

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО–НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра інженерії програмного забезпечення

Пояснювальна записка

до магістерської роботи
на ступінь вищої освіти магістр

на тему: «**МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ОПТИМІЗАЦІЇ
ІНТЕРФЕЙСІВ ДОДАТКІВ ДЛЯ ЛЮДЕЙ З ОСОБЛИВИМИ ПОТРЕБАМИ**»

Виконав: студент 6 курсу, групи ПДМ–61

спеціальності

121 Інженерія програмного забезпечення

(шифр і назва спеціальності)

Нікітченко А.Ю.

(прізвище та ініціали)

Керівник Щербина І.С.

(прізвище та ініціали)

Рецензент _____

(прізвище та ініціали)

Київ – 2021

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

Кафедра Інженерії програмного забезпечення

Ступінь вищої освіти - «Магістр»

Спеціальність - 121 «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ

Завідувач кафедри

Інженерії програмного забезпечення

О.В. Негоденко

— ___ || ___ 20__ року

**ЗАВДАННЯ
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ**

Нікітченку Андрію Юрійовичу

(прізвище, ім'я, по батькові)

1. Тема роботи: «Методи машинного навчання для оптимізації інтерфейсів додатків для людей з особливими потребами»

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

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

затверджені наказом вищого навчального закладу від «13» го жовтня 2020 року №230.

2. Строк подання студентом роботи 24.12.2020

3. Вихідні дані до роботи:

3.1 Методи машинного навчання;

3.2 ІТ-технології для людей з особливими потребами;

3.3 Технічні вимоги;

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

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

4.1 Порівняльний аналіз результатів, отриманих іншими авторами;

4.2 Методика дослідження;

4.3 Результати дослідження;

4.4 Висновки/

5. Перелік демонстраційного матеріалу (назва основних слайдів)

1. Актуальність

2. Існуюче програмне забезпечення

3. Технологія машинного навчання

4. Методи кластеризації та трансферного навчання

5. Представлення роботи GNMT

6. Інтерфейс додатку для людей з особливими потребами з використанням мережі GNMT

7. Відповідність розробленого інтерфейсу міжнародним стандартам доступності

6. Дата видачі завдання 02.11.2020.

КАЛЕНДАРНИЙ ПЛАН

№ з / п	Назва етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1	Підбір науково-технічної літератури	02.11.20	Виконано
2	Огляд існуючих рішень та літератури	03.11.20	Виконано
3	Огляд розумних будинків на основі IoT	12.11.20	Виконано
4	Розробка системи оцінки безпеки	16.11.20	Виконано
5	Дослідження ризиків безпеки за допомогою розробленої системи	24.11.20	Виконано
6	Вступ, висновки, реферат	08.12.20	Виконано
7	Розробка обов'язкових демонстраційних матеріалів	11.12.20	Виконано
8	Здача роботи	24.12.20	

Студент _____
(підпис) (прізвище та ініціали)

Керівник роботи _____
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Текстова частина магістерської роботи складає 87 с, 20 рис, 25 джерел.

Об'єкт дослідження – дослідження застосування методів машинного навчання при розробці додатку для людей з обмеженими можливостями.

Предмет дослідження – мобільний додаток для людей з особливими потребами.

Мета роботи – дослідити ефективність впровадження технологій машинного навчання при розробці додатку для людей з особливими потребами.

Методи дослідження – системний аналіз проведення порівняльного аналізу, метод узагальнення.

У даній магістерській роботі, мною було проведена певна кількість досліджень інтегрованості людей з особливими потребами і інформаційну діяльність, а також роль ІТ-технологій в житті та побуті таких людей.

Також було проведено дослідження та порівняння систем штучного інтелекту та методів машинного навчання при розробці мобільного додатку для людей з особливими потребами

Було поставлене та вирішене завдання з розробки зручного та зрозумілого Android-додатку онлайн перекладача для людей з особливими потребами.

Задля створення максимальної доступності та зручності в користуванні, були реалізовані такі інформаційні рішення як наявність голосового синтезатору, наявність технології розпізнавання мови(Text-To-Speech), наявність голосових інструкцій і підказок, а також – впровадження методів машинного навчання для розробки зручного UI-інтерфейсу та його відповідності міжнародному стандарту WCAG (Web Content Accessibility Guidelines).

Ключові слова: ІТ-ТЕХНОЛОГІЇ, ШТУЧНИЙ ІНТЕЛЕКТ, APPINVENTOR, МАШТННЕ НАВЧАННЯ, UI-ІНТЕРФЕЙС.

ЗМІСТ

ВСТУП	8
1. ДОСЛІДЖЕННЯ ВПЛИВУ ІТ-ТЕХНОЛОГІЙ ТА СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ НА ЖИТТЯ ЛЮДЕЙ З ОСОБЛИВИМИ ПОТРЕБАМИ	9
1.1. Використання сучасних технологій людьми з обмеженими можливостями	9
1.2. Доступність сучасних ІТ-технологій для людей з особливими потребами ..	13
1.3. Дослідження ринку ІТ-технологій для людей з особливими потребами	19
1.4. Огляд та характеристика мобільних додатків, призначених для людей з обмеженими можливостями	27
1.5. Застосування нейронних мереж при розробці високотехнологічних рішень для людей з обмеженими можливостями	31
2. ДОСЛІДЖЕННЯ СПОСОБІВ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ	38
2.1. Дослідження способів технології машинного навчання	38
2.2. Дослідження методів машинного навчання	53
3. ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ПРИ РОЗРОБЦІ МОБІЛЬНОГО ДОДАТКУ ДЛЯ ЛЮДЕЙ З ОБМЕЖЕНИМИ МОЖЛИВОСТЯМИ	67
3.1. Розробка мобільного додатку - перекладача для людей з особливими потребами	67
3.2. Огляд та застосування нейронної мережі GNMT	77
3.3. Впровадження інтерфейсу, заснованого на технології машинного навчання	90
ВИСНОВКИ	97
ПЕРЕЛІК ПОСИЛАНЬ	98

ВСТУП

Об'єкт дослідження – дослідження застосування методів машинного навчання при розробці додатку для людей з обмеженими можливостями.

Предмет дослідження – мобільний додаток для людей з особливими потребами.

Мета роботи – дослідити ефективність впровадження технологій машинного навчання при розробці додатку для людей з особливими потребами.

Методи дослідження – проведення порівняльного аналізу, метод узагальнення.

В сучасному світі, ми практично не можемо уявити собі життя без технологій. Саме технологічний прогрес допомагає людству відкривати нове і досягати небувалих висот в різноманітних видах діяльності. З часом нові прогресивні технології стають доступними для все більшої кількості людей.

Одну з ключових ролей в цьому процесі відіграють смартфони. Станом на сьогодні, смартфон – це повноцінний портативний комп'ютер. Він має свою операційну систему (Android, IOS, Windows Phone), можливість завантажувати та видаляти різноманітні програмні продукти, можливість налаштування системних інтерфейсів для зручності роботи користувача і т.д.

Метою даної роботи є спрощення користування сервісом онлайн-перекладу для людей з обмеженими можливостями, шляхом розробки кастомізованого мобільного Android-дodatку та застосування методів машинного навчання.

Актуальність: Для людей з обмеженими можливостями сучасні технології мають набагато більше значення ніж для інших, адже вони не просто допомагають комфортніше облаштувати свій побут, а й надають можливість жити максимально повним життям, наскільки це можливо. На сьогоднішній день, люди з обмеженнями слухового та зорового апарату мають змогу використати цифрові технології у своєму житті. В даній роботі буде проведено дослідження та порівняння різноманітних методів машинного навчання та їх вплив на розробку сучасних мобільних додатків для людей з особливими потребами.

1 ДОСЛІДЖЕННЯ ВПЛИВУ ІТ-ТЕХНОЛОГІЙ ТА СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ НА ЖИТТЯ ЛЮДЕЙ З ОСОБЛИВИМИ ПОТРЕБАМИ

1.1 Використання сучасних технологій людьми з обмеженими можливостями

Сьогодні ми не можемо уявити наше життя без технологій. Вони проникли у всі сфери людської діяльності. Без них, багато процесів у світі не могли б відбуватися взагалі. Саме сучасні передові технології дозволяють нам дізнаватися про те, що відбувається в світі, дозволяють нам спілкуватись між собою в незалежності від географічного розташування, нівелюють такі поняття як “мовний бар’єр”, стирають кордони і дають можливість самовиражатися всім без виключення людям. Також, сучасні технології допомагають нам у повсякденних справах і роблять наш побут простішим і приємнішим. Отже, можна стверджувати, що технології стали нашим вірним помічником у сучасному світі.

Проти є великий соціальний прошарок, який потребує турботи та допомоги набагато більше ніж звичайні люди. Це – інваліди або люди з обмеженими можливостями. Для них звичні побутові речі, які ми робимо кожен день, не звертаючи уваги, можуть бути серйозною проблемою. Такі люди значно більше потребують наявності сучасних технологій, що допомагатимуть їм як у повсякденних справах так і в глобальному сенсі, тобто робитимуть їх життя максимально активним та повноцінним [2].

Такі технології називаються Допоміжними технологіями (Assistive Technologies). Вони призначені для підвищення функціональних можливостей людей з інвалідністю. Вони можуть бути дуже різноманітними, (прикладі показано на рис 1.1). Для них звичні побутові речі, які ми робимо кожен день, не звертаючи уваги, можуть бути серйозною проблемою. Деякі з них відносно прості та низькотехнологічні, такі як, наприклад, окуляри для читання, милиці, слухові апарати

або інвалідні візки. Інші – більш просунуті, розроблені з використанням передових наукових досягнень та сучасних технологій.

Деякі популярні ІТ-технології також перетворились із нішових (розроблених спеціально для людей з обмеженими можливостями) на загальноживані. Наприклад, такі технології як “text-to-speech” – читання вголос написаного на екрані тексту та технологія голосових команд смартфонів були спочатку розроблені спеціально розроблені для людей з вадами зору. Такі технології називаються Допоміжними технологіями (Assistive Technologies) [1].

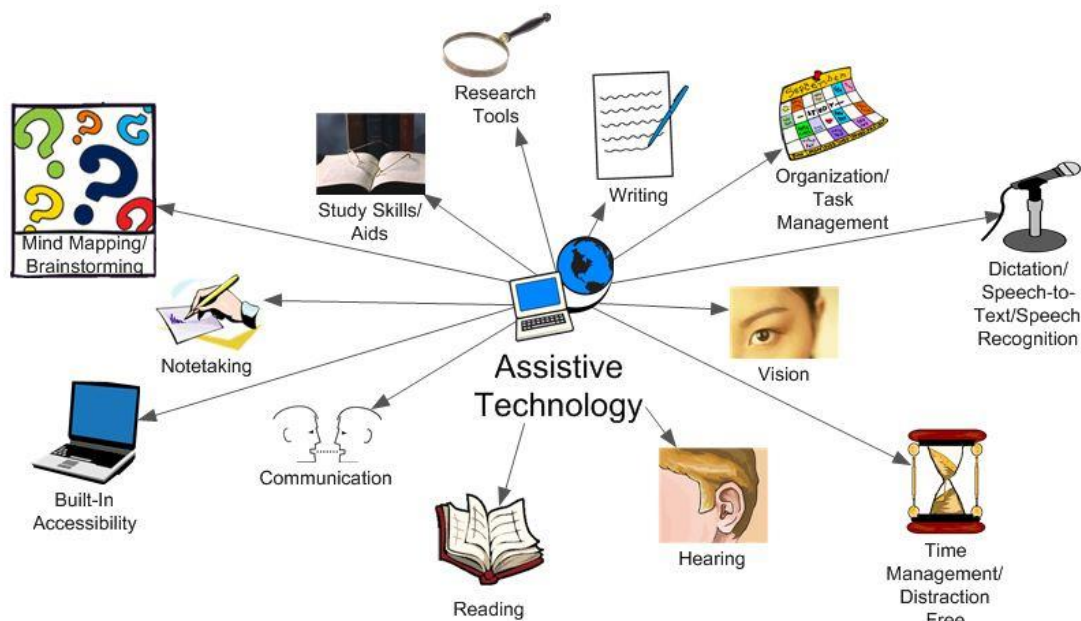


Рисунок 1.1 – Допоміжні технології

Деякі з них відносно прості та низькотехнологічні, такі як, наприклад, окуляри для читання, милиці, слухові апарати або інвалідні візки. Інші – більш просунуті, розроблені з використанням передових наукових досягнень та сучасних технологій.

Справді, поняття допоміжних технологій доволі широке. Вони призначені для підвищення функціональних можливостей людей з інвалідністю. Вони можуть бути дуже різноманітними. Допоміжною технологією може бути будь-який елемент,

частина обладнання, комп'ютерна програма або будь-яка система, що використовується для збільшення, підтримки чи вдосконалення функціональних можливостей людей з обмеженими можливостями.

Види допоміжних технологій(рис. 1.1):

— допоміжні технології можуть бути низькотехнологічними: картонні дошки для спілкування;

— допоміжні технології можуть бути високотехнологічними: спеціалізовані комп'ютери;

— допоміжними технологіями може бути спеціалізоване обладнання: протези, милиці, візки;

— допоміжними технологіями може бути комп'ютерне обладнання: спеціалізовані вимикачі, клавіатури та інші спеціальні контролери;

— допоміжними технологіями можуть бути комп'ютерні програми: зчитувачі тексту з екрану, та спеціальні програми для комунікації;

— допоміжними технологіями можуть бути інклюзивні спеціалізовані навчальні програми чи посібники;

— допоміжними технологіями можуть бути спеціалізовані навчальні комп'ютерні програми;

— допоміжні технології можуть бути дуже різноманітними: електронні пристрої, інвалідні коляски, ходунки, фіксатори, підйомники, тримачі для олівців та багато іншого.

