшення шуму, нормалізацію та сегментацію тексту, що суттєво покращує результати розпізнавання.

Розглянуто методи автоматичного аналізу тексту в реальному часі. Автоматичний аналіз включає процеси розпізнавання тексту, його синтаксичної та семантичної обробки, а також інтеграції з мовними моделями для забезпечення контекстуального розуміння. Важливим аспектом є здатність системи працювати в реальному часі, що дозволяє використовувати її в мобільних додатках, системах введення даних та інших інтерактивних платформах. Дослідження показало, що використання сучасних технологій, таких як штучний інтелект та машинне навчання, дозволяє досягти високої швидкості та точності обробки тексту в реальному часі, що є надзвичайно важливим для багатьох сучасних застосувань.

3 ДОСЛІДЖЕННЯ ОПТИЧНОЇ СИСТЕМИ ОБРОБКИ РУКОПИСНОГО ТЕКСТУ

3.1 Аналіз обраного методу розпізнавання

Ми можемо навчитися робити це так, щоб люди могли читати та розуміти рукописні документи. Машина також може навчитися читати рукописний текст, використовуючи спеціальну технологію, яка називається глибоким навчанням та штучним інтелектом. Ця технологія називається OCR і відноситься до оптичного розпізнавання символів. Це система, яка допомагає машинам перетворювати електронні та графічні документи на прості для розуміння слова.

Розпізнавання тексту-це комп'ютерний спосіб розуміння та читання різних типів тексту. Існує 2 типи розпізнавання тексту: для рукописного введення і для друку. Навіть якщо його пише одна і та ж людина, писати від руки може бути важко, оскільки почерк варіюється від людини до людини. Є також 2 способи розпізнавання тексту: офлайн і онлайн. Автономне Оптичне розпізнавання символів перевіряє вже створений текст, тоді як онлайн-розпізнавання тексту дозволяє читати текст точно так, як він написаний.

Основною метою цього дослідження є вивчення автономного підходу до розпізнавання символів, необхідного для перетворення рукописних конкретних документів, книг та різних інших матеріалів у цифрові формати. Крім того, цей метод значно прискорює вивчення літератури, яка не була виявлена або не може бути зрозуміла автономно. жовтень. (Рис. 3.1)

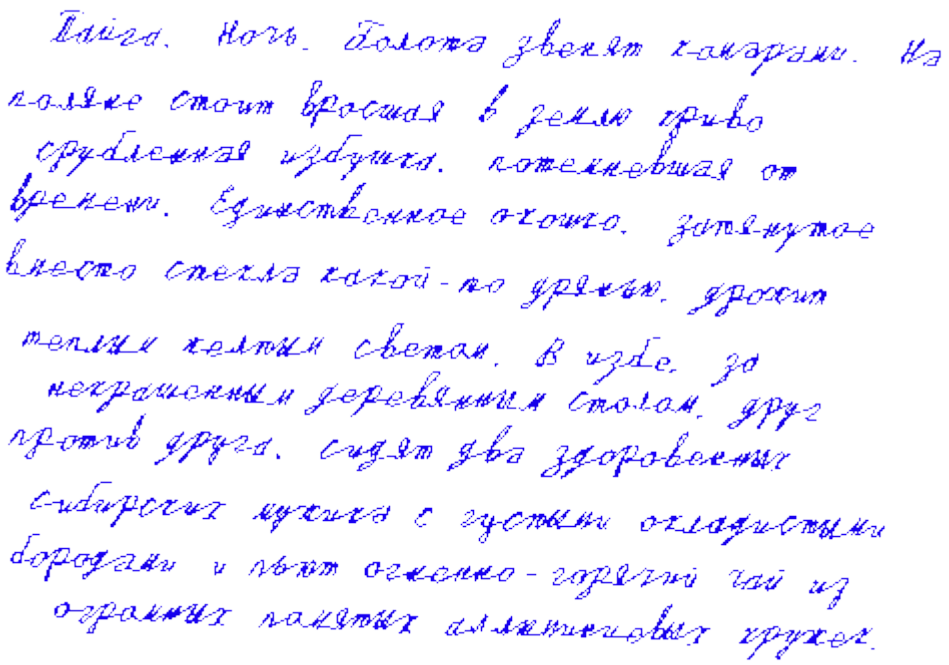


Рис. 3.1 – Приклад рукописного тексту

Згорткові нейронні мережі (CNN) широко використовуються як важливий алгоритм глибокого навчання в галузі класифікації зображень і використовуються в різних програмах комп'ютерного зору.

Однією з ключових переваг використання згорткових нейронних мереж (CNN) є можливість автономного визначення цінних функцій на основі вхідних даних, процес, відомий як відображення функцій, без нагляду людини, на відміну від попередніх моделей. Крім того, CNN зменшує просторовий розмір під час навчання та Жовтня, що прискорює обчислювальну потужність моделі. Нарешті, прихований мережевий рівень забезпечує перетворене значення, яке проходить через функцію активації, щоб правильно класифікувати різні класи.

Згорткові нейронні мережі (CNN) продемонстрували успішне застосування в системах рекомендацій, обробці природної мови та інших областях. Ця модель є високоефективною та ефективною, оскільки вона може автоматично призначати функції та досягати рівня точності, порівнянного з людськими характеристиками.

Платформа TensorFlow, розроблена Google Brain, визнана відмінним інструментом для розробки штучних нейронних мереж (ADN). Він спеціалізується на дослідженнях машинного навчання та штучного інтелекту на глибокому рівні.

Спочатку розроблений для вирішення внутрішніх проблем, з якими стикається Google, TensorFlow буде випущений як програмне забезпечення з відкритим кодом, яке потім розповсюджуватиметься під ліцензією Apache2.0. Таким чином, продукт тепер підтримується спільнотою розробників, які вносять свій внесок у його постійне вдосконалення.

TensorFlow-це універсальна програма, яка полегшує розробку нейронних мереж в операційних системах Windows та UNIX. Ця функція є важливою перевагою фреймворку. tensorflow також сумісний з різними апаратними платформами, такими як ПК та мобільні пристрої, які можуть використовувати різні процесори, такі як процесори, сумісні з платформою, графічні процесори та TPU. Крім того, структура забезпечує спрощений підхід до інтеграції різних функцій штучного інтелекту, таких як обробка мови, комп'ютерний зір, проблеми класифікації та машинний переклад, в існуюче програмне забезпечення. Ця універсальність робить tensorflow придатним для різних завдань. жовтні жовтня. Крім того, TensorFlow включає в себе власну вбудовану систему ведення журналів і відображення, яка дозволяє вам спостерігати за покроковим виконанням обчислень без необхідності використання додаткового програмного забезпечення. Ці функції ілюструють переваги фреймворку.

Коннекціоністи Класифікація часу (CTC) є одним із висновків нейронної мережі, який корисний для вирішення проблем, пов'язаних із порядком, таких як завдання розпізнавання рукописного тексту та мовлення у змінному часі. Використання CTC усуває потребу в попередньо узгоджених наборах даних та спрощує процедури навчання.

Звичайно, одним із можливих підходів до вирішення цієї проблеми є створення набору даних, що містить зображення, що містить текстовий рядок.1 потім ви можете позначити відповідний символ у кожному горизонтальному положенні цих зображень. Навчивши нейронну мережу (N n) цьому набору даних, можна було попросити оцінку кількості символів для кожної горизонтальної позиції. Для цього використовуйте крос-ентропію категорії як функцію втрат.

Однак, незважаючи на життєздатність цього запропонованого рішення, важливо визнати та вирішити деякі пов'язані з цим проблеми.

Анотація набору даних на рівні символів може бути дуже довгим і виснажливим процесом. Це пов'язано з тим, що Ви отримуєте бали лише за окремі символи, і вам знадобиться додаткова обробка, щоб отримати остаточний текст. 1. Однією з проблем цього процесу є видалення повторюваних символів. Це особливо складно, коли один символ може займати більше однієї горизонтальної позиції. Наприклад, впізнаваний текст "Heeloo" може бути отриманий за допомогою широкої літери "e". Однак, якщо розпізнаний текст є "хорошим", видалення всього повторюваного "o" призведе до неправильних результатів. Тоді виникає питання: Як ми впораємося з цією ситуацією?

Варто зазначити, що втрата CTC, яка зазвичай використовується в програмах перетворення мови в текст, також підходить для вирішення цієї проблеми. У цих програмах як аудіосигнал, так і відповідний текст можуть використовуватися в якості навчальних даних. Однак прямого зіставлення між певною частиною аудіосиг деки і окремими символами в тексті немає. Наприклад, ви не можете анотувати відповідну звукову форму кожної літери в хронологічному порядку. Це ще більше підкреслить складність завдання. Таким чином, процес анотування набору даних на рівні символів не тільки трудомісткий і трудомісткий, але й вимагає повторної обробки символів та відповідних голосових сегментів, але такі методи, як втрата CTC, забезпечують ефективне рішення для програм перетворення мови в текст із використанням існуючих навчальних даних. без необхідності жовтневих анотацій точного часу.

Алгоритм враховує символи з найбільшою ймовірністю за кожен часовий інтервал, щоб визначити оптимальний маршрут. грудень 2015 р Алгоритм враховує символи з найбільшою ймовірністю за кожен часовий інтервал для визначення оптимального маршруту. Щоб змінити кодування, алгоритм радикально видаляє повторювані символи та порожні місця. Решта символів представляють певний текст (рис. 3.2).

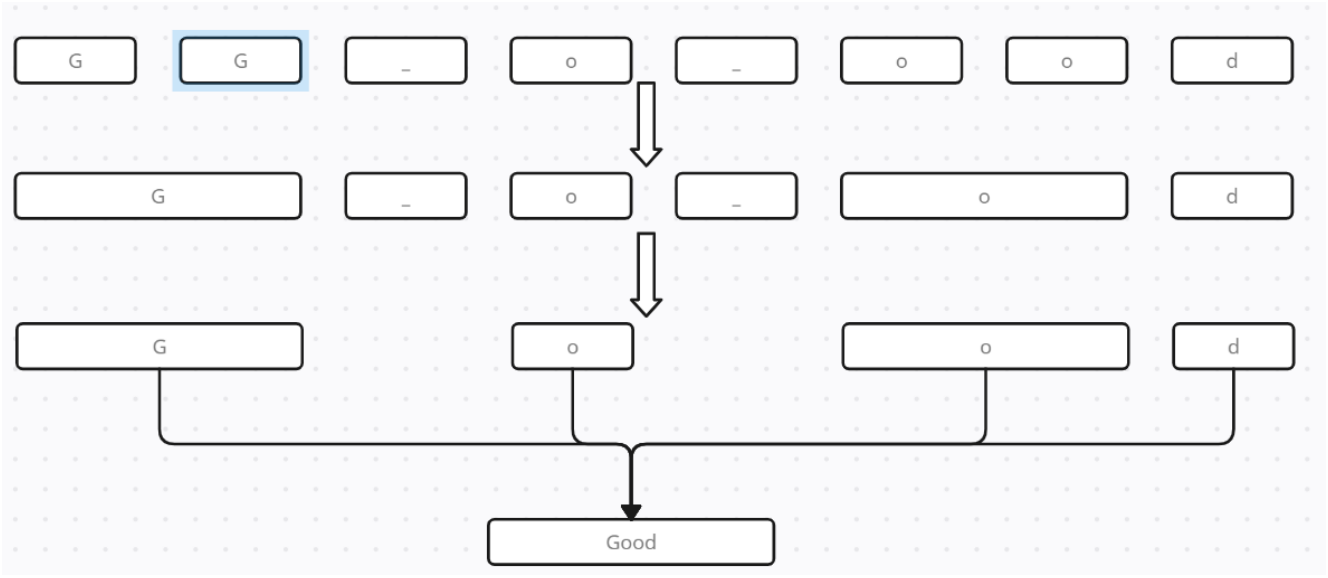


Рис. 3.2 – Декодування слова «Good» використовуючи CTC втрати

Щоб вирішити проблему дублювання символів, ми застосовуємо метод інтелектуального кодування. При кодуванні тексту ви можете вставити будь-яку кількість пробілів в будь-якому місці. Крім того, ви можете повторювати кожного персонажа скільки завгодно разів. Однак важливо розуміти, що в одному прикладі пробіл повинен повторюватися з буквою "Привіт". Ось кілька прикладів:

– Із слова Home: H-o-m-e, H-oo-m-eeee, HHH-o-m-e, тощо
– Із слова Good: G-o-o-d, GG-oo-dd, G-oo-ooo-dd, тощо

Етап перефразування може бути використаний для розширення текстових варіантів, які нейронна мережа може ефективно розпізнавати, тим самим покращуючи загальну продуктивність розпізнавання тексту. Цей метод забезпечує надійну та надійну роботу нейронної мережі в різних сценаріях, дозволяючи обробляти широкий спектр зображень з різною орієнтацією тексту, вирівнюванням та макетом. 1. Однією з переваг цього методу є те, що легко створювати різні макети тексту, кожен із супровідними зображеннями та чіткими макетами. Ця універсальність підвищує загальну гнучкість і адаптивність нейронної мережі. Завдяки навчанню нейронна мережа навчилася створювати зашифрований текст.

Таким чином, ми можемо використовувати здатність нейронної мережі розпізнавати текст на невидимих зображеннях, додаючи перефразування до кращих алгоритмів декодування шляхів, створюючи альтернативне вирівнювання

та варіації вихідного тексту. Такий підхід підвищує адаптивність і точність мережі і робить її цінним інструментом для завдань розпізнавання тексту в різних реальних програмах. Однак важливо зазначити, що перефразування слід робити обережно, щоб не зіпсувати суть тексту та передбачуване повідомлення. Мета полягає в тому, щоб створити альтернативний текстовий вираз, який може бути розпізнаний і зрозумілий нейронною мережею в процесі декодування.

3.2 Підготовка набору даних для аналізу

Ця модель використовує набір даних IAM, який потрібно завантажити вручну. База даних і AMI містить 657 зображень рукописних текстових рядків, написаних групою з 13 353 авторів. Ці тексти взяті з британського англійського корпусу Ланкастер-Осло / Берген і охоплюють загалом 115 320 рукописних сторінок, що складаються з 1539 слів. База даних класифікується як частина сучасної колекції, а анотації додаються на рівні речень, рядків та слів.