Розробка таких сучасних допоміжних технологій неодмінно матиме великий вплив на наше життя. Деякі з них зроблені спеціально під окремі потреби людей з обмеженими можливостями, інші ж мають багато спільного із загальноновживаними технологіями. Люди з обмеженими можливостями стикаються з бар'єрами усіх типів. Але сучасні технології дають їм змогу зменшити багато з цих бар'єрів [5].

Насправді, деякі з них є стандартними продуктами масового ринку, що налаштовуються під потреби користувача як і всі інші. Наприклад, планшети можуть бути налаштовані для інтерпретації мови жестів в режимі реального часу [3].

Звісно, допоміжні технології дуже необхідні людям з обмеженими можливостями при вирішенні різноманітних завдань. Але, окрім іншого, одна з найбільших переваг цих технологій – це можливість їх використання при навчанні дітей з особливими потребами.

Використання допоміжних технологій при навчанні допомагає дітям працювати над їх слабкостями і вдосконалювати сильні сторони. Це особливо важливо для дітей з вадами розумового розвитку. Саме такі технології допомагають таким дітям оволодівати фундаментальними навичками читання, письма, рахування та ін. Допоміжні технології дають змогу дітям з особливими потребами розвиватись і ставати більш успішними і продуктивними учнями та студентами. Саме це допомагає їм розвивати в собі впевненість у власних силах і відчуття психологічної незалежності.

Нажаль, не всі люди (особливо консервативно налаштовані) визнають велику користь у впровадженні сучасних допоміжних технологій у навчання дітей з особливими потребами.

Одним з найбільших міфів є те, що використання допоміжних технологій не дозволить дитині здобувати елементарні необхідні навички. Проте вчені та експерти наголошують на хибності таких тверджень, адже, наприклад, прослуховування аудіокниг не заважає дитині вчитися читати. Тобто, допоміжні технології при правильному не заміщують собою якісь навички під час навчання, а лише дають дитині з особливими потребами можливість навчатися та розвиватися в незалежності від розладів її здоров'я. Звісно, допоміжні технології дуже необхідні людям з обмеженими можливостями при вирішенні різноманітних завдань. Саме такі технології допомагають таким дітям оволодівати фундаментальними навичками читання, письма, рахування та ін. Допоміжні технології дають змогу дітям з

особливими потребами розвиватись і ставати більш успішними і продуктивними учнями та студентами

Важливо пам'ятати, що не зважаючи на велику кількість переваг, допоміжні технології не можуть повністю “вилікувати” проблеми навчання та уваги всіх дітей з особливими потребами. Також треба розуміти, що вони не можуть замінити хороше професійне педагогічне навчання. І в будь-якому разі, такі діти потребують підвищеної уваги та турботи з боку батьків, а також професійного лікування.

1.2 Доступність сучасних ІТ-технологій для людей з особливими потребами

ІТ – технології стрімко розвиваються та інтегруються у всі сфери людського життя. Їх стрімкий розвиток можна помітити неозброєним оком. Ще 10-15 років тому, персональні комп'ютери були далеко не на всіх підприємствах. Ще менше їх було у людей вдома. Сьогодні ж, людина, що не має ПК в себе вдома, виглядатиме “білою вороною”. Люди з обмеженими можливостями стикаються з бар'єрами усіх типів. Але сучасні технології дають їм змогу зменшити багато з цих бар'єрів. Використовуючи комп'ютерні технології для таких завдань, як читання та написання документів, спілкування з іншими людьми та пошук інформації в Інтернеті, студенти та працівники з обмеженими можливостями здатні самостійно працювати з більш широким спектром даних. Тим не менш, люди з інвалідністю стикаються з різними бар'єрами для використання комп'ютера [7].

Ці бар'єри можна розділити на три функціональні категорії:

- перешкоди, пов'язані із пристроями вводу інформації;
- інтерпретація виведеної на екран інформації;
- читання супровідної документації.

Апаратні та програмні засоби (відомі як адаптивні або допоміжні технології) розроблені для забезпечення функціональних альтернатив для цих стандартних операцій.

Для всебічного дослідження доступності сучасних ІТ-технологій для людей з обмеженими можливостями, ми візьмемо з них декілька найуразливіших груп і дослідимо для них технологічну доступність згідно трьох вищевказаних критеріїв.

Люди з обмеженою рухливістю. Деякі інваліди-колясочники не можуть комфортно сидіти за стандартним комп'ютерним столом, адже він не підходить їм по висоті. А деякі користувачі комп'ютера не достатньо володіють своїми руками щоб користуватись стандартною клавіатурою або мишкою.

Введення даних. Обладнання, яке забезпечує гнучкість у розташуванні моніторів, клавіатур, документації та стільниць, корисно для багатьох людей з обмеженими можливостями. Підключення всіх комп'ютерних компонентів до розетки з доступними вимикачами дає можливість деяким особам самостійно включати та вимикати обладнання [6].

Та чи стали ІТ-технології більш доступними для людей з обмеженими можливостями? Можна сказати, що доступність ІТ-технологій за останні роки стрімко зросла.

Деякі технології допомагають людям з обмеженими можливостями практично не користуватися своїми руками при використанні стандартної клавіатури.

Особи, які використовують один палець, або мають доступ до спеціальної палички, що одягається на голову або користуються іншим вказівним пристроєм, можуть керувати комп'ютером, натискаючи клавіші за його допомогою. Програмні утиліти можуть створювати "липкі клавіші", які в електронному вигляді фіксують SHIFT, CONTROL та інші ключові клавіші, щоб дозволити їх послідовним натисканням вводити команди, які зазвичай вимагають одночасного натискання двох або більше клавіш. Функція повторення клавіш може бути відключена для тих, хто не може достатньо швидко відпустити клавішу, щоб уникнути декількох виділень.

Клавіатура (тверді шаблони з отворами над кожною клавішею для точного вибору) можуть використовуватися користувачами з обмеженим керуванням моторними функціями.

Іноді зміна розташування клавіатури та монітора може підвищити доступність. Наприклад, встановлення клавіатури, перпендикулярно столу, може допомогти особам з обмеженою рухливістю, які використовують вказівні пристрої для натискання клавіш. Інші прості модифікації обладнання можуть допомогти людям з порушенням рухливості. Наприклад, напрямлювачі для дисків можуть допомогти вставити та видалити диски.

Для людей, яким потрібно керувати комп'ютером однією рукою, доступні ліво та правостороння клавіатура. Вона забезпечує більш ефективні розташування клавіш, ніж стандартні клавіатури, розроблені для дворуких користувачів. Спеціалізовані ж трекболи та джойстики можуть повністю замінити мишу для людей з обмеженими можливостями.

Розпізнавання мови є ще одним механізмом, що допомагає людям з порушеннями рухливості набирати текст та керувати комп'ютером. Спеціальне програмне забезпечення може додатково допомогти тим, хто має порушення рухливості. Розширення аббревіатур (макро) та програмне забезпечення для прогнозування слів дозволяють зменшити потреби в запитах для часто використовуваних текстових та клавіатурних команд. Наприклад, програмне забезпечення для прогнозування слів передбачає цілі слова після декількох натискань клавіш і збільшує швидкість введення.

Виведення на екран. Зчитування інформації з екрану не становить проблеми для людей з обмеженнями рухливості, але, наприклад використання такого приладу виводу інформації, як принтер може вимагати сторонньої допомоги.

Документація. Екранна довідка забезпечує ефективний доступ до посібників користувача особам, які не можуть перегорнути сторінки книзі.

Особи з вадами зору (сліпота). Особи, які є сліпими, не можуть отримати доступ до візуальних матеріалів, представлених на екрані комп'ютера або в друкованому вигляді.

Введення даних. Більшість людей з вадами зору користуються звичайними клавіатурами, проте на ринку представлені і спеціальні клавіатури зі шрифтом Брайля, що надають сліпим людям можливість повноцінно користуватися таким пристроєм вводу даних як клавіатура.

Виведення на екран. Системи виводу мови використовуються для читання екранного тексту користувачами з вадами зору. Спеціальні програмні продукти зчитують інформацію з екрану комп'ютера, а голосовий синтезатор “промовляє” текст. Навушники дають людям, що користуються подібними програмами, можливість не турбувати синтетичним голосом оточуючих. Брайлівські дисплеї дозволяють переробляти відсканований програмою текст в шрифт Брайля. Такі дисплеї добре підходять для людей, що добре знайомі з цим шрифтом, адже це дозволяє читати екранний текст набагато швидше ніж за допомогою голосового синтезатора. Також вони незамінні для детального редагування тексту (програмування, створення документації). Окрім цього також існують спеціальні принтери для друкування тексту шрифтом Брайля.

Документація. Сканери з оптичним розпізнаванням символів можуть зчитувати друковані матеріали і зберігати їх на комп'ютер, де вони можуть бути прочитані за допомогою мовного синтезатора, Брайлівського дисплея, або роздруковані на принтері з підтримкою шрифту Брайля. Електронні версії документацій можуть бути прочитані користувачами з вадами зору за допомогою вищезазначених спеціалізованих програмних продуктів.

Для деяких людей з порушеннями зору стандартний розмір літер на екрані або надрукований у документах замалий для читання. Деякі люди не можуть відрізнити одного кольору від іншого.

Введення даних. Більшість людей, які мають порушення зору, можуть використовувати стандартні клавіатури, але іноді корисно використовувати великі клавіші, розроблені спеціально для таких людей.

Виведення на екран. Спеціальне обладнання для людей з вадами зору може змінити вигляд дисплея комп'ютера. Комп'ютерні символи, як текст, так і графічні зображення, можуть бути збільшені на моніторі чи принтері, тим самим дозволяючи людям з низьким рівнем зору використовувати стандартний текстовий редактор, електронні таблиці, електронну пошту та інші програмні додатки. Для людей з деякими порушеннями зору, здатність налаштувати колір монітора або змінити колір переднього плану та фону також має значення. Наприклад, спеціальне програмне забезпечення може перетворити екран з чорного-білого на біло-чорний для людей, які чутливі до світла. Екрани з антибліковим покриттям полегшують читання з них. Системи голосового виводу також використовуються людьми з низьким рівнем зору.

Особи з проблемами слухового та мовного апаратів. Порушення мови та слуху самі по собі не дуже заважають користуванню комп'ютером. Проте розширені синтезатори мовлення достатньо близькі до якості голосу людини, щоб діяти як замітники голосу, і таким чином надавати компенсаційний інструмент для людей, які не можуть спілкуватися у вербальній формі.

Введення даних. Люди з порушеннями слуху або мови зазвичай використовують стандартну клавіатуру та мишу.

Виведення на екран. Альтернативні варіанти виводу звуку можуть допомогти комп'ютерному користувачеві, з порушеннями слуху. Наприклад, якщо гучність звуку повернута до нуля, комп'ютер може вивести на екран панель меню, де можна керувати налаштуваннями аудіовиходу.

Документація. Особи, що мають порушення слухових або мовних функцій, як правило, не стикаються зі складностями при використанні стандартної письмової (друкованої) чи екранної документації.

Особи зі специфічними розладами при навчанні. Сучасне спеціалізоване освітнє програмне забезпечення, завдяки якому комп'ютер забезпечує мульти-сенсорний досвід користувача, різноманітну взаємодію з ним, індивідуалізоване навчання та повторення, що може бути корисним для набуття різноманітних вмінь та навичок користувачам з розладами навчання. Деякі з них, мають труднощі з обробкою письмової інформації. Вони також можуть отримати користь від виконання завдань письмової форми опрацювання, завдяки комп'ютерним утилітам. Наприклад, стандартний текстовий редактор може бути цінним інструментом для людей з дисграфією, неможливістю писати вручну.

Введення даних. Тихі робочі зони та засоби для захисту вух можуть полегшити використання комп'ютера для людей з обмеженими фізичними можливостями, які мають підвищену чутливість до фонового шуму. Програмне забезпечення, яке допомагає ефективному та точному вводу даних, також може допомогти. Деякі люди можуть знизити показники вхідних помилок, використовуючи перевірки правопису та граматики. Крім того, програми для прогнозування слів (програмне забезпечення, яке передбачає цілі слова з фрагментів) успішно використовуються особами з обмеженими можливостями навчання. Аналогічним чином, спеціальне програмне забезпечення, що розширює аббревіатури, може зменшити необхідність запам'ятовування команд клавіатури та полегшити написання будь-якого тексту.

Виведення на екран. Програмне забезпечення, яке допомагає ефективному та точному вводу даних, також може допомогти. Деякі особи з труднощами в навчанні використовують корисні адаптивні пристрої, призначені для тих, хто має порушення зору. Зокрема, великі дисплеї, альтернативні кольори на екрані комп'ютера та голосовий вивід інформації (мовний синтезатор) можуть компенсувати деякі проблеми читання. Люди, з ускладненою інтерпретацією візуального матеріалу, можуть краще розуміти, виявляти та виправляти помилки, коли візуальна інформація подається в аудіовигляді або друкується у великих шрифтах та символах. Програмне

забезпечення, яке допомагає ефективному та точному вводу даних, також може допомогти.

Документація. Деяким особам з труднощами в навчанні, важко читати. Комп'ютерна документація, зроблена в електронній формі, може бути збільшена на екрані та / або читатися вголос системами синтезу мовлення, задля того, щоб зробити її доступнішою для таких людей [4].

1.3 Дослідження ринку ІТ-технологій для людей з особливими потребами

На сьогоднішній день стає зрозумілим той факт, що життя сучасного суспільства неможливе без інформаційного простору. Люди з обмеженими можливостями, я ніхто, потребують сучасних ІТ – технологій для спрощення життя і вирішення багатьох нагальних проблем. Тому, на ринку інформаційних технологій у великій кількості представлені продукти для людей з особливими потребами.

Розрізняють два основних види продуктів для людей з обмеженими можливостями на ринку ІТ – технологій:

- фізичні продукти;
- програмні продукти;
- комбіновані продукти.