Після завантаження вам потрібно підготувати набір даних, створивши повний шлях до кожного зображення з відповідними жовтневими анотаціями. Для цього було реалізовано ітеративний цикл (рис. 3.3).

```
sentences_txt_path = sentences_folder_path
stow.join('Datasets', 'IAM_Sentences', 'ascii', 'sentences.txt') stow.join('Datasets', 'IAM_Sentences',
    'sentences')
=
dataset, vocab, max_len = [], set(), 0
words = open(sentences_txt_path, "r").readlines()
for line in tqdm(words):
if line.startswith("#"):
continue
line_split = line.split(" ") if line_split[2] == "err": continue
folder1
line_split[0][:3]
folder2 line_split[0][:8]
file_name = line_split[0] + ".png" label line_split[-1].rstrip('\n')
=
# recplace '' with in label label = label.replace('', '')
rel_path = stow.join(sentences_folder_path, folder1, folder2, file_name) if not stow.exists(rel_path):
continue
dataset.append([rel_path, label]) vocab.update(list(label))
max_len = max(max_len, len (label))
```

Рисунок 3.3 – Ітераційний цикл

Цей код використовується для підготовки набору даних, призначеного для моделі машинного навчання. Набір даних складається із зображень, що містять рукописний текст, і відповідного тексту в цих зображеннях (рис. 3.4).

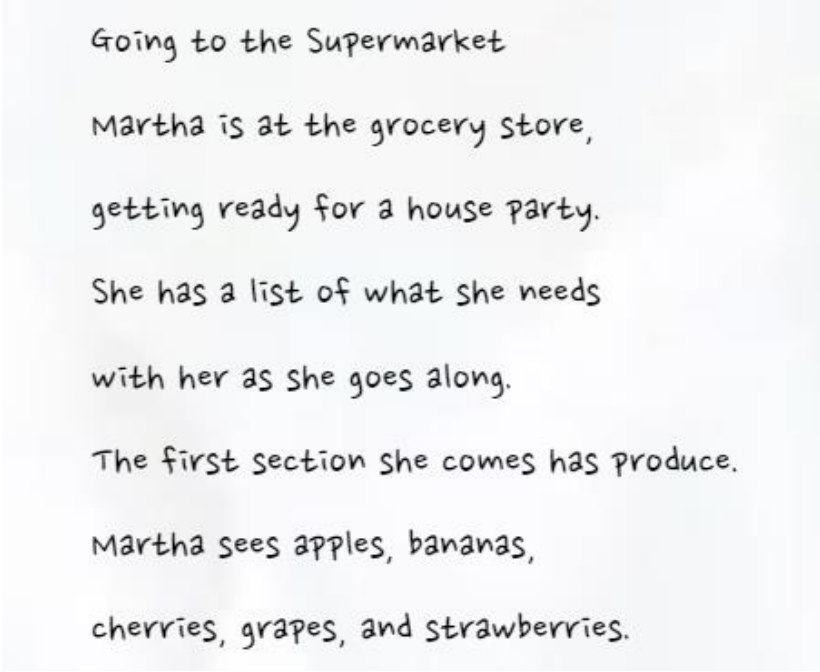


Рис. 3.4 – Приклад тексту взятий із датасету

Код визначає розташування текстового файлу та папки, де містяться зображення. Потім він ініціалізує три порожні змінні: `Dataset`, `vocab`, і `max_len`. Код читає рядки з текстового файлу. Для кожного рядка перевіряється, чи починається він з символу "#", або третє слово у рядку є "err"; якщо так, рядок пропускається. В іншому випадку витягуються перші три та перші вісім літер першого слова, із яких створюється шлях до папки зі зображенням. Ім'я файлу формується, додаючи до першого слова розширення ".png". Текст на зображенні, який є останнім словом у рядку, також вилучається, прибираючи символ нового рядка. Шлях до зображення та відповідна мітка додаються до набору даних, а всі символи мітки додаються до словника. Код також відстежує найдовшу мітку у наборі даних.

Існує велика кількість різноманітних відкритих датасетів, що використовуються для навчання та оцінювання моделей розпізнавання рукописних текстів включаючи:

– Набір даних IAM, отриманий в престижному інституті аналізу документів і розпізнавання образів iam при Базельському університеті, містить добірку рукописних зображень. Німецька англійська, німецька та французька мови цей набір даних охоплює значну кількість 3 сторінок, що охоплюють вміст, написаний 1539 мовами. Дивно, але він охоплює понад 6000 рядків тексту і є безцінним ресурсом для вчених та дослідників машинного навчання та комп'ютерного зору.

– Бентам. набір даних Бентама містить колекцію рукописів Джеремі Бентама, видатного англійського філософа, юриста та соціального реформатора. Цей набір даних, що містить понад 600 рукописів у значній кількості, розділений на чотири окремі розділи: політичні статті, юридичні статті, психологічні статті та різні роботи. Ці рукописи різняться за розміром, від стислих нотаток до повністю переплетених томів. Крім того, кожна стаття в наборі даних містить вступну частину, зміст та детальний опис кожного дослідження. жовтень.

– Рими це велика колекція рукописних текстів французькою мовою, що містить понад 100 мільйонів анотованих слів. Його основна мета-сприяти дослідженню розпізнавання рукописного тексту, що охоплює як окремі літери, так і рядки тексту. Набір даних пропонує різні рівні складності з точки зору розміру символу, технології набору тексту та загальної якості рукописного вводу.

– Вашингтон. Набір даних Вашингтона містить 1800 рукописних зображень англійською мовою, спеціально відібраних для дослідження розпізнавання рукописного введення та машинного навчання. Цей набір даних містить як розділені Символи, так і рядки тексту. Він також розділений на 3 рівні складності: легкий, середній і жорсткий.

– Священна мета. Св. набір даних Галлії, 10. і 6000. Він містить 8 зображень рукописних латинських текстів, що датуються століттями. Цей набір даних має важливе значення для аналізу історичних документів та розпізнавання

рукописного вводу, оскільки він складається з рукописів, що збереглися протягом століть. Він також використовувався для навчання моделей у галузі обробки природної мови.

Наступним кроком є попередня обробка та підготовка набору даних, який буде використовуватися в моделі. Цей процес включає такі дії, як аналіз зображень, перетворення зображень рукописних речень у сумісні формати, сортування тегів, додавання тегів та розбиття наборів даних на навчальні та тестові підмножини. Як і на попередньому кроці, ці дії виконуються за допомогою персоналізованого об'єкта dataProvider (рис. 3.5).

```
# Create a data provider for the dataset
data_provider = DataProvider(
dataset-dataset,
skip_validation=True,
batch_size=configs.batch_size,
data_preprocessors=[ImageReader() ]
transformers=
ImageResizer(configs.width, configs.height, keep_aspect_ratio=True),
Label Indexer (configs.vocab) ,
LabelPadding(max_word_lengt!
]

=configs.max_text_length, padding_value=len(configs.vocab) )

)

# Split the dataset into training and validation sets
train_deta_provider, val_date_provider = data_provider.split(split = 0.9)
```

Рис. 3.5 – Створення об’єкта DataProvider

3.3 Встановлення структури моделі

Після того, як ви підготували набір даних, наступним кроком буде розробка та навчання моделі машинного навчання з використанням тензорного потоку та втрат CTC. Це включає вибір відповідної архітектури моделі та вибір

гіперпараметрів, таких як швидкість навчання та розмір пакета. Ось кілька пропозицій щодо створення моделі.

Визначте вхідний шар моделі як 4-мірний тензор з розміром партії, шириною, висотою та розмірами каналу. Розмір пакета означає кількість зображень у пакеті, а ширина та Висота представляють розміри зображення. Параметр канали вказує кількість кольорових каналів у зображенні зі значенням 1 для зображень у відтінках сірого та 3 для зображень RGB.

Додавання шару CNN: додавання шару CNN до вашої моделі дуже важливо, оскільки воно відіграє важливу роль у витягуванні функцій із зображень. Звичайний підхід до проектування шару CNN включає комбінацію згорткових шарів, сполучних шарів та повністю зв'язаних (щільних) шарів.

Додавання шару Rnn: вам потрібно додати шар RNN до вашої моделі, щоб обробити впорядковану структуру функції та створити символи, що проектуються в тексті. Для цієї мети зазвичай використовуються мережі з довгостроковою короткочасною пам'яттю (LSTM). На малюнку 3.6 показана Архітектура моделі.

Скомпілювати модель шляхом визначення рівнів CNN та RNN з подальшим використанням функції втрат CTC та оптимізатора, такого як Adam.

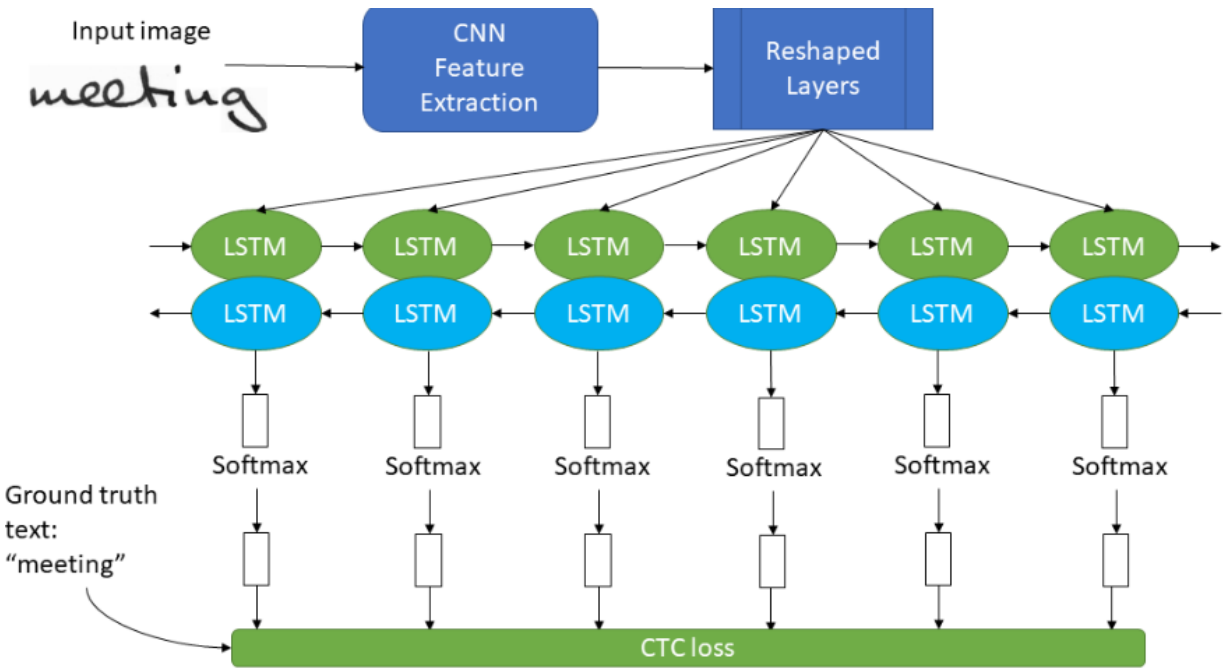


Рис. 3.6 – Архітектура використаної моделі

Показана Архітектура передбачає створення моделі розпізнавання. Як тільки ця модель буде досягнута, ви зможете тренуватися за допомогою тренувального набору. Цей процес можна виконати за допомогою алгоритмів оптимізації, таких як Adam, для підвищення продуктивності моделі. Крім того, жовтневі оптимізатори, функції втрат, метрики та зворотні виклики повинні бути вказані для подальшого вдосконалення процесу навчання (рис. 3.7)

```
model = train_model(
input_dim (configs.height, configs.width, 3),
output_dim= len(configs.vocab),
model.compile(
optimizer=tf.keras.optimizers.Adam(learning_rate=configs.learning_rate),
loss=CTCloss(),
metrics=[
CERMetric (vocabulary-configs.vocab),
WERMetric(vocabulary-configs.vocab)
run_eagerly=False
model.summary(line_length=110)
earlystopper EarlyStopping (monitor="val_CER", patience-20, verbose=1, mode="min")
checkpoint = ModelCheckpoint (f"{configs.model_path}/model.h5", monitor="val_CER", verbose=1,
    save_best_only=True, mode="min") trainLogger
TrainLogger(configs.model_path)
tb_callback TensorBoard (f" (configs.model_path}/logs", update_freq=1)
reduceLROnPlat = ReduceLROnPlateau (monitor="val_CER", factor=0.9, min_delta-le-10, patience-5, verbose
    =1, mode="auto") model2onnx = Model2onnx("{configs.model_path}/model.h5")
model.fit(
train_data_provider,
validation_data=val_data_provider,
epochs-configs.train_epochs,
callbacks=[earlystopper, checkpoint, trainLogger, reduceLROnPlat, tb_callback, model2onnx], workers
    -configs.train_workers
)
```

Рис. 3.7 – Створення архітектури моделі, її компіляція та навчання

У процесі навчання модель генерує оцінки me на основі вхідних даних. Використовуйте функцію декомутації, щоб рекурсивно змінювати модель, щоб мінімізувати розбіжності між прогнозованими значеннями та фактичними мітками.

3.4 Експериментальне тестування та отримання висновків

Після завершення процесу навчання моделі ви можете перевірити його за допомогою нових даних. Щоб правильно оцінити продуктивність моделі, потрібно виконати кілька кроків. По-перше, тестові дані повинні виконувати ті самі етапи попередньої обробки, які застосовуються до навчальних даних. Це гарантує, що тестові дані мають послідовний та послідовний формат для роботи моделі. Після попередньої обробки текстових даних оцініть продуктивність моделі, використовуючи показники втрат CTC, створені на етапі навчання. Крім того, інші вимірювання, визначені під час навчання, можуть бути використані для оцінки ефективності моделі порівняно з жовтневими даними. Ці кроки використовуються для визначення того, наскільки добре навчена модель працює з невидимими даними, що дозволяє проводити подальший аналіз та уточнення за необхідності.