До фізичних продуктів відноситься різноманітне устаткування та технологічні пристрої, до програмних продуктів відносяться програми, що полегшують людям з особливими потребами роботу з обчислювальною технікою, до комбінованих рішень відносяться технологічні комплекси, що включають до себе як фізичні прилади, так і програмне устаткування, пристосоване до потреб певних категорій людей з обмеженими можливостями. Прикладом такої технології є розробка команди Boston Retinal Implant Project. Вони створили біонічний імплант ока, призначений для людей, що втратили зір. Реагуючий на світло мініатюрний пристрій розміщується на тильній

стороні очного яблука позаду сітківки. Сигнали з цього пристрою передаються електричними імпульсами до мозку людини за допомогою дуже тонких проводів, з'єднаних з необхідними нервами. В ході розробки даного пристрою вченим та інженерам прийшлося стикнутися не лише с проблемою мініатюарізації імпланта, а й з проблемою надійності та довговічності його роботи. До цього високотехнологічного пристрою висуваються серйозні умови – він має бути вологостійким та надійно функціонувати на протязі не менше десяти років.

За словами розробників, імплант поки не здатний повернути інвалідам повноцінний зір, але дозволяє бачити обриси предметів в чорно-білому кольорі. Звісно, не можна назвати розробку Boston Retinal Implant Project ідеальною у всіх аспектах. Однією з причин цього є вузька направленість даного приладу. Тобто люди з глаукомою не зможуть його використовувати. Також він не зможе допомогти людям, що народилися вже незрячими через атрофію або вроджену відчуженість зорових нервів.

Отже даний технологічний прилад справді полегшує життя певної групи людей з обмеженими можливостями, проте в нього є серйозні недоліки через вузьку направленість даної технології та її високу вартість, що робить її малодоступною для більшості людей, які її потребують.

До даної категорії фізичних продуктів сучасного ринку інформаційних технологій відносяться різноманітні “розумні” протези, візки, ліжка, та інші прилади, що заточені під потреби різноманітних груп людей з особливими потребами. Такі технологічні і прості у використанні. Так, деякі з них, в залежності від потреб користувача керуються за допомогою рухів пальцями, м'язами обличчя або інших частин тіла, голосом, або навіть думками, за допомогою сучасних нейро-розпізнавальних систем. Такі пристрої значно спрощують життя людей з обмеженими можливостями, але далеко не завжди є доступними для них, з огляду на їх досить високу вартість.

Альтернативні пристрої введення даних: Деякі користувачі не можуть використовувати мишу або клавіатуру для роботи на комп'ютері. Однією з причин цього є вузька направленість даного приладу. Ці люди мають змогу користуватися різноманітними технологічними пристроями для таких цілей. Головні вказівники: паличка або об'єкт, встановлений безпосередньо на голові користувача, який можна використовувати для натискання клавіш на клавіатурі. Цей пристрій використовується особами, які не користуються своїми руками.

Відстеження руху очей користувача - ця технологія може включати в себе пристрої, які спостерігають за очима користувача, щоб інтерпретувати, де він хоче розмістити вказівник миші і переміщує його для користувача.

Також яскравими прикладами технологій для спрощення життя людей з обмеженими можливостями є такі прилади: Пристрій ReadingPen від Wizcomtech сканує написаний текст і перетворює його в мову. Орієнтований на людей з обмеженими можливостями в навчанні, ReadingPen використовується дуже просто. Достатньо пронести прилад над текстом для того, щоб він конвертував і озвучив написане.

Intel Reader - це пристрій розміром з кишеньковий комп'ютер, який сканує текст і конвертує його у цифровий стан, а потім перетворює отриманий текст на мову. Це допомагає людям з обмеженими можливостями читання, слабким зором і людям зі сліпотою. Intel Reader також може взаємодіяти з комп'ютером через USB-з'єднання.

До програмних продуктів ринку ІТ-технологій відносяться

- програми для збільшення або зчитування інформації з екрану;
- програми для розпізнавання мови;
- віртуальні маніпулятори та пристрої вводу;
- програмне забезпечення для читання тексту;
- програмні комплекси для комунікації.

Програми для збільшення інформації на екрані є необхідними для людей з деякими вадами зору та для слабозорих. Такі програми допомагають у користуванні

браузерами, текстовими редакторами та іншими програмами. Програмне забезпечення для збільшення екрану дозволяє користувачеві збільшити зображення тексту і графіки, що відображаються на екранах їх комп'ютерів. Це робиться шляхом перехоплення відеосигналу, що надходить на монітор, потім збільшення і посилення його перед відправкою на дисплей. Як комп'ютери з операційною системою IOS, так і комп'ютери під управлінням Windows пропонують базову функцію збільшення екрану в своїх операційних системах [10].

Вхідні пристрої для одного перемикача. Ці види пристроїв можуть використовуватися з іншими альтернативними пристроями введення або самі по собі. Зазвичай вони використовуються з екранними клавіатурами. Екранна клавіатура має опцію переміщення курсору по клавішах, і коли клавіша, яку хоче використати користувач, знаходиться у фокусі, він натискає перемикач. Ця технологія також може працювати на веб-сторінці: курсор може переміщатися по веб-сторінці, і якщо користувач хоче натиснути на посилання або кнопку, йому треба дочекатися, коли це посилання або кнопка знаходиться у фокусі, і активувати перемикач.

Деякі програми доступні безкоштовно, але вони також мають обмежені функції і обмежений діапазон збільшення. Відносно дешеві програми можна завантажити з інтернету, але вони також пропонують лише мінімальні можливості. Як показало опитування людей з вадами зору та слабозорих, що користуються подібними програмними рішеннями, більшість із них, вважають такі програми не дуже прийнятними, адже люди потребують більш складних функцій і глибокого індивідуального налаштування, що пропонують їм більш дорогі комерційними аналоги.

Деяким користувачам зі слабким зором важко читати чорний текст на яскравому білому тлі, тож програми збільшення екрану дозволяють користувачеві вибирати різні комбінації тексту та кольору фону. Найбільш ефективний спосіб взаємодії з програмами збільшення екрану - використання клавіатурних команд. Проте деяким людям зручніше користуватися мишею для цих потреб. Проте не всі люди з вадами

зору можуть із легкістю знайти на екрані курсор миші. З цієї причини виробники включають функції, що дозволяють користувачам вибирати розмір і колір, а також інші візуальні аспекти курсора, що полегшує їм роботу з комп'ютером. Яскравими прикладами подібних технологій є такі програми: "Zoom Text Express Zoom", "Lightning Plus", "Lunar", "Zoom Text Magnifier"[9].

Програмне забезпечення для розпізнавання мови надає користувачам, що мають труднощі при введенні інформації, альтернативний спосіб введення тексту та керування комп'ютером. Користувачі можуть надавати системі певні обмежені команди для виконання дій миші. Користувачі можуть наказати системі натиснути посилання або кнопку або скористатися пунктом меню. Прикладом може бути програма "Dragon Naturally Speaking" для операційних систем Windows або IOS. Самі ж операційні системи Windows і IOS мають деякі власні утиліти для розпізнавання мови, але вони не можуть бути використані для перегляду веб-сторінок[8].

Деякі системи розпізнавання мови вимагають «навчання» (також званого «enrollment»), де окремий оратор читає текст або ізольований словник у систему. Система аналізує голос конкретної людини і використовує його для точного налаштування розпізнавання мови цієї особи, що призводить до підвищення точності. Системи, які не використовують навчання, називаються "незалежними від диктора". Системи, що використовують навчання, називаються "залежними від диктора".

Програми розпізнавання мовлення включають голосові користувальницькі інтерфейси, такі як голосовий набір, маршрутизація викликів, пошук, просто введення даних, підготовка структурованих документів, визначення характеристик мовника, обробка мовлення в текст, та інші функції.

Термін розпізнавання голосу або ідентифікація спікера відноситься до визначення того хто говорить, а не до того, що вони говорять. Розпізнання особи, що говорить може спростити завдання перекладу мови в системах, які пройшли навчання за голосом конкретної людини, або може бути використана для аутентифікації або перевірки особистості мовника як частини процесу безпеки.

З технологічної точки зору, розпізнавання мови має довгу історію з кількома хвилями значних нововведень. Зовсім нещодавно ця сфера отримала стрімкий розвиток завдяки досягненню успіхів у технології глибокого навчання (“Deep learning”) та наявності великих обсягів голосових даних. Неймовірні темпи розвитку цієї технології зумовлені не лише її користувацькою популярністю, але що більш важливо, впровадженням різноманітних методів глибокого навчання при проектуванні та розгортанні систем розпізнавання мови в усьому світі.

Програмне забезпечення для читання тексту забезпечує доступ до письмового матеріалу людям з порушеннями зору, рухового апарату та вадами навчання, усуваючи необхідність читати та інтерпретувати текст. Завдяки зручності цифрових технологій і їх доступності через мережу інтернет, на сьогоднішній день, більшість текстових даних знаходиться в електронному вигляді. Саме тому, дана технологія набула дуже широкого розповсюдження по всьому світу. Подібні програмні рішення включають безкоштовне програмне забезпечення, комерційне програмне забезпечення та комбінації апаратного / програмного забезпечення [5].

Однією з найяскравіших представників інформаційних рішень даного типу є програма “Windows Narrator”. Розпізнавання особи, що говорить може спростити завдання перекладу мови в системах, які пройшли навчання за голосом конкретної людини, або може бути використана для аутентифікації або перевірки особистості мовника як частини процесу безпеки.

Вбудована в операційну систему Windows програма включає в себе базовий екранний зчитувач під назвою “Narrator”, який конвертує текст і описує користувачеві стан робочого екрану та все що на ньому відбувається. Перекладач вмикається, комбінацією клавіш "Windows+Enter". Програма відкриється та почне зчитувати інформацію що знаходиться у вікні “Narrator” . Ви можете застосовувати різноманітні команди, налаштувати параметри або перейти до інших сторінок з підтримкою зачитування тексту на домашньому екрані “Narrator”.

Також, одним з найпопулярніших програмних рішень в даному сегменті є програма “NVDA”. NonVisual Desktop Access (NVDA) - це безкоштовний портативний зчитувач тексту з відкритим вихідним кодом для Microsoft Windows. Проект розпочав Майкл Керран у 2006 році. NVDA запрограмований у Python. В даний час він працює виключно з доступними API, такими як UI Automation, Microsoft Active Accessibility, IAccessible2 та Java Access Bridge, замість використання спеціалізованих відео драйверів для "перехоплення" та інтерпретації візуальної інформації. Основні переваги даної платформи виглядають таким чином:

- підтримка популярних програм, включаючи веб-браузери, такі як Mozilla Firefox і Google Chrome, клієнти електронної пошти, програмне забезпечення для чату в Інтернеті, музичні програвачі та офісні програми, такі як Microsoft Word і Excel;

- вбудований синтезатор мовлення, що підтримує більше 50 мов, а також підтримку різних голосів;

- звітування про текстове форматування, якщо такі є, наприклад, ім'я та розмір шрифту, стиль та орфографічні помилки;

- автоматичне оголошення тексту під мишкою і додатковий звуковий індикатор позиції миші;

- підтримка багатьох оновлюваних Брайлівських дисплеїв, включаючи введення шрифту Брайля через брайлівські дисплеї, які мають клавіатуру Брайля;

- можливість повноцінно працювати з флеш-пам'яттю USB або інших портативних носіїв без необхідності установки;

- простий у використанні розмовний інсталятор;

- перекладено на більш ніж 50 мов;

- підтримка сучасних операційних систем Windows, включаючи 32 і 64-розрядні варіанти;

- можливість запускати під час входу в систему Windows та інших безпечних екранів;

— оголошення елементів керування та тексту під час взаємодії з жестами на сенсорних екранах.

Програмних продуктів для комунікації людей з обмеженими можливостями між собою та з навколишнім світом є доволі багато. Вони всі дуже різноманітні, адже спеціалізовані під конкретні потреби певних груп людей з особливими потребами. Наприклад люди, які мають проблеми з мовою, можуть легко розчаруватися, намагаючись спілкуватися з вихователями, друзями, сім'єю та близькими. Незалежно від того, чи вони не можуть говорити через інвалідність, чи їхні мовні здібності є занадто слабкими, а слова незрозумілими, нездатність повідомляти свої власні думки, почуття, потреби та бажання неймовірно засмучує і обмежує. Батьки і опікуни також можуть розчаруватися, коли вони не в змозі зрозуміти свого близького. Завдяки технології, яка доступна сьогодні, засоби комунікації стають набагато більш поширеними та доступними. Зараз існує багато програм, які допомагають людям з обмеженими можливостями говорити та спілкуватися легше та ефективніше. Залежно від інвалідності людини або конкретної боротьби, програми з різними функціями працюватимуть краще для різних людей. Ось перелік деяких програм, що полегшують життя таким людям:

iConverse - інструмент, призначений для людей з комунікативними обмеженнями. iConverse - це додаток AAC для iPhone та iPod Touch, який замінює громіздкі та дорогі комунікаційні пристрої, які використовувалися в минулому. Ця програма надає користувачам можливість переносити пристрій для AAC - Аугментативної альтернативної комунікації (Augmentative Alternative Communication) у свою кишеню, створюючи відповідний, ефективний і зручний засіб спілкування. Він поставляється з 6 основними "плитками" для зв'язку, які представляють найбільш основні потреби людини. При активації на дотик, піктограми дають як слухове, так і візуальне представлення конкретної потреби або бажання. iConverse дозволяє батькам, вчителям та широкій громадськості зрозуміти основні потреби та бажання людей, які не в змозі усвідомлено спілкуватися;

Open Sesame - Мільйони людей мають обмеження у використанні своїх власних рук, що перешкоджає їм виконувати діяльність, яку більшість з нас приймає як належне, а звичайні телефони не обладнані для потреб людей з обмеженою рухливістю. Open Sesame використовує передню камеру будь-якого пристрою Android для відстеження руху голови. Програма допомагає таким людям розблокувати смартфон без дотиків, вводити текст, здійснювати телефонні дзвінки та використовувати інші функції.