Зображення та результати розпізнавання показані нижче (рис. 3.8-3.10)

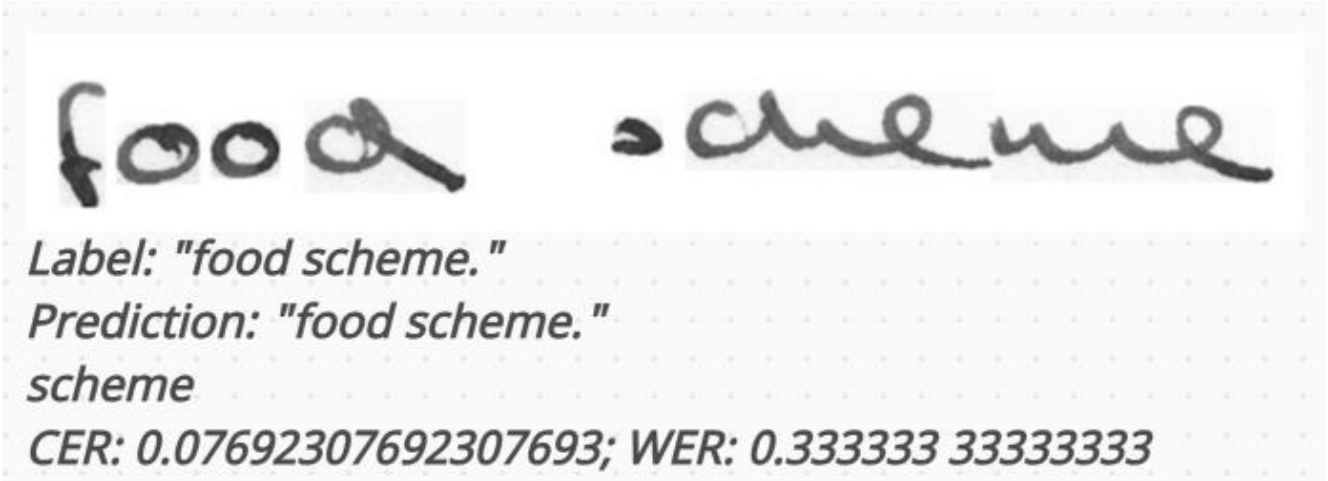


Рис. 3.8 – Результат розпізнавання 1

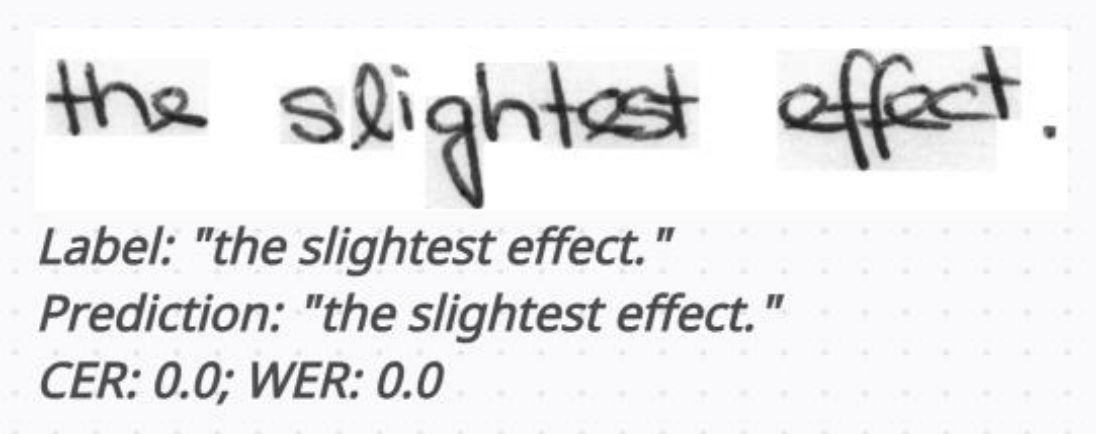


Рис. 3.9 – Результат розпізнавання 2

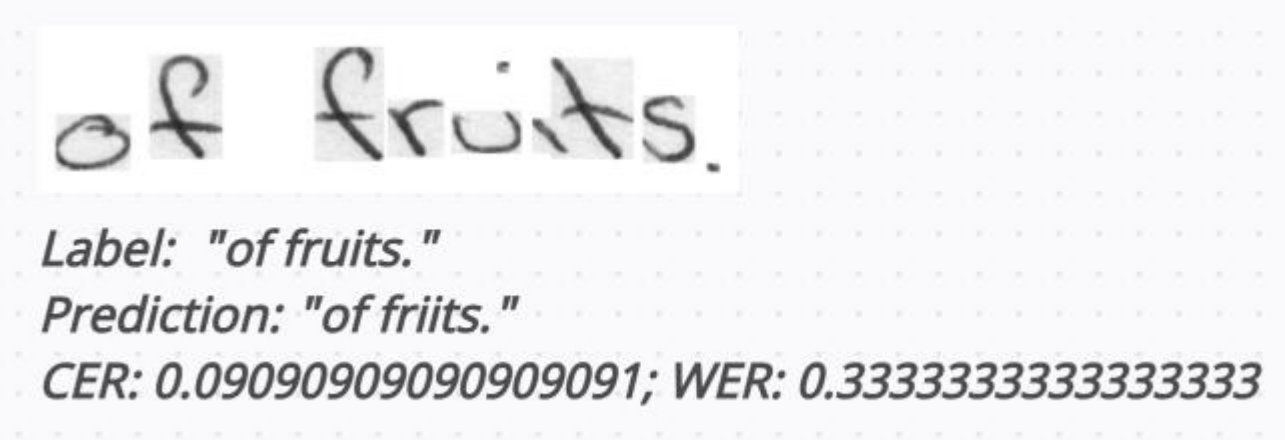


Рис. 3.10 – Результат розпізнавання 3

Наведені вище результати чітко демонструють точність прогнозів нашої моделі. Іноді можуть бути невеликі обмеження, але є багато можливостей для підвищення його ефективності. Потенціал для покращення та покращення продуктивності моделі величезний завдяки вивченню різних будівельних проектів, інтеграції додаткових навчальних даних та використанню різних методів розширення даних.

Висновок до розділу 3

У розділі "Дослідження оптичної системи обробки рукописного тексту" ми перш ніж все зосереджуємося на підготовці набору даних для подальшого аналізу. Це включає в себе збирання та обробку текстових файлів, що містять інформацію про зображення, а також виокремлення тексту, який буде використовуватися для навчання та тестування моделі розпізнавання рукописного тексту.

Після підготовки даних ми переходимо до встановлення структури моделі. Це включає в себе вибір архітектури моделі, визначення вхідних та вихідних шарів, а також налаштування параметрів навчання моделі.

Завершальним етапом є експериментальне тестування та отримання висновків. Під час цього етапу ми навчаємо модель на підготованому наборі даних, тестуємо її на відокремленому тестовому наборі та аналізуємо отримані результати. На основі цих результатів ми можемо робити висновки щодо ефективності моделі та її здатності розпізнавати рукописний текст.

ВИСНОВОК

Розглянуто основні концепції штучного інтелекту (ШІ) та його значення у розпізнаванні тексту. ШІ відіграє ключову роль у сучасних системах обробки тексту, завдяки своїй здатності аналізувати та обробляти великі обсяги даних. Застосування ШІ дозволяє досягти високої точності та ефективності у розпізнаванні тексту, що є важливим для багатьох галузей, таких як оцифровка документів, автоматизація роботи з текстами, створення систем перекладу та інше. Використання нейронних мереж, алгоритмів машинного навчання та інших технологій ШІ значно підвищує можливості розпізнавання тексту, роблячи цей процес більш надійним і точним.

Розглянуто принципи та методи, які використовуються для розпізнавання тексту за допомогою ШІ. Було проаналізовано такі методи, як оптичне розпізнавання символів (OCR), глибинне навчання, трансферне навчання та натуральна обробка мови (NLP). Крім того, розглянуто етапи збирання та підготовки даних, попередньої обробки зображень, сегментації тексту та постобробки результатів. Поєднання цих методів дозволяє створювати високоефективні системи розпізнавання тексту, які здатні працювати з різними типами текстових даних та забезпечувати високу точність результатів.