Predictable – програмне рішення, що призначене для грамотних людей, які не мають здатності говорити, через церебральний параліч, хворобу моторних нейронів, ларингектомію, аутизм, інсульт, апраксію, черепно-мозкову травму або будь-який інший стан, що впливає на мову людини. Коли користувачі починають набирати текст, програма " Predictable " передбачає, що саме людина хоче сказати і формує пропозицію повного слова, або навіть речення. Також програма має можливість озвучити запропоноване слово чи речення, надаючи голос людям, які його не мають.

1.4 Огляд та характеристика мобільних додатків, призначених для людей з обмеженими можливостями

Оскільки, на сьогоднішній день інтегрованість людей з обмеженими можливостями в цифровий світ є досить високою, можна стверджувати, що більшість з них мають в своєму розпорядженні смартфони. Це означає, що допоміжні технології (Assistive Technologies) мають знайти своє відображення на ринку мобільних додатків. Тож далі розглянемо найяскравіших представників цього сегменту ринку. Одним з них є Talkitt.

Статистика говорить про те, що 1,5% населення світу, через бічний аміотрофічний склероз, церебральний параліч, аутизм, травматичну черепно-мозкову травму, інсульт, хворобу Паркінсона або інші недуги, живуть з порушеннями мовлення. "Talkitt" - це новаторська технологія від компанії "Voiceitt", яка може

перекласти нерозбірливе мовлення в чітке для повноцінного спілкування з навколишнім середовищем.

В даному мобільному додатку використовується не звичайна технологія розпізнавання голосу. Наприклад, людина з церебральним паралічем може сказати "шо-ко-ла ", маючи на увазі "шоколад ". Додаток розпізнає, це завдяки персоналізованому словнику конкретного користувача. Завдяки інтегрованій системі "Deep learning", додаток працює з будь-якою мовою. Алгоритм машинного навчання спочатку аналізує введене користувачем голосове повідомлення(після прохання вимовити декілька стандартних фраз), потім створює для нього особистий словник, а потім розпізнає слова та фраз завдяки використанню розпізнавання образів, що персоналізовані для диктора. На даний момент, даний додаток доступний лише в тестовій версії. Даний мобільний додаток Sesame дозволяє перейти на повне управління смартфоном за допомогою голосу та жестів.

Додаток працює на ОС "Android" і фіксує рухи людини на відстані 40 – 60 см. За допомогою цього, власник додатку має можливість користуватися різноманітними функціями смартфона та відкривати інші додатки. Перевагою даного мобільного додатку є його ексклюзивні функціональні можливості, надійність та зручність в роботі.

Даний додаток підходить для користування людям з такими обмеженнями можливостей:

- травми хребта;
- розсіяний склероз;
- хвороба Лу Геріга;
- нервово-м'язові захворювання;
- церебральний параліч (залежно від контролю голови);
- ампутації;
- попередній хід;

— будь-який стан, що призводить до втрати рухомості в руках чи пальцях.

Перевагою даного мобільного додатку є його ексклюзивні функціональні можливості, надійність та зручність в роботі. До недоліків можна віднести вузьку направленість даного додатку, тобто він заточений на лише на виконання вузького набору функцій, втім така особливість характерна для всіх мобільних додатків, спеціалізованих для роботи з людьми з обмеженими можливостями. Цікаво, що першочергова ідея та принцип роботи були взяті з ігрової індустрії, де аналіз жестів користувача користується величезною популярністю, особливо на ігрових консолях.

Цей мобільний додаток був створений задля заміни слухових апаратів групою студентів із Берклі та Сан-Франциско. Програма допомагає людям з вадами слуху налагодити спілкування з навколишнім світом за допомогою конвертування мовлення в текстову форму на екрані смартфона.

Програма чутливо реагує на мовлення людини в реальному часі. Навіть при групових бесідах, реалізовані в програмі алгоритми дозволяють розпізнати репліки всіх учасників бесіди, визначити, кому саме вони належать та вивести результат на екран користувача у вигляді діалогу (подібного до месенджерів) з виокремленням різними кольорами цитат, що належать різним учасникам розмови. Таким чином, користувач може з легкістю зрозуміти не лише промовлену фразу, але й те, ким саме вона була промовлена, що є дуже корисною функцією під час бесіди за участю великої кількості людей [10].

Цікавим є той факт, що кожен з початкових розробників цього додатку має проблеми зі слуховим апаратом, проте природа захворювань у всіх різна. Саме через це, на думку експертів, молодим програмістам вдалось розробити такий універсальний програмний продукт, що підходить людям з будь-якими порушеннями та вадами слухового апарату. Ця програма сконструйована ізраїльським розробником ігор, якого звать Оded Бен Дов і паралізованим інженером-електриком Гіором Лівні. Вони працювали над складним програмним забезпеченням, використовуючи в якості апаратної платформи смартфон “Nexus 5” від “Google”.

Такий додаток є справді універсальним у своєму сегменті ринку, адже задовольняє потреби декількох підгруп людей з обмеженими можливостями.

Для глухих і людей з проблемами слухового апарату, таке завдання, як комунікація по телефону може стати дуже складним або навіть неможливим завданням. Хоча багато людей з вадами слуху здатні розмовляти по телефону за допомогою слухових апаратів, є й ті хто все ж не був в змозі користуватися телефонними дзвінками або жертвувати приватністю та ефективністю такого спілкування.

Останніми роками послуги текстових месенджерів і програм для відеочату (наприклад, Skype, FaceTime) стали безцінними засобами зв'язку для тих, хто не може користуватися традиційними голосовими викликами. Проте голосові дзвінки відіграють важливу роль у повсякденному житті більшості людей, як на робочому місці, так і за його межами. Тому, додаток під назвою “RogerVoice” створений для вирішення саме цієї проблеми.

“RogerVoice” використовує технологію передачі голосу по IP (VoIP) для передачі субтитрів для голосових дзвінків у режимі реального часу. Дане технологічне рішення робить можливими голосові розмови для мільйонів людей з різноманітними вадами та проблемами слухового апарату. Співзасновник та виконавчий директор Олівер Дженел, який сам має вади слуху. Вперше даний IT- продукт впевнено показав себе під час своєї кампанії на краудфайдинговій платформі “Kickstarter” у 2014 році [9]. Під час цієї кампанії, французький стартап отримав майже вдвічі більше фінансових інвестицій ніж розраховував на початку.

Однією з найбільш привабливих можливостей додатку “RogerVoice” є те, що вона може використовуватися для виклику будь-якого користувача, незалежно від того, чи має він, цей же додаток, встановлений на своєму телефоні. Після того, як встановлено з'єднання, людині, що отримує дзвінок надходить звуковий запит, про те, що виклик транскрибується. Після цього все, що говорить одержувач, переписується в режимі реального часу і відображається на екрані телефона

користувача (функція, відома як "автоматизовані субтитри в режимі реального часу"). Опція перетворення тексту в мову ("автоматизований синтез голосу") також доступна для користувачів даного мобільного додатку. Даний мобільний додаток є універсальним засобом для спілкування по телефону людьми з вадами слуху та проблемами слухового апарату. Він простий і ефективний в користуванні, що є дуже важливим критерієм при роботі з людьми із особливими потребами. Також тут наявна функція голосового синтезатору, що розширює кількість потенціальних користувачів даного програмного рішення [11].

1.5 Застосування нейронних мереж при розробці високотехнологічних рішень для людей з обмеженими можливостями

Особливі потреби людей з інвалідністю дуже різняться між собою. Тому і високотехнологічні рішення, спеціалізовані для таких людей є дуже різноманітними. Розглянемо приклади застосування нейронних мереж та технології машинного навчання для IT-рішень для людей з обмеженими можливостями.

Однією із методик застосування систем штучного інтелекту при розробці програмних комплексів для людей з особливими потребами є метод Обробки природної мови "Natural Language Processing(NLP)". Наразі дослідники працюють над аспектами ситуаційної моделі за допомогою привчання нейронних систем давати пояснення подіям описаним у певному тексті. ШІ має зробити висновок про мотиви людей, згаданих у тексті, при тому що в самому тексті їх мотиви описані не були. Також дослідники працюють над впровадженням технології представлення абстрактного значення "Abstract Meaning Representation (AMR)". Це підхід, що може використовуватися для побудови ситуаційної моделі на основі певного тексту з огляду на певні знання, здобуті читачем раніше.

Прогрес у дослідженні даної технології є дуже стрімким, проте залишаються неопрацьованими деякі аспекти. Одним з них є відсутність визначених та чітких

показників зрозумілості, незалежно від того, базуються вони на судженнях людини або на автоматизованих методах. Без них важко оцінити та порівняти нові підходи. Загальноживані критерії читабельності, хоча і корисні для порівняння ймовірної зрозумілості природних текстів, не підходять для порівняння методів спрощення, оскільки вони базуються на факторах, які співвідносяться із зрозумілістю, але самі по собі не важливі. Наприклад, природні тексти з коротшими реченнями, швидше за все, будуть зрозумілішими, але простішими. Наприклад, природні тексти з коротшими реченнями, швидше за все, будуть зрозумілішими, але просте розбиття довгих речень у тексті часто призводить до гіршої зрозумілості тексту.

Другим аспектом, що потребує доопрацювання можна вважати так звану “Мавп’ячу лапу(Monkey’s Paw)”. NLP -системи іноді виконують завдання, як зазначено, але не так, як би бажали або очікували дослідники. Наприклад, було проведено змагання між системами штучного інтелекту, яке мало на меті сприяти прогресу в розумінні сенсу історій. Від ШІ вимагалось запропонувати продовження тексту який би людські читачі вважали правдоподібними чи неправдоподібними. Один з найкращих “конкурсантів” навіть не читав оповідань, розсудивши, що самі продовження містять достатньо реплік, щоб добре виконати завдання. Ця система виконала чітку мету завдання, уникнувши обробки даних, на яку очікували дослідники [12], [13].

Дослідження показують, високу точність у передбаченні людського сприйняття певних текстових матеріалів. Такого результату біло досягнуто шляхом необхідної кількості навчальних ітерацій обробки даних. Науковці також зазначають, що продуктивність навчання подібних NLP- систем дає змогу впроваджувати такі рішення при розробці спеціалізованих програмних продуктів для людей з обмеженими можливостями. Проте, також науковці зазначають, що системам штучного інтелекту та машинного навчання ще необхідно пройти певний шлях розвитку та становлення аби забезпечити стабільно високоефективний результат при роботі з людьми з особливими потребами.

Також однією із широко розповсюджених технологій є віртуальні помічники(агенти), навчені за допомогою процедури машинного навчання. Зазвичай виділяють два типи таких асистентів із штучним інтелектом. У першому випадку, людина, яка бажає допомоги, ініціює взаємодію зі штучним інтелектом. Потім асистент повинен інтерпретувати запит, який робить особа, разом із відповідним контекстом, а потім виконати відповідні дії. Такий тип помічника називається активним, оскільки його не потрібно активувати за запитом від особи.

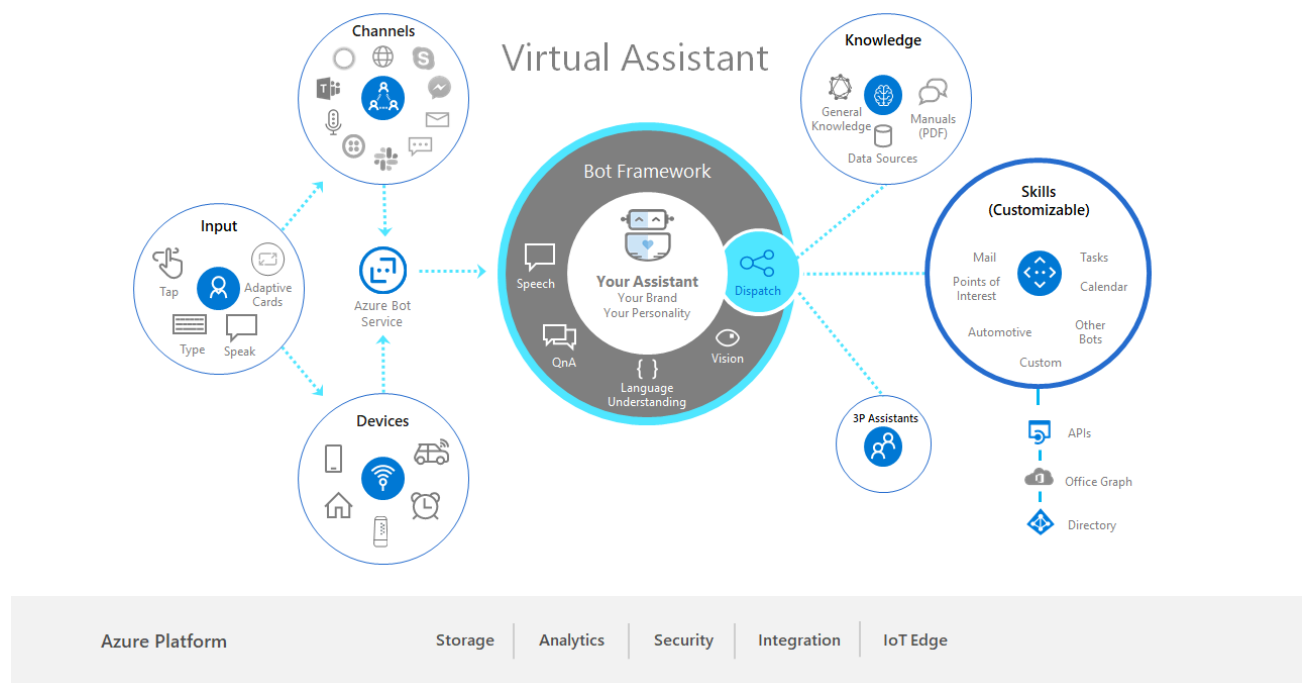


Рисунок 1.2 – Схема застосування активного агента

Конкретне застосування активного агента для людей з особливими потребами, може бути служба віддаленого догляду(рис.1.2).. Так само й активний штучний інтелект може навчитися виконувати цю роль, допомагаючи доглядачеві людини помічати ситуації, які потребують уваги?

Ще одним застосуванням для активного агента може бути навчання “м’яким” навичкам роботи, наприклад, робота в команді, прийняття рішень та спілкування. Наявність або відсутність цих навичок часто має вирішальне значення для

працевлаштування. Чи міг би активний штучний інтелект контролювати взаємодію людини на робочому місці та надавати пропозиції щодо вдосконалення? Дослідники машинного навчання сьогодні, як правило, песимістично ставляться до впровадження ШІ помічників як повноцінних співробітників. Вони відзначають дві ключові проблеми - виклик даних та виклик розуміння. Такий тип помічника називається пасивним тому, що він нічого не робить, поки його не активує людина. У другому випадку, штучний інтелект виявляє ситуацію, в якій людині потрібна допомога, без необхідності запиту зі сторони самої людини [14].

Проблема з даними впливає з того факту, що системи машинного навчання, як правило, вимагають великої кількості прикладів того, що система повинна спостерігати, разом із відповідною реакцією на кожен приклад. Оскільки різноманіття ситуацій, в яких може діяти штучний когнітивний асистент, є дуже великим збирати корисний набір прикладів є складним завданням. Однак є кілька способів вирішити цю проблему. Одним із них є використання технології навчання без вчителя, в якому приклади не потрібно пов'язувати з відповідними реакціями, що значно спрощує проблему. Сама система визначає якість можливих реакцій, часто перевіряючи якість моделі ситуації, яку вона будує. Для цього не потрібен людський вклад. Сьогодні доглядачі за людьми стежать за житловою площею клієнта, спостерігаючи за ознаками небезпеки чи лиха, які клієнт може не розпізнати або не мати змоги впоратись.

Як приклад, вчені описують камеру спостереження, яка вчиться виявляти незвичну поведінку на стоянці, будуючи модель поведінки, яку вона бачить. На прикладі автостоянки камера дізнається, що об'єкти певного розміру рухаються по певних шляхах, тоді як об'єкти різного розміру рухаються по різних шляхах. Ми знаємо, але камера не знає, що ці об'єкти - це машини та пішоходи.

Оскільки різноманіття ситуацій, в яких може діяти штучний когнітивний асистент, є дуже великим збирати корисний набір прикладів є складним завданням.

Наприклад, якщо об'єкт розміром з машину рухається в зону, де ці об'єкти зазвичай не рухаються, це можна помітити. Це можна дослідити як підхід до створення активних когнітивних помічників, для віддаленого догляду або для моніторингу активності на робочому місці, попередження тренера про щось незвичне, що відбувається.

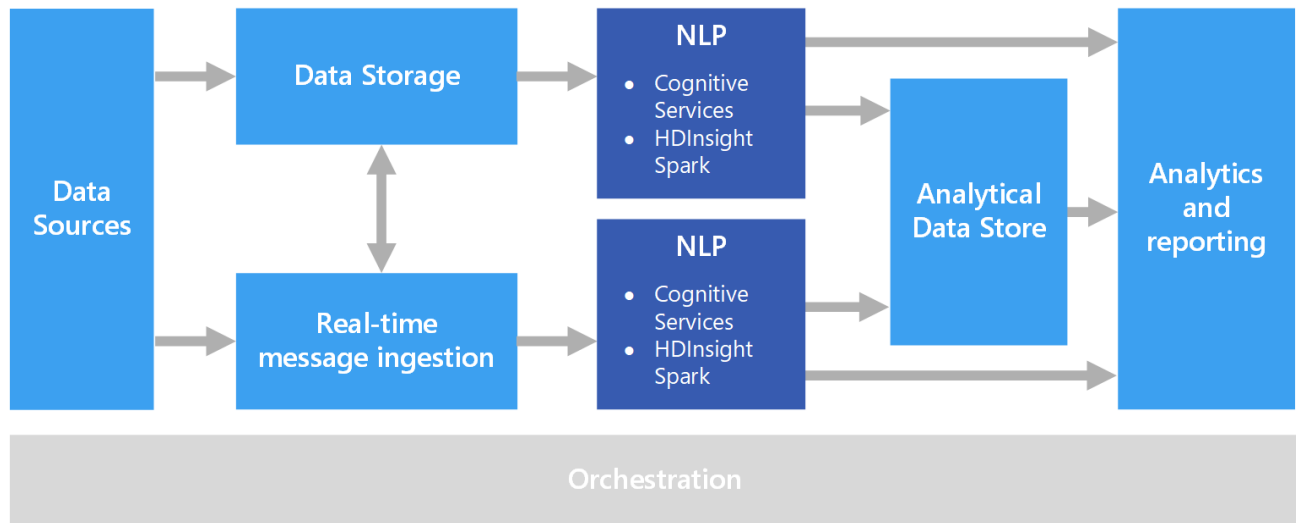


Рисунок 1.3 – Схема роботи системи ІІІ-помічника

Іншим підходом до обробки даних є попередня підготовка системи для обробки широкого кола різних ситуацій, а потім її спеціальне навчання для роботи в певній конкретній ситуації. Результат такої методики мета-навчання полягає в тому, що попередня глибока підготовка дозволяє системі в конкретній ситуації швидко розробити необхідний алгоритм дій маючи невелику кількість вхідних даних. Попередня підготовка може застосовувати або не використовувати спеціальну обробку даних, щоб підготувати основу для подальшого навчання[15].

Глибинна попередня підготовка буде кошовною, але вона буде давати можливість системі швидкого знаходити оптимальний алгоритм дій у більшій кількості ситуацій. Наприклад, віддалений монітор догляду може бути навчений у багатьох різних будинках клієнтів; тоді він може швидко і дешево навчитися

відстежувати події для нового клієнта в новому будинку. Проблемою в навчальних системах для вирішення дуже широкого кола завдань є ефект “забування”. Зміни в системі, необхідні для обробки нового завдання, можуть суперечити даним та алгоритмам, набутими за попередні завдання. Науковці повідомляють про прогрес у розробці новітніх рішень, які дозволяють вивчати нові завдання, не конфліктуючи зі старими [17].

Проблему розуміння можна проілюструвати на прикладі камери спостереження. Хоча камера може ідентифікувати незвичну поведінку, вона взагалі не розуміє того, що бачить. Як уже згадувалося, хоча ШІ-модель відрізняє предмети різного розміру, проте вона не має інформації про те, якими є ці об’єкти. Вона також не трактує різних траєкторій, описаних у її моделі, наприклад, що деякі траєкторії малих предметів (пішоходів) ведуть до входу в будівлю. Ці недоліки означають різницю між тим, що може зробити камера, і тим, що може зробити людський спостерігач. Спостерігаючи за пішоходом, людина-спостерігач розцінюватиме рухи не просто по тому, чи звичні вони, а по тому про які наміри, на її думку, вони свідчать. Переміщення біля вікна будівлі трактуватимуться інакше, ніж рух до незвичного місця на зовнішній стороні стоянки. При віддаленому догляді камері також не вистачає можливості відрізнити незвичні, але неважливі події від важливих. Все ж, необхідно зауважити, що така камера все ще може бути корисною як допомога для спостерігача-людини, хоча вона не може повністю її замінити.

Інший вплив відсутності “розуміння” виникає для активного агента, який визначає, коли клієнту потрібна допомога. З огляду на відповідні навчальні дані, агент міг би виявити видимі ознаки біди, такі як міміка чи голосові ознаки. Але система не може робити таких висновків без попереднього навчання і пристосування до особливості поведінки конкретного користувача, як це міг би зробити спостерігач-людина. Сама система визначає якість можливих реакцій, часто перевіряючи якість моделі ситуації, яку вона буде. Хороша модель зможе передбачити, що буде далі, або може дозволити реконструювати деталі ситуації за спрощеним поданням. Як уже

згадувалося, хоча ШІ-модель відрізняє предмети різного розміру, проте вона не має інформації про те, якими є ці об'єкти [16]. Вона також не трактує різних траєкторій, описаних у її моделі, наприклад, що деякі траєкторії малих предметів (пішоходів) ведуть до входу в будівлю.

рЗ огляду на відповідні навчальні дані, агент міг би виявити видимі ознаки біди, такі як міміка чи голосові ознаки. Але система не може робити таких висновків без попереднього навчання и пристосування до особливосте поведінки конкретного користувача, як це міг би зробити спостерігач-людина. Сама система визначає якість можливих реакцій, часто перевіряючи якість моделі ситуації, яку вона будує. Хороша модель зможе передбачити, що буде далі, або може дозволити реконструювати деталі ситуації за спрощеним поданням

Проблему розуміння можна проілюструвати на прикладі камери спостереження. Хоча камера може ідентифікувати незвичну поведінку, вона взагалі не розуміє того, що бачить. Як уже згадувалося, хоча ШІ-модель відрізняє предмети різного розміру, проте вона не має інформації про те, якими є ці об'єкти. Вона також не трактує різних траєкторій, описаних у її моделі, наприклад, що деякі траєкторії малих предметів (пішоходів) ведуть до входу в будівлю.

З проведеного дослідження можна зробити висновок, що застосування систем штучного інтелекту та методів машинного навчання є найефективнішим способом для підлатування програмних рішень під конкретні потреби користувачів з обмеженими можливостями.

2 ДОСЛІДЖЕННЯ СПОСОБІВ ТА МЕТОДІВ МАШИННОГО НАВЧАННЯ

2.1 Дослідження способів технології машинного навчання

Технологія машинного навчання передбачає тренування математичної моделі на історичних даних для того, щоб прогнозувати деяку подію або явище на основі нововведених даних. Машинне навчання досліджує вивчення та побудову алгоритмів, що мають змогу вчитися на введених даних і прогнозувати дані що ще не введені. Такі алгоритми працюють шляхом побудови математичної моделі на прикладі навчального набору вхідних даних та спостережень для того, щоб, керуючись введеними даними, приймати певні рішення або робити прогнози. Неодмінна риса таких рішень та прогнозів полягає в тому, що вони базуються на результаті обробки та навчання на введених даних, а не шляхом строгого слідування статичним програмним інструкціям.

Машинне навчання означає велику кількість математичних, статистичних і обчислювальних методів для розробки алгоритмів, здатних вирішити задачу не прямим способом, а на основі пошуку закономірностей у різноманітних вхідних даних. Рішення обчислюється не за чіткою формулою, а за встановленою залежністю результатів від конкретного набору ознак та їх значень.

На сьогоднішній день технологія машинного навчання широко розповсюджена у всьому світі. Завдяки своїй адаптивності під різноманітні вимоги, ця технологія використовується у різноманітних сферах людської діяльності – від медицини до банківської та охоронної діяльності.

Розрізняють такі способи машинного навчання:

— навчання зі вчителем (Supervised learning) – цей спосіб є оптимальним якщо відомо чому конкретно має навчитись програма. Алгоритм навчається на основі величезної кількості вхідних навчальних даних. Параметри даних треба змінювати до отримання необхідних вихідних параметрів. Після цього процесу можна перевірити чи успішно пройшло навчання наказавши програмі спрогнозувати результат для

контрольних даних які ще не були введені[18]. Найчастіше такий спосіб машинного навчання використовується для задач класифікації і прогнозування. Наприклад його можна використовувати для визначення фінансового ризику приватних осіб чи організацій на основі наявних відомостей про минулу фінансову діяльність, або при прогнозуванні споживацького попиту на певні товари на основі попередніх закономірностей;

— навчання без вчителя (Unsupervised learning) – суть цієї технології полягає в тому, що програма досліджує набір даних та виявляє приховані закономірності кореляції між різноманітними змінними. Цей спосіб можна використовувати для групування даних в кластери на основі одних лиш їх статистичних ознак. Хороше застосування навчання без вчителя - алгоритм кластеризації, який використовується для імовірнісного з'єднання записів. Визначаються зв'язки між елементами даних, і на підставі цих відносин виявляються зв'язки між людьми і організаціями в фізичному або віртуальному світі. Варіант особливо корисний компаніям, яким потрібно, наприклад, об'єднати дані з різнорідних джерел або по різних структурним підрозділам, щоб побудувати загальну картину клієнтури. Навчання без вчителя також можна використовувати для аналізу тональності висловлювань, щоб визначати емоційний стан людей на основі їх постів в соціальних мережах, повідомлень електронної пошти та інших записів. Сьогодні, зокрема, в компаніях, що спеціалізуються на фінансових послугах, за допомогою навчання без вчителя все частіше оцінюють рівень задоволеності клієнтів;

— навчання із частковим залученням вчителя (Semi-supervised learning) - це гібрид навчання з учителем і без. Розмітивши невелику частину даних, учитель дає машині зрозуміти, яким чином кластеризувати інше. Спосіб можна застосовувати для розпізнавання шахрайств зі спробами видати себе за іншого. Шахрайства можна класифікувати як аномалію на тлі звичайної активності. Методи машинного навчання з частковим залученням вчителя дозволяють створювати моделі, що розпізнають такі аномалії. Відповідні системи нерідко застосовуються для виявлення спроб шахрайств

при онлайн-угодах. Навчання з частковим залученням вчителя також можна застосовувати у випадках, коли є готові набори частково розмічених даних, що характерно для великих підприємств. В Amazon, наприклад, поліпшили можливості цифрового асистента Alexa розуміти природну мову, навчаючи алгоритми штучного інтелекту на сполученнях розмічених і нерозмічених даних. Завдяки цьому вдалося підвищити точність відповідей Alexa[19];

— навчання з підкріпленням (Reinforcement learning) – при навчанні з підкріпленням машині дозволяють взаємодіяти з оточенням (наприклад, скидати браковану продукцію з конвеєра в кошик) і «винагороджують», коли вона правильно виконує завдання. Автоматизувавши підрахунок винагород, можна дати можливість машині навчатися самостійно. Одне із застосувань навчання з підкріпленням - сортування товарів в роздрібних магазинах. Деякі компанії експериментують з роботизованими системами сортування предметів одягу, взуття та аксесуарів. Різновид цього способу, глибинне навчання з підкріпленням, добре підходить для автономного прийняття рішень у випадках, коли можливостей навчання з учителем і без недостатньо;

— глибинне навчання (Deep learning) – глибинне навчання може проходити як без вчителя, так і з підкріпленням. При глибинному навчанні частково імітуються принципи навчання людей - використовуються нейронні мережі для все більш докладного уточнення характеристик набору даних. Глибинні нейронні мережі застосовуються, зокрема, для прискорення скринінгу великих обсягів даних при пошуку лікарських засобів. Такі нейромережі здатні обробляти безліч зображень за короткий час і отримати більше ознак, які модель в кінцевому рахунку запам'ятовує. Цей спосіб теж широко застосовується для боротьби з шахрайствами, так як покращує точність розпізнавання завдяки автоматизації. Глибинне навчання може використовуватися в автомобільній галузі при виконанні ремонту і профілактичного обслуговування.