Досліджено стилі почерку та виявлено унікальні характеристики, які дозволяють ідентифікувати особу за її рукописом. Аналіз почерку враховує різні параметри, такі як форма, нахил, розмір символів, інтервали між літерами та словами. Виявлення цих характеристик є критично важливим для розпізнавання рукописного тексту, оскільки кожна людина має свій унікальний стиль написання. Результати дослідження показали, що використання моделей машинного навчання та нейронних мереж значно покращує точність ідентифікації стилів почерку, що є корисним у різних прикладних сферах, включаючи безпеку та автентифікацію документів.

Розглянуті основні принципи розпізнавання тексту та ключові алгоритми, що використовуються у цій галузі. Було проаналізовано алгоритми оптичного розпізнавання символів (OCR), нейронних мереж, зокрема конволюційних та рекурентних нейронних мереж, а також методи глибинного навчання. Дослідження показало, що поєднання цих алгоритмів дозволяє досягти високої точності розпізнавання тексту, навіть у випадках зі складним почерком або низькою якістю зображення. Важливим аспектом є також попередня обробка даних, яка включає зменшення шуму, нормалізацію та сегментацію тексту, що суттєво покращує результати розпізнавання.

Розглянуто методи автоматичного аналізу тексту в реальному часі. Автоматичний аналіз включає процеси розпізнавання тексту, його синтаксичної та семантичної обробки, а також інтеграції з мовними моделями для забезпечення контекстуального розуміння. Важливим аспектом є здатність системи працювати в реальному часі, що дозволяє використовувати її в мобільних додатках, системах введення даних та інших інтерактивних платформах. Дослідження показало, що використання сучасних технологій, таких як штучний інтелект та машинне навчання, дозволяє досягти високої швидкості та точності обробки тексту в реальному часі, що є надзвичайно важливим для багатьох сучасних застосувань.

У розділі "Дослідження оптичної системи обробки рукописного тексту" ми перш ніж все зосереджуємося на підготовці набору даних для подальшого аналізу. Це включає в себе збирання та обробку текстових файлів, що містять інформацію про зображення, а також виокремлення тексту, який буде використовуватися для навчання та тестування моделі розпізнавання рукописного тексту.

Завершальним етапом є експериментальне тестування та отримання висновків. Під час цього етапу ми навчаємо модель на підготованому наборі даних, тестуємо її на відокремленому тестовому наборі та аналізуємо отримані результати. На основі цих результатів ми можемо робити висновки щодо ефективності моделі та її здатності розпізнавати рукописний текст.

ПЕРЕЛІК ПОСИЛАНЬ

1. Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta, “Handwritten Character Recognition using Neural Network and Tensor Flow”, Computer Science and Engineering, SRM IST Ghaziabad, India, International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, pp 1445

2. Plamondon R.A.: Online and off-string handwriting recognition: a comprehensive survey. Pattern Analysisand Machine Intelligence. vol. 1 (22), 63-84 (2000)

3. Nguyen. Cuong Tuan, Bilan Zhu, and Masaki Nakagawa. "A semi-incremental recognition method foron-line handwritten English text”, 14th InternationalConference on Document Analysis and Recognition, 2014.

4. Nair. Pranav P, Ajay James, and C saravnan. "MalayalamHandwritten Character Recognition Using ConvolutionalNeural Network", International Conference on InventiveCommunication and Computational Technologies (ICICC),2017.

5. H. Scheidl, S. Fiel Wien, and H. Scheidl Robert Sablatnig, “Handwritten Text Recognition in Historical Documents DIPLOMARBEIT zur Erlangung des akademischen Grades Visual Computing eingereicht von,” pp. 90– 96, 2011 – Режим доступу до ресурсу:

6. Park, Jaehwa, Venu Govindaraju, and Sargur N. Srihari."OCR in a hierarchical feature space." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.4 (2000)

7. Bunke, H., Samy Bengio, and A. Vinciarelli. "Offlinerecognition of unconstrained handwritten texts using HMMsand statistical language models." Pattern Analysis and Machine Intelligence, IEEE Transactions on 26.6 (2004)

8. Zhizhen Liang and Pengfei Shi. A metasynthetic approach for segmenting handwritten Chinese characterstrings. In Pattern Recognition Letters, volume 26, pages 1498–1511, 2005

G. Seni E. Cohen "External Word Segmentation ofOff-Line Handwritten Text Lines" Pattern Recognition vol. 27 no. 1.

9. M. Gilloux J.M. Bertille M. Leroux "Recognitionof Handwritten Words in a Limited DynamicVocabulary" Proc. Third Int'l Workshop Frontiers ofHandwriting Recognition (IWFHRIII)

A. Graves, S. Fernandez, M. Liwicki, H. Bunke, and J. Schmidhuber. Unconstrained online ´ handwriting recognition with recurrent neural networks. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, Advances in Neural Information Processing Systems 20. MIT Press, Cambridge, MA

10. Singh Abnikant, Singh Markandey, Ratan Ritwaj, Kumar S. 2005, “Optical Character Recognition for printed Tamil text”, Journal of Zhejiang University, 6A(11)

11. Schantz H. F. History of OCR, Optical Character Recognition / Schantz H. F. // Recognition Technologies Users Association, 1982. – 114 p

12. Lam S. W. An Adaptive Approach to Document Classification and Understanding / S. W. Lam // IAPR Workshop on Document Analysis Systems, Kaiserslautern, Germany. – 1994. – P. 231-251.

13. R. E. Howard, W. Hubbard, L. D. Jackel, H. S. Baird // Proc. of International Conference on Pattern Recognition, Atlantic City. – 1990. – P. 541-551.

14. Machine Learning with Python - from Model to Web-Service by Aleksei Tiulpin. URL: https://youtu.be/9sFUlR-CS7Y. (дата звернення: 02.04.2024)

15. Support Vector Machine Learning MNIST Handwritten Digits. URL: https://youtu.be/s8q_OQBJpwU. (дата звернення: 05.04.2024)

16. K-Nearest Neighbor Classification MNIST Handwritten Digits. URL: https://youtu. be/ooQtUaCExa8. (дата звернення: 10.04.2024)

17. Machine learning in Python with scikit-learn. URL: https://www.youtube. com/playlist?list=PL5- da3qGB5ICeMbQuqbbCOQWcS6OYBr5A. (дата звернення: 12.04.2024)

18. Розпізнавання рукописного тексту при застосуванні нейронної мережі. ДУІКТ А. Гудз. Київ 2023

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ

ДЕРЖАВНИ Й УНІВЕРСИТЕТ ІНФОРМАЦІЙНО - КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО -НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

Кваліфікаційна робота
на тему:

«Дослідженняалгоритміврозпізнаваннярукописнихсимволівв системахтестування»

Виконав: Ушаков Олександр

Науковий керівник: Ільїн Олег

КИЇВ - 2024

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

- *Мета роботи* – аналіз принципи та методи роботи нейронних мереж та машинного навчання з метою розробки ефективної моделі для розпізнавання рукописного тексту.
- *Об'єкт дослідження* – нейронні мережі та методи машинного навчання для розпізнавання рукописного тексту.
- *Предмет дослідження* – принципи роботи нейронних мереж та алгоритми машинного навчання, які використовуються для ефективного розпізнавання рукописного тексту.

Підхід до аналізу тексту з використанням

Штучний інтелект відіграє значну роль у розпізнаванні тексту, що дозволяє автоматизувати та полегшити обробку великих обсягів текстової інформації. Однією з ключових задач є автоматичне розпізнавання тексту зі зображень або сканованих документів за допомогою методів машинного навчання.

Рисунок 1 ілюструє анатомію типового нейрона, показуючи його три окремі домени. Дендрити характеризуються розгалуженими та дерев'янистими структурами та виконують важливу функцію передачі електрохімічних сигналів, отриманих від сусідніх нервових клітин, до тіла клітини.

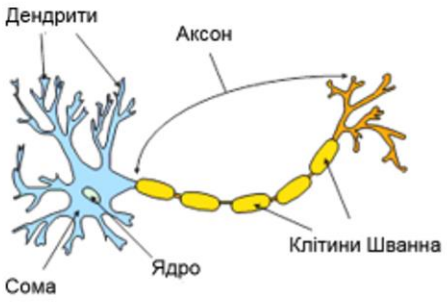


Рис. 1 – Анатомія типового нейрона

Ці канали часто є місцем, де ініціюються потенціали дії. Точніше, коли електрохімічні сигнали досягають тіла клітини, вони консолідуються на аксональному горбку. Якщо сума цих сигналів перевищує певний поріг, генерується потенціал дії, який передається по аксону. На рисунку показано візуальний поріг для спрацьовування потенціалу дії та результуюча вихідна напруга.

На рисунку показано візуальний поріг для спрацьовування потенціалу дії та результуюча вихідна напруга.

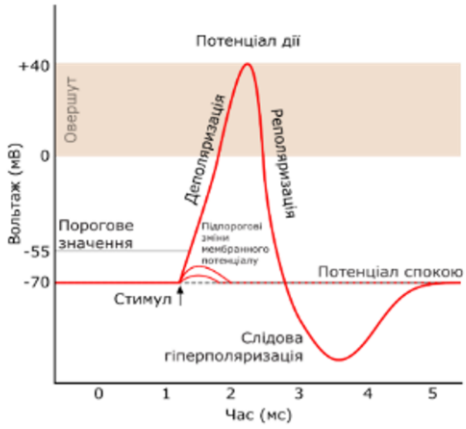


Рис. 2 – Візуальний поріг для спрацьовування потенціалу

На рисунку 3 показана конфігурація одного нейрона, що складається з трьох вхідних вузлів x1, x2 і x3, і їх відповідних ваг w1, w2 і w3. Зверніть увагу, що нижній індекс і опущено, оскільки є вузол, хоча він буде представлений як w11, w12 і w13. Важливо розуміти, що коли мережа стає складнішою, знаки ваг і вузлів також ставатимуть складнішими.

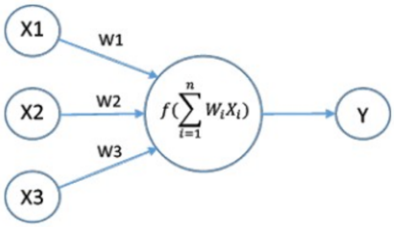


Рис. 3 – Математична конфігурація одного нейрона

Принципи розпізнавання тексту

1. Збирання та підготовка даних, процес розпізнавання тексту починається зі збирання високоякісних зображень тексту. Ці зображення повинні бути чіткими та добре освітленими. Крім того, важливим етапом є анотація зображень, що може виконуватися вручну або автоматично, для точного визначення розташування тексту на зображенні.
2. Попередня обробка зображень, цей етап включає нормалізацію зображення для вирівнювання освітлення та контрасту, фільтрацію шуму для видалення зайвих артефактів та бінаризацію, що перетворює зображення в чорно-біле для полегшення процесу розпізнавання.
3. Сегментація тексту на етапі зображення розбивається на окремі символи або слова. Виділяються рядки тексту, ідентифікуються окремі символи, що значно полегшує подальший процес розпізнавання.

Методи розпізнавання тексту

1. Оптичне розпізнавання символів (OCR), цей метод використовує машинне навчання та комп'ютерний зір для розпізнавання окремих символів з зображень. OCR є основою більшості сучасних систем розпізнавання тексту.
2. Глибинне навчання, використовується для виділення ознак з зображень тексту. Рекурентні нейронні мережі (RNN) обробляють послідовності символів і слів, а архітектури типу Transformer, завдяки механізмам уваги, забезпечують високу ефективність обробки тексту.
3. Трансферне навчання, використання попередньо навчених моделей, таких як Tesseract, дозволяє прискорити процес розпізнавання та поліпшити точність результатів. Це досягається завдяки вже накопиченим знанням і здатності адаптуватися до нових даних.
4. Натуральна обробка мови (NLP), після розпізнавання тексту застосовуються методи натуральної обробки мови для корекції можливих помилок розпізнавання та аналізу контексту. Це допомагає краще зрозуміти і інтерпретувати розпізнаний текст.
5. Постобробка результатів, завершальним етапом є видалення артефактів, усунення неправильних розпізнань та форматування тексту. Це включає відновлення структури тексту, такої як абзаци, заголовки.

Дослідження стилів почерку: Виявлення унікальних характеристик рукопису особи

- Представлене візуальне уявлення (рис. 5) описує різні методи правопису, що використовуються в англійській мові. Перші 3 рядки - це стиль письма, відомий як друкований або дискретний почерк, при якому кожна буква складається в певну область або чітко розділяє кожну букву 4. Цей рядок є прикладом того, як писати, який часто називають чистим курсивом або пов'язаним почерком, і більшість людей пишуть 5 рядків. Буквально так, щоб всі малі літери були з'єднані разом. Варто відзначити , що в ньому використовується змішаний шрифт, що поєднує елементи стилю друку і почерку, а також Символи, доступні для очей.

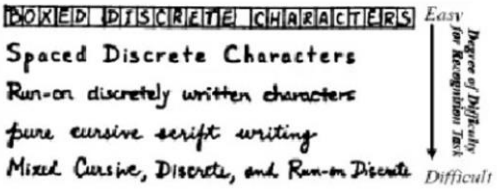


Рис. 5 – Стилы написання English

Аналіз обраного методу розпізнавання

- Основною метою цього дослідження є вивчення автономного підходу до розпізнавання символів, необхідного для перетворення рукописних конкретних документів, книг та різних інших матеріалів у цифрові формати. Крім того, цей метод значно прискорює вивчення літератури, яка не була виявлена або не може бути зрозумілаавтономно. жовтень. (Рис. 6)
- Алгоритмвраховуєсимволи з найбільшою ймовірністю за кожен часовий інтервал, щоб визначити оптимальний маршрут.грудень2015 р Алгоритмвраховуєсимволи з найбільшою ймовірністю за кожен часовий інтервал для визначення оптимального маршрутуЩоб змінити кодування, алгоритм радикальновидаляє повторюванісимволи та порожні місця. Решта символів представляютьпевний текст (рис. 7).