Штучний інтелект (ШІ) є скрізь. Одна з дуже успішних форм ШІ, моделі машинного навчання (ML), вже допомагають лікарям виявити меланому, рекрутерам знайти перспективних кандидатів та банкам вирішити, кому надавати позику. Вони використовуються в рекомендаціях щодо продуктів, цільовій рекламі, оцінці есе, просуванні та утриманні співробітників, маркуванні зображень, відеоспостереженні, самохідні машини та безліч інших додатків. Деякі з цих заявок мають великі наслідки для залучених осіб.

ШІ також трансформує спосіб взаємодії з машинами, оскільки вони вчаться перекладати наші слова в текст, інтерпретувати наші жести та розпізнавати нас та наші емоції. Розпізнавання мови досягло майже співвідношення з людською ефективністю деяких наборів даних (Xiong et al., 2017), а методи AI здатні ідентифікувати людей та предмети на фотографіях, відео або датчиках даних. Автономні транспортні засоби вже там. Моделі машинного навчання навчаються на вивченні великої кількості прикладів. Вони вивчають закономірності в навчальних даних, використовуючи їх для класифікації нових прикладів. Про потенціал цих моделей кодувати, продовжувати і навіть посилювати дискримінацію маргіналізованих груп у суспільстві. Дослідники виявили гендерну дискримінацію, расову дискримінацію та вікову дискримінацію в моделях ML.

Дискримінації щодо інвалідності не було досліджено в літературі на сьогоднішній день, але існує очевидна можливість упередженості щодо людей з обмеженими можливостями в системах ШІ. Як і вік, стать та раса, статус інвалідності є захищеною характеристикою. У Сполучених Штатах Америки Закон про американців з інвалідністю (ADA) 1990 р. Забороняє дискримінацію за ознакою інвалідності при працевлаштуванні, доступ до державних програм та послуг, громадський транспорт, громадський транспорт та телекомунікації. У Європі інвалідність визнана підставою для дискримінації, а Рамкова директива захищає людей з інвалідністю від дискримінації у сфері зайнятості та занять (Рада

Європейського Союзу, 2000 р.). Конвенцію ООН про права інвалідів (Генеральна Асамблея ООН, 2007 р.) ратифікували 177 країн.

У міру поширення моделей ШІ важливо, щоб вони виконували морально-правові зобов'язання суспільства щодо справедливого ставлення до громадян, особливо стосовно захищених груп, які в минулому зазнавали дискримінації. Додатки, засновані на ШІ, не повинні упереджувати будь-яку захищену групу. Однак досвід показав, що таке упередження може існувати і існує. Одним з відомих прикладів є модель COMPAS, яка використовується для прогнозування рецидиву злочину - ймовірність того, що в'язень, звільнений, скоює злочин. Порівнюючи фактичні вчинені злочини, було виявлено, що COMPAS має расовий ухил, при якому частіше помилково відмовляють у звільненні темношкірим, ніж білим (Angwin, Larson, Mattu, & Kirchner, 2016). Систематичне упередження може виникнути, якщо дані, що використовуються для підготовки моделі, містять упереджені людські рішення, а упередженість передається вивченій моделі. Наприклад, якщо вербувальники коледжів систематично пропускають заяви студентів з обмеженими можливостями, або медичний страховик регулярно відмовляє в покритті людям з обмеженими можливостями, модель, навчена за цими даними, повторюватиме таку саму поведінку.

Іншим джерелом упередженості може бути відсутність представлення даних у наборах даних. Наприклад, було встановлено, що у відомих систем аналітики обличчя рівень чорношкірих жінок набагато вищий, ніж у білих чоловіків (Buolamwini & Gebru, 2018). Коли цю проблему було виявлено, IBM швидко здійснила виправлення упередженості в своїй системі, збалансувавши навчальні дані, зменшивши невідповідність на порядок (Puri, 2018). Використання репрезентативних даних є важливим кроком до усунення потенційних упереджень щодо людей з інвалідністю. Бюро перепису населення США підраховує, що кожен п'ятий чоловік має певну форму інвалідності в США (Brault, 2012). Якщо набори даних включають цю сукупність, отримані моделі, швидше за все, будуть ефективними для них.

Нарешті, упередження може виникнути, коли справжня кількість відсотків не вимірюється безпосередньо, а інші дані використовуються як проксі (Eubanks, 2018; O'Neil, 2017). Eubanks (Eubanks, 2018) описує інструмент скринінгу графства Аллегані, систему прогнозування ризику жорстокого поводження з дітьми, яка широко використовує атрибути, пов'язані з використанням програм державної допомоги, як функції передбачення (Eubanks, 2018, p156). Ці заходи пов'язані з бідністю, що веде до упередженого ставлення до бідніших сімей у моделі. Подібна ситуація може легко виникнути щодо інвалідності. Наприклад, якщо час, необхідний для проходження онлайн-тестування, інтерпретується як відображення рівня кваліфікації учасника тестування, це призведе до втрати людей, які використовують допоміжні технології для доступу до тесту, особливо якщо тест не був повністю доступним.

Чесність для людей з обмеженими можливостями. Чи справедливість для людей з обмеженими можливостями відрізняється від справедливості щодо інших захищених груп на основі статі, раси, віку чи інших ознак? Ми стверджуємо, що в багатьох випадках це так. Є кілька вирішальних відмінностей між групами інвалідності та іншими захищеними групами.

Різноманітність та викиди. Найбільша різниця полягає у різноманітності, нюансі та динамічному характері самої інвалідності. Інвалідність - це не проста змінна з невеликою кількістю дискретних значень. Він має багато вимірів, і люди можуть зазнати численних вад. Ми визначаємо інвалідність відповідно до Конвенції ООН про права людей з інвалідністю (КПІ) як невідповідність між наявними інфраструктурами та потреби людини. Як описано в CPRD, інвалідність не є атрибутом людини, а «концепцією, що розвивається, і є результатом взаємодії між людьми з порушеннями та перешкодами щодо ставлення та навколишнього середовища, що заважає їх повній та ефективній участі у суспільстві нарівні з іншими ».

Умови, які можуть призвести до інвалідності, не тільки різноманітні, але різняться за інтенсивністю та впливом і часто змінюються з часом. Це різноманіття

також стосується груп, які спочатку можуть здатися однорідними. Популярна приказка в аутизмі: "Коли ви зустрічаєте одну людину з аутизмом, ви зустрічаєте одну людину з аутизмом". Замість формування згуртованої групи, спільнота інвалідів включає в себе багато випадків. Це створює виклик для машинного навчання, яке працює шляхом пошуку закономірностей та формування груп.

Дані, що перебувають поза межею, часто трактуються як „шум” і не враховуються. Навіть з урахуванням даних, у наборі даних для машини не може бути достатньо людей з даним типом та ступенем тяжкості інвалідності для ідентифікації шаблону. Включення чужих даних ускладнює навчальне завдання і веде до складніших моделей, які можуть перевершити дані. Перевагу надають більш простим моделям, і їх прогнози для цих «незвичних» осіб можуть бути неякісними або несправедливо негативними. Порівняно зі статтю, расою чи віком, непросто вирішити упереджені результати, зібравши збалансований набір навчальних даних, оскільки існує так багато форм та ступенів інвалідності.

Конфіденційність. Багато людей турбуються щодо конфіденційності щодо обміну інформацією про інвалідність. Люди з обмеженими можливостями відчують дискримінацію та відторгнення у повсякденному житті. В одному з недавніх польових досліджень, розкриття інвалідності (пошкодження спинного мозку або синдром Аспергера) у супровідному листі про роботу призводило до зменшення позитивних відповідей роботодавців на 26%, хоча інвалідність, ймовірно, не впливала на продуктивність роботи на цій посаді (Ameri et al ., 2018). Інше дослідження показало, що різницю в оплаті праці людей з інвалідністю та без неї не можна пояснити різницею в продуктивності праці, і, ймовірно, є дискримінацією (Kruse, Schur, Rogers, & Ameri, 2017). Особи справедливо бажають мати контроль над тим, що відомо про них і як ця інформація використовується. Як приклад, нещодавно з'ясувалося, що стандартизована організація тестування АСТ у США не тільки передавала інформацію про інвалідність про абітурієнтів у коледжі, але навіть продавала її третім особам (Jaschik, 2018). Студенти з обмеженими можливостями подають позов проти

АСТ за незаконне розкриття своїх інвалідностей до коледжів, до яких вони подають документи. Використання коледжами цієї інформації для прийняття рішень про вступ також було б незаконним згідно із Законом США про інваліди (ADA).

У Європі нові правила GDPR дають людям право знати, які дані про них зберігаються, і як вони використовуються, а також вимагати їх видалення. Оскільки організації рухаються до обмеження інформації, яку вони зберігають, і способів її використання, це означає, що системи ШІ часто можуть не мати явної інформації про інвалідність, яку можна використовувати для застосування встановлених тестів на справедливість та виправлень.

Алгоритмічна справедливість. Коли підходи ШІ використовуються для підтримки спілкування на основі мовлення, письма чи жестів, основним завданням справедливості є алгоритмічна справедливість. Моделі повинні працювати однаково добре членів різних груп. Наприклад, система розпізнавання мовлення повинна працювати так само добре, як для жіночої мови, так і для чоловічої. Залежно від програми, методи ШІ можуть бути неточними або просто зовсім не працювати для деяких людей, оскільки їх зовнішній вигляд, мова чи поведінка перебувають поза навчальними даними ШІ. Існують анекдотичні докази глухонімих, наприклад, що сучасні системи розпізнавання мови мають дуже високий рівень помилок для розуміння глухої мови. Люди, які мають порушення мовлення, часто не можуть користуватися такими системами.

Цей аспект справедливості можна покращити, збираючи дані про навчання від широкого кола груп, і забезпечуючи, щоб процес «очищення» даних зберігав достатню різноманітність. Однак можливо, що використання більш різноманітних навчальних даних може погіршити загальну ефективність деяких моделей. У таких випадках може знадобитися побудова спеціалізованих моделей для відомих груп, таких як розпізнавання глухої мови. Дослідникам слід вивчити методи правильної обробки даних, отриманих непередбаченими людьми та групами, які говорять, пишуть, дивляться або поведуться не так, як звичайна людина.

Обмежена справедливість. Що стосується алгоритмів, які розподіляють людей до улюблених або менш сприятливих груп, багато дослідників досліджували різні способи вимірювання та забезпечення справедливості при визначенні сприятливих результатів.

Якому визначенню справедливості слід віддати перевагу в просторі інвалідності? Тут ми розглядаємо це питання, досліджуючи деякі загальноживані підходи, а також їх сильні та слабкі сторони. Як запущений приклад, уявіть оголошення про роботу з 1000 претендентами, 50 з яких є інваліди. Усім претендентам пропонується пройти онлайн-тест на знання домену роботи, а їх результати, резюме та листи заявок використовуються системою рішень на основі ШІ, щоб запропонувати шорт-лист із 20 заявників для співбесіди.

Найпростіший підхід до справедливості - це «чесність через несвідомість», коли інформація про захищені ознаки (наприклад, стать, вік, інвалідність) не збирається та не використовується для прийняття рішень. Це може бути ефективним у деяких ситуаціях і узгоджується з деякою існуючою практикою. Наприклад, багато осіб, які шукають роботу, вирішують не розкривати інвалідність у початкових заявках на роботу, а Закон про американців з інвалідністю (ADA) спеціально забороняє майбутнім роботодавцям запитувати про інвалідність заявників.

Однак це не є гарантією справедливості. Він може зазнати невдачі, коли інвалідність заявника впливає на іншу інформацію, яка використовується у рішенні. Наприклад, якщо п'ять наших претендентів на роботу використовують допоміжні технології, такі як зчитувач з екрана чи лупу, а сам онлайн-тест не є повністю доступним, то довгий час відгуку може призвести до систематичного виключення цих п'яти претендентів із використанням допоміжних технологій, навіть якщо їх інвалідність невідома.

Без даних про інвалідність важко оцінити, чи є систематичною дискримінація. Як результат, дослідники ШІ виступають за „чесність через обізнаність” (Dwork et al.,

2012), де членство в захищеній групі явно відомо, а справедливість може бути офіційно визначена, перевірена та застосована до алгоритму.

Коли інформація про інвалідність доступна, з упередженнями даних можна вирішити чисельні методи, а вихід моделей можна скорегувати для зменшення упередженості. У деяких додатках інформація про результат може використовувати для оцінки точності прогнозів, зроблених моделями для різних. Набір інструментів AI Fairness 360 від IBM містить великий набір методів, застосовних до ситуацій із даними про результати та без них, пропонуючи цілий ряд різних заходів справедливості, які можна використовувати для виявлення упередженості та оцінки ефективності пом'якшення упередженості кроки.

Одним із способів виміряти справедливість за наявності інформації є групова справедливість або статистичний паритет.