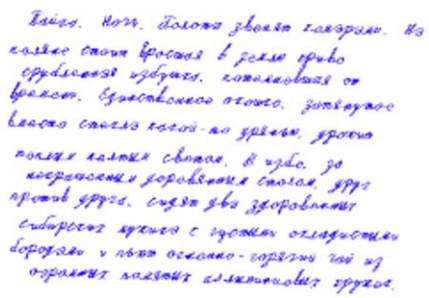


Рис. 6 – Приклад рукописного тексту

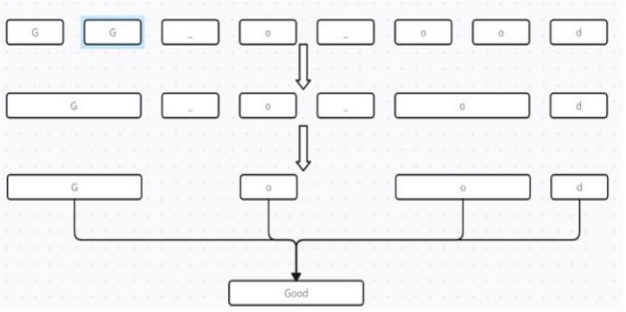


Рис. 7 – Декодування слова «Good» використовуючиCTC втрати

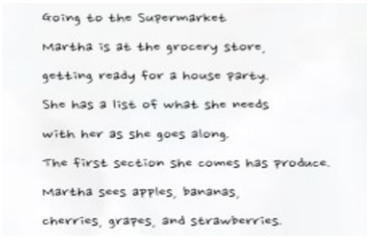
Підготовка набору даних для аналізу

Після завантаження вам потрібно підготувати набір даних, створивши повний шлях до кожного зображення з відповідними жовтневими анотаціями. Для цього було реалізовано ітеративний цикл (рис. 8).

```
sentences_txt_path = sentences_folder_path
stow.join('Datasets', 'IAM_Sentences', 'ascii', 'sentences.txt') stow.join('Datasets', 'IAM_Sentences',
    'sentences')
#
dataset, vocab, max_len = [], set(), 0
words = open(sentences_txt_path, "r").readlines()
for line in tqdm(words):
if line.startswith("#"):
continue
line_split = line.split(" ") if line_split[2] == "err": continue
folder1
line_split[3][:3]
folder2 line_split[3][:8]
file_name = line_split[3] + ".png" label line_split[-1].rstrip('\n')
#
# replace '' with in label label = label.replace('|', ' ')
rel_path = stow.join(sentences_folder_path, folder1, folder2, file_name) if not stow.exists(rel_path):
continue
dataset.append([rel_path, label]) vocab.update(list(label))
max_len = max(max_len, len (label))
```

Рис. 8 – Ітераційний цикл

Цей код використовується для підготовки набору даних, призначеного для моделі машинного навчання. Набір даних складається із зображень, що містять рукописний текст, і відповідного тексту в цих зображеннях (рис. 9).


Going to the Supermarket

Martha is at the grocery store,
getting ready for a house party.
She has a list of what she needs
with her as she goes along.
The first section she comes has produce.
Martha sees apples, bananas,
cherries, grapes, and strawberries.

Рис. 9 – Приклад тексту взятий із датасету

- Наступним кроком є попередня обробка та підготовка набору даних, який буде використовуватися в моделі. Цей процес включає такі дії, як аналіз зображень, перетворення зображень рукописних речень у сумісні формати, сортування тегів, додавання тегів та розбиття наборів даних на навчальні та тестові підмножини . Як і на попередньому кроці, ці дії виконуються за допомогою персоналізованого об'єкта dataProvider (рис. 10).

```
# Create a data provider for the dataset
data_provider = DataProvider(
dataset-dataset,
skip_validation=True,
batch_size=configs.batch_size,
data_preprocessors=[ImageReader() ]
transformers=
ImageResizer(configs.width, configs.height, keep_aspect_ratio=True),
Label Indexer (configs.vocab) ,
LabelPadding(max_word_lengt!
]

=configs.max_text_length, padding_value=len(configs.vocab) )

)

# Split the dataset into training and validation sets
train_deta_provider, val_date_provider = data_provider.split(split = 0.9)
```

Рис. 10 – Створення об’єкта DataProvider

Встановлення структури моделі

- Скомпілювати модель шляхом визначення рівнів CNN та RNN з подальшим використанням функції втрат CTC та оптимізатора, такого як Adam.

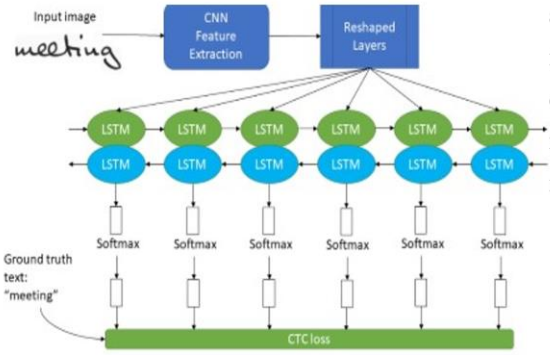


Рис. 11 – Архітектура використаної моделі

Цей процес можна виконати за допомогою алгоритмів оптимізації, таких як Adam, для підвищення продуктивності моделі. Крім того, жовтневі оптимізатори, функції втрат, метрики та зворотні виклики повинні бути вказані для подальшого вдосконалення процесу навчання (рис. 12)

Рис. 12 – Створення архітектури моделі, її компіляція та навчання

Експериментальне тестування та отриманнявисновків

- Зображення та результати розпізнавання показані нижче (рис. 13-15)

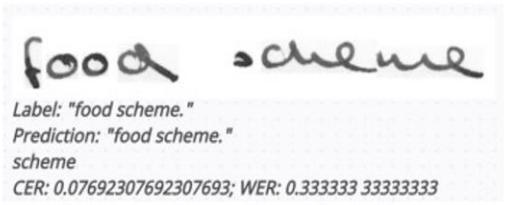


Рис. 13 – Результат розпізнавання 1

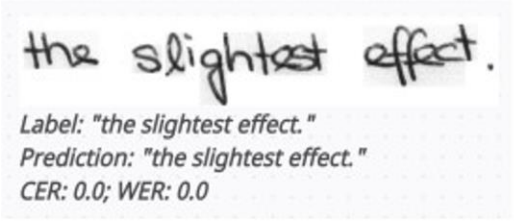


Рис. 14 – Результат розпізнавання 2

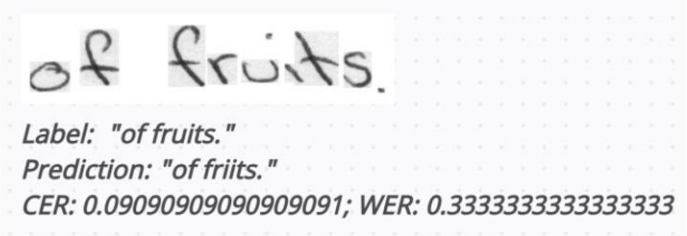


Рис. 15 – Результат розпізнавання 3

Висновок

- Розглянуто основні концепції штучного інтелекту (ШІ) та його значення у розпізнаванні тексту. ШІ відіграє ключову роль у сучасних системах обробки тексту, завдяки своїй здатності аналізувати та обробляти великі обсяги даних.
- Розглянуто принципи та методи, які використовуються для розпізнавання тексту за допомогою ШІ. Було проаналізовано такі методи, як оптичне розпізнавання символів (OCR), глибинне навчання, трансферне навчання та натуральна обробка мови (NLP).
- Розглянуто поточні тенденції та напрями розвитку в області розпізнавання тексту з використанням ШІ. Однією з ключових тенденцій є зростання використання глибинного навчання та архітектур Transformer, що дозволяє досягти високих результатів у розпізнаванні тексту.
- Досліджено стилі почерку та виявлено унікальні характеристики, які дозволяють ідентифікувати особу за її рукописом. Аналіз почерку враховує різні параметри, такі як форма, нахил, розмір символів, інтервали між літерами та словами.
- Розглянуті основні принципи розпізнавання тексту та ключові алгоритми, що використовуються у цій галузі. Було проаналізовано алгоритми оптичного розпізнавання символів (OCR), нейронних мереж, зокрема конволюційних та рекурентних нейронних мереж, а також методи глибинного навчання.