Відповідно до цієї політики частка відібраних осіб у захищеній групі (людей з інвалідністю) повинна бути приблизно такою ж, як частка відібраних осіб у незахищеній групі. Іншими словами, 20 претендентів, відібраних для співбесіди, повинні включати одного заявника з інвалідністю. Це політика позитивних дій, але не гарантує, що будуть обрані найбільш кваліфіковані кандидати.

Групову справедливість може бути важко застосувати до інвалідності:

— покладається на наявність явної інформації про інвалідність. Як описано вище, ця інформація може бути надзвичайно конфіденційною, і через існуючий правовий захист вона може бути недоступною (як у нашому прикладі заявки на роботу);

— тестування на сукупному рівні може приховувати дискримінацію щодо окремих підгруп (наприклад, тих, хто використовує допоміжні технології для перевірки кваліфікації). З іншого боку, тестування всіх можливих підгруп інвалідності є недоцільним, і для цього рідко буде достатньо точок даних або детальної інформації про інвалідність;

— не може врахувати справжні відмінності між групами.

Альтернативним підходом є «індивідуальна справедливість», який визначає, що подібні особи повинні мати подібні результати. Цей підхід використовує метрику, щоб визначити, наскільки схожі дві особи щодо даного завдання (Dwork et al., 2012). Якщо можна визначити відповідний (і справедливий!) Спосіб вимірювання подібності, справедливість можна перевірити за допомогою цієї метрики. Однак такий підхід ставить проблему визначення метрики справедливої подібності, і виникає питання, чи і як інформація про інвалідність включена в такий показник.

Метрика справедливої подібності для прикладу заявки на роботу може використовувати показники відповідного досвіду та результати тестів, але не використовувати захищену інформацію, таку як статус інвалідності.

Таким чином, кандидати з однаковими оцінками та кваліфікацією повинні бути однаково ймовірними обрані, незалежно від того, мають вони інвалідність чи ні. Це має ту перевагу, що справедливе ставлення до будь-якої особи можна перевірити, не знаючи, чи є у них інвалідність, шляхом порівняння результатів для «подібних» заявників. Тестові набори, що представляють широкий спектр людей з обмеженими можливостями, можуть бути використані для перевірки справедливості в отриманих моделях. Завдання цього підходу полягає в тому, що кожна програма потребує власної чітко визначеної та справедливої метрики подібності. Упередження можна легко ввести до такої метрики, використовуючи атрибути, на які впливає інвалідність, наприклад, час проведення тесту.

Забезпечення справедливості. У доступі до технологій ми покладаємось на стандарти, що є посередником між різними потребами та кроками, які повинен зробити розробник. Розробникам рішень для машинного навчання може знадобитися подібна допомога, яка допоможе їм зрозуміти та оцінити вплив їхніх заявок на людей з обмеженими можливостями.

Щоб покращити справедливість для людей з обмеженими можливостями в рішеннях на основі ШІ, пропонується розробникам рішень, керованих ШІ, які використовують дані про людей.

Поміркувати над наслідками їх вирішення для людей з обмеженими можливостями, задаючи питання, серед яких: Хто представив би незвичні схеми даних, які можуть бути погано оброблені підходом машинного навчання, навченим типовим даним? Чи існує група інвалідності, яка зазнає серйозних змін? Включення членів таких груп у розробку та розробку додатків, а також збір даних про навчання спеціально від цих груп, де це можливо, може допомогти вирішити проблеми.

Використовувати існуючі методи для перевірки упередженості та пом'якшення упередженості по всьому конвеєру машинного навчання (у даних, моделях та результатах), коли доступна інформація про інвалідність. Наприклад, набір інструментів IBM AI Fairness 360 включає низку заходів справедливості та методи попередньої обробки даних для зменшення упередженості, навчання справедливих моделей та обробка вихідних даних моделі для покращення справедливості.

Поміркувати, чи не може інвалідність впливати і на інші змінні у вхідних даних. Поширеним прикладом може бути час, затрачений на виконання завдання. Для застосувань цього класу підхід «індивідуальної справедливості» може бути відповідним способом визначення та вимірювання справедливості.

Підходи до машинного навчання добре працюють для типових вхідних даних і погано справляються з викидами. Хоча деякі аспекти справедливості для людей з обмеженими можливостями можна було б покращити, збільшивши представленість у навчальних даних, завжди будуть люди, які не відповідають вимогам. Чесність для людей з обмеженими можливостями та надійність рішень для машинного навчання в реальних додатках може бути покращена шляхом розробки підходів, які розпізнають та обробляють викиди.

Нарешті, задокументувати послугу, щоб інші могли оцінити справедливість щодо людей з обмеженими можливостями. Сюди входить: чи були люди з інвалідністю включені до збору даних, проектування та тестування; результати перевірки справедливості; будь-які кроки, спрямовані на зменшення упередженості щодо людей з інвалідністю; і відомі обмеження послуги для конкретної групи

інвалідності.

Незважаючи на те, що в системах ШІ вже багато писано про гендерні, расові та вікові упередження, дотепер у дослідницькій літературі мало уваги приділяється потенціалу зменшення або закріплення дискримінації щодо інвалідності. Тревіран (Treviranus, 2017) висвітлює потенційні небезпеки покладання на підходи чистого ML для програм, які повинні працювати для всіх, а не лише для звичайної людини.

Типовий тренувальний процес для моделей ML оптимізує ефективність для типових випадків, за рахунок крайніх випадків, однак люди з обмеженими можливостями дуже різноманітні, і часто можуть фігурувати як дані крайніх випадків у даних.

Для систем, які прийматимуть або впливатимуть на рішення, що зачіпають людські життя, дуже важливо, щоб у розробці було залучено широке коло зацікавлених сторін користувачів, включаючи людей з обмеженими можливостями, які можуть допомогти розробникам продумати можливі наслідки технології та протестувати технологію. ефективність на крайніх випадках та недостатньо представлених групах населення.

В результаті ігор ми бачимо, як суспільство стає набагато більш розуміючим та інклюзивним для спільноти інвалідів, але з повідомленням ООН, що близько 15% світового населення страждає певним видом інвалідності, зростає відчуття, що нам потрібно зробити більше, щоб полегшити життя людей з обмеженими можливостями.

Наука даних та штучний інтелект вийшли на перший план технологій за останні кілька років, і кілька практиків ставляться до більш благодійних поглядів на життя, підтримуючи людей, які страждають як фізичними, так і психічними вадами.

Одним із напрямків, де машинне навчання відіграє визначну роль, є підтримка людей, які страждають на розлад аутистичного спектра (РАС), а це стан, який страждає приблизно у 1 на кожні 100 людей, і у чоловіків частіше діагностують цей стан ніж жінки. Це впливає на дітей приблизно у віці 3 років і призводить до того, що у дитини виникають труднощі з обробкою або залученням до людської взаємодії чи

емоцій, що може дуже ускладнити інтеграцію в групи інших дітей. Однак Лондонська лабораторія знань мала неймовірний успіх, познайомивши групу дітей з АСД із віртуальним автономним роботом на ім'я Енді. Експеримент показав, що діти охоче взаємодіяли з Енді, уважно слухаючи, задаючи питання та відповідаючи на них набагато вільніше, ніж якщо б це був дорослий чоловік.

Подібним чином, робот зі штучним інтелектом під назвою Майло був розповсюджений у понад 50 школах США, щоб заохотити дітей з РАС до більш легкої взаємодії віч-на-віч, а в деяких випадках навіть дозволити фізичний контакт робота, що немислимо для більшості хворих на РАС. Хоча в даний час не існує визнаного ліки від РАС, експеримент дає надію сім'ям дітей та дорослих, які страждають на РАС, на те, що колись їм буде так зручно спілкуватися із програмним забезпеченням для штучного інтелекту, щоб вони могли збільшити впевненість у своїй взаємодії з іншими Люди.

Інтелектуальне протезування - ще одна сфера, в якій Data Science впроваджено для того, щоб полегшити життя інвалідів. Для багатьох ампутованих або людей, які живуть без використання однієї або декількох кінцівок, можливості дуже обмежені і часто дуже елементарні, зосереджуючись на голому, а не на повсякденному використанні. Однак компанія Deka Research, очолювана винахідником Segway Діном Каменом, вперше створила протез, який вони називають рукою «Люк» (названа в честь Люка Скайуокера). Унікальний аспект цієї технології полягає в тому, що вона більш точно змодельована на внутрішній роботі людської руки, з механічним еквівалентом сухожилів, м'язів та кісток, що дозволяє користувачеві набагато природніший діапазон рухів, ближчий до людської руки, ніж до гачків і пазурів, які були більш поширеними протягом останніх 20 років.

Частина технології, яку на популярність вивів один із найвідоміших учених у світі, професор Стівен Хокінг, - електронна доповнювальна та альтернативна система зв'язку (ААС) стала невід'ємною частиною життя людей, які не можуть говорити через такі умови. як хвороба моторних нейронів або церебральний параліч.

Самою рукою можна керувати кількома способами: за допомогою мікроскопічних нервових закінчень, прикріплених до основи руки, або контролерів у взутті користувача. Проект отримав значний фінансовий поштовх з боку Служби досліджень армії США, проте виготовлення протезів, як і раніше, буде надзвичайно дорогим у масовому масштабі, за деякими оцінками, приблизно в 50 000 доларів. Хоча це може бути не відповіддю для всіх, хто потребує протезування, але діапазон рухів та вдосконаленість представлених тут технологій є блискучим знаком для майбутнього, коли подібні технології будуть доступні за менші витрати та, можливо, з ще більшим використанням.

Частина технології, яку на популярність вивів один із найвідоміших учених у світі, професор Стівен Хокінг, - електронна доповнювальна та альтернативна система зв'язку (ААС) стала невід'ємною частиною життя людей, які не можуть говорити через такі умови, як хвороба моторних нейронів або церебральний параліч.

Одним з останніх нововведень на цьому ринку, який досі тісно пов'язаний із пристроєм «Еквалайзер», який використовував Хокінг з 1986 року, є DynaVox EyeMax. Цей пристрій використовує прийоми Computer Vision через фронтальну камеру для відстеження рухів очей користувача на екрані команд, і навіть може бути запрограмований на використання певної інтонації, щоб мова не сприймалася неправильно. Кожен пристрій повинен бути запрограмований індивідуально для відповідного користувача, включати визначні місця, імена друзів та сім'ї та іншу унікальну інформацію, а вартість пристроїв не часто покривається медичною страховкою. Однак очікується, що включення технологій глибокого навчання до пристроїв у майбутньому має скоротити час, необхідний для програмування пристрою, що може знизити вартість пристроїв до більш доступного рівня.

Це не лише спеціалізовані медичні фірми, котрі бачать лише фінансову цінність, але й благодійні фонди, що використовують описані технології для допомоги при вирішенні багатьох проблем людей з особливими потребами.

2.2 Дослідження методів машинного навчання

Розрізняють наступні методи машинного навчання:

- регресія;
- класифікація;
- кластеризація;
- зменшення розмірності;
- ансамблеві методи;
- нейронні мережі та глибоке навчання;
- трансферне навчання;
- навчання з підкріпленням;
- обробка природної мови;
- вкладення слів;

Регресія. Методи регресії відносяться до категорії навчання зі вчителем. Вони допомагають передбачити або пояснити конкретне числове значення на основі набору попередніх даних, наприклад, прогнозування ціни на нерухомість на основі попередніх даних про ціни на подібні властивості [20].

Найпростіший метод - це лінійна регресія, де ми використовуємо математичне рівняння прямої ($y = m * x + b$) для моделювання набору даних. Ми тренуємо модель лінійної регресії з великою кількістю пар даних (x, y), обчислюючи положення та нахил лінії, яка мінімізує загальну відстань між усіма точками даних та лінією. Іншими словами, ми обчислюємо нахил (m) і перетин y (b) для лінії, яка найкраще наближує спостереження в даних.

Давайте розглянемо більш конкретний приклад лінійної регресії. Якщо застосувати лінійну регресію для прогнозування споживання енергії (y кВт-год) певних будівель, збираючи разом вік будівлі, кількість поверхів, площу та кількість вбудованого обладнання. Принцип є таким самим, як і проста лінійна регресія, але в

цьому випадку створена "лінія" знаходиться у багатовимірному просторі, що базується на основі кількості змінних. Методи регресії відносяться до категорії навчання зі вчителем. Вони допомагають передбачити або пояснити конкретне числове значення на основі набору попередніх даних, наприклад, прогнозування ціни на нерухомість на основі попередніх даних про ціни на подібні властивості.

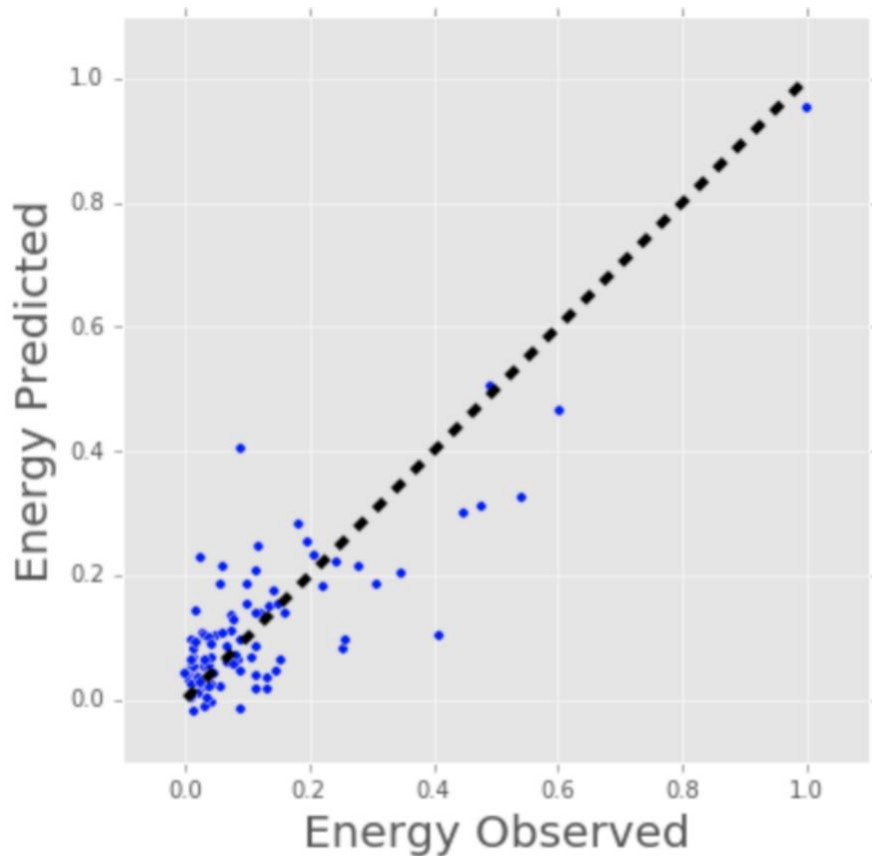


Рисунок 2.1 – Графічне представлення методу лінійної регресії

Наведений вище графік (рис.1) показує, наскільки модель лінійної регресії відповідає фактичному енергоспоживанню будівлі. А тепер уявіть, що у вас є доступ до характеристик будівлі (вік, площа тощо), але ви не знаєте, як споживається енергія. У цьому випадку ми можемо використовувати встановлену лінію для приблизного споживання енергії конкретної будівлі.

Класифікація. Інший клас навчання зі вчителем - методи класифікації, що передбачають або пояснюють значення класу. Але методи класифікації не обмежуються двома класами. Наприклад, метод класифікації може допомогти оцінити, чи містить дане зображення автомобіль чи вантажівку. У цьому випадку на виході будуть 3 різні значення:

- зображення містить автомобіль;
- зображення містить вантажівку;
- зображення не містить ні автомобіля, ні вантажівки.

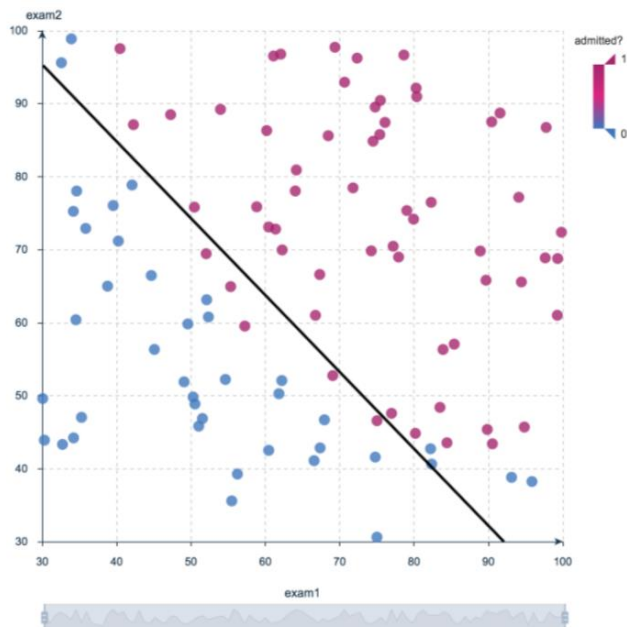


Рисунок 2.2 – Графічне представлення методу класифікації

У представленому вище рис. 2.2 наведені оцінки попередніх студентів разом із тим, чи були вони прийняті. Логістична регресія дозволяє нам провести лінію, яка представляє межу рішення. Метод класифікації може допомогти оцінити, чи містить дане зображення автомобіль чи вантажівку. Наприклад, вони можуть допомогти передбачити, чи купуватиме продукт онлайн-клієнт. Результатом може бути так чи ні: покупець чи ні покупець. Найпростіший алгоритм класифікації - це логістична

регресія, яка звучить як метод регресії, але це не так. Логістична регресія оцінює ймовірність настання події на основі сукупності даних. Але методи класифікації не обмежуються двома класами [21], [22].

Логістична регресія є найпростішою моделлю класифікації. Окрім неї існують також нелінійні класифікатори, такі як дерева рішень, випадкові ліси, опорні векторні машини та нейронні мережі. Наприклад, логістична регресія може взяти в якості вхідних даних два результати іспитів студента, щоб оцінити ймовірність того, що студент потрапить до певного коледжу. Оскільки оцінка є ймовірністю, результат є числом від 0 до 1, де 1 означає повну впевненість. Для студента, якщо передбачувана ймовірність перевищує 0,5, тоді ми прогнозуємо, що він чи вона будуть прийняті. Якщо передбачувана ймовірність менше 0,5, ми прогнозуємо, що йому буде відмовлено.

Кластеризація. Методи кластеризації відносяться до категорії навчання без вчителя, оскільки їх метою є групування або кластеризація спостережень, що мають спільні характеристики.

Алгоритм методу “К-середніх”, працює так: випадково вибирає К-центри серед існуючих даних після цього відносить кожен одиницю даних до найближчого із випадково створених центрів. Потім повторно обчислює центр кожного кластера. Якщо центри не змінюються (або змінюються дуже мало), процес закінчується. В іншому випадку ми повертаємося до кроку 2. (Щоб запобігти потраплянню в нескінченний цикл, якщо центри продовжують змінюватися, заздалегідь встановити максимальну кількість ітерацій.)

Найпопулярнішим методом кластеризації є метод “К-середніх”, де “К” означає кількість кластерів, які користувач вирішив створити. Методи кластеризації не використовують вихідні дані для навчання, а натомість дозволяють алгоритму визначати вихідні дані самостійно. У методах кластеризації ми можемо використовувати лише візуалізації для перевірки якості рішення.

Досліджуючи кластеризацію, можна зустріти дуже корисні алгоритми, такі як просторова кластеризація для додатків із шумами, заснована на щільності (DBSCAN), кластеризація зсувів середнього значення, ієрархічна кластеризація, кластеризація методом EM (Expectation-Maximization), використовуючи серед інших моделі Гаусової суміші.

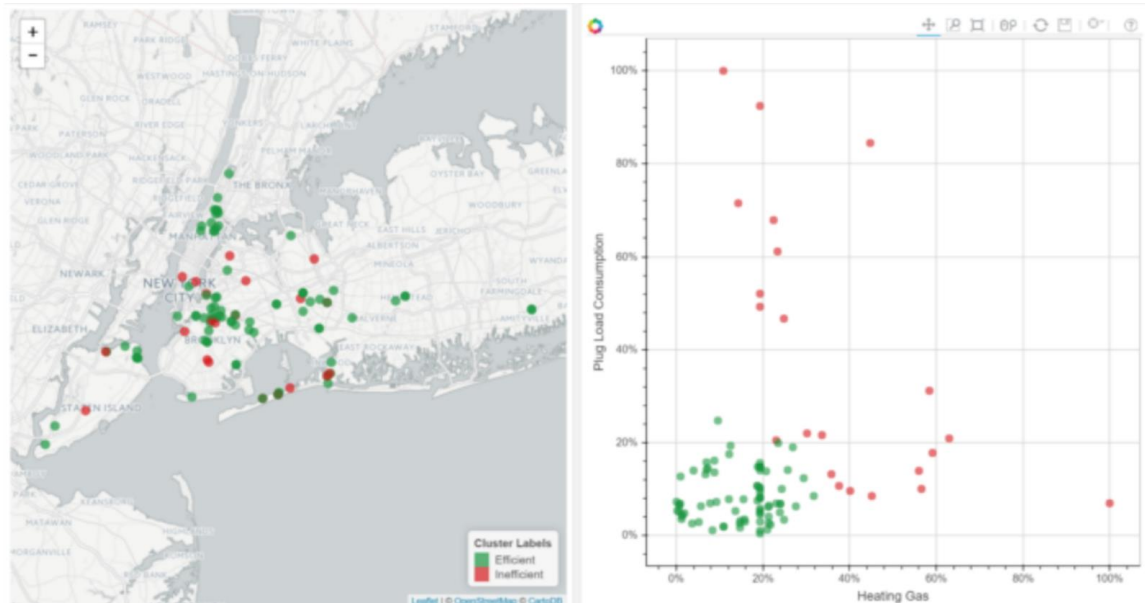


Рисунок 2.3 – Графічне представлення методу кластеризації

Для прикладу, на наведеному графіку (рис 2.3) застосовується метод “К-середніх” для набору даних з будівель. Кожна колонка на ділянці показує ефективність для кожної будівлі. Чотири вимірювання стосуються кондиціонування повітря, підключеного електрообладнання (мікрохвильові печі, холодильники тощо ...) побутового газу та опалювальних систем [24].

У цьому випадку параметр “К” = 2 для кластеризації, що дозволяє легко інтерпретувати один із кластерів як групу ефективних будівель, а інший кластер як групу неефективних будівель. Зліва наведено розташування будівель, а праворуч - два з чотирьох вимірів, які використовувались як вхідні дані: підключене обладнання та опалювальні системи.

Досліджуючи кластеризацію, можна зустріти дуже корисні алгоритми, такі як просторова кластеризація для додатків із шумами, заснована на щільності (DBSCAN), кластеризація зсувів середнього значення, ієрархічна кластеризація, кластеризація методом EM (Expectation-Maximization), використовуючи серед інших моделі Гаусової суміші [22].

Зменшення розмірності (Dimensionality Reduction). Як випливає з назви, ми використовуємо зменшення розмірності, щоб видалити найменш важливу інформацію (іноді надлишкові колонки) з набору даних. На практиці, часто трапляються набори даних із сотнями або навіть тисячами стовпців, тому зменшення загальної кількості є життєво важливим. Наприклад, зображення можуть містити тисячі пікселів, не всі з яких мають значення для вашого аналізу. Або під час тестування мікросхем у процесі виробництва у вас можуть бути тисячі вимірювань та тестів, застосованих до кожного чіпу, багато з яких надають надлишкову інформацію. У цих випадках вам потрібні алгоритми зменшення розмірності, щоб зробити набір даних зручним для опрацювання.

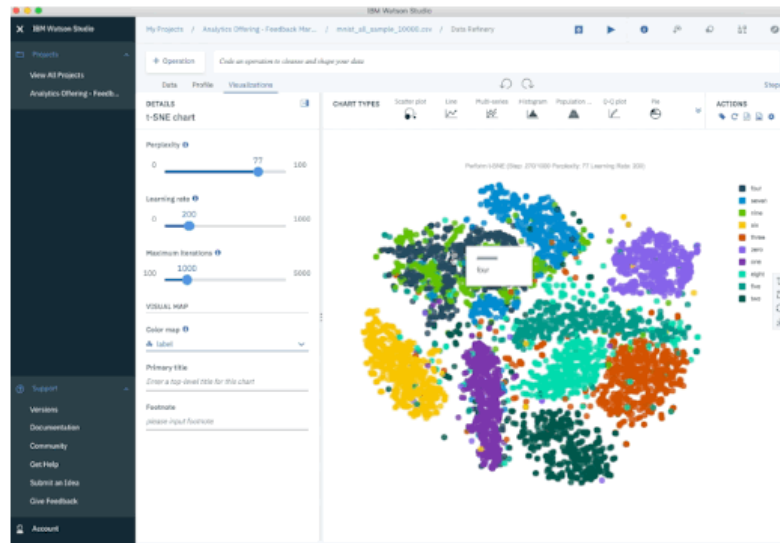
Найпопулярнішим методом зменшення розмірності є Аналіз основних компонентів (PCA), який зменшує розмірність простору ознак шляхом пошуку нових векторів, які максимізують лінійну варіацію даних. PCA- метод може різко зменшити розмір даних і не втратити занадто багато інформації, якщо лінійні кореляції даних сильні. Також можна виміряти фактичний ступінь втрати інформації та відповідно скорегувати алгоритм.

Іншим популярним методом є t-стохастичне вбудовування сусідів (t-SNE), який здійснює нелінійне зменшення розмірності. Зазвичай такий метод використовують для візуалізації даних, але він також підходить для завдань машинного навчання, серед яких - зменшення простору функцій та кластеризації.

Наступний графік (рис 2.4) показує аналіз бази даних рукописних цифр "MNIST". "MNIST" містить тисячі зображень цифр від 0 до 9, які дослідники використовують для перевірки своїх алгоритмів кластеризації та класифікації. Кожен

рядок набору даних - це векторизована версія оригінального зображення (розмір $28 \times 28 = 784$ пікселі) та мітка для кожного зображення (нуль, один, два, три,..., дев'ять). В цьому випадку ми зменшуємо розмірність із 784 (пікселі) до 2 (виміри даної візуалізації).

Проектування набору даних на двовимірному просторі дозволяє нам візуалізувати оригінальний набір даних досить великого розміру.



t-SNE Iterations on MNIST Database of Handwritten Digits.

Рисунок 2.4 – Графічне представлення роботи методів кластеризації та класифікації

Ансамблеві методи (Ensemble methods). Ансамблеві методи використовують ідею поєднання кількох прогнозованих моделей (машинне навчання зі вчителем), щоб отримати прогнози вищої якості, ніж кожна з моделей могла б надати самостійно. Наприклад, алгоритми випадкового лісу - це ансамблевий метод, який поєднує багато “натренованих” дерев рішень з різними зразками наборів даних. Як результат, якість прогнозів випадкового лісу вища, ніж якість прогнозів, здобутих за допомогою одного дерева рішень.

Ансамблеві методи – це хороший спосіб для зменшення дисперсії(похибки) та

упередженості однієї моделі машинного навчання. Це важливо, оскільки будь-яка з оглянутих моделей може бути точною за певних умов, але неточною за інших. Поєднуючи дві різні моделі, якість прогнозів буде більш високою та збалансованою.

На відміну від лінійних та логістичних регресій, які розглядаються як лінійні моделі, метою нейронних мереж є фіксація нелінійних шаблонів у даних шляхом додавання шарів параметрів до моделі. На зображенні нижче (рис 2.5) проста нейронна мережа має три входи, один прихований шар з п'ятьма параметрами та вихідний шар.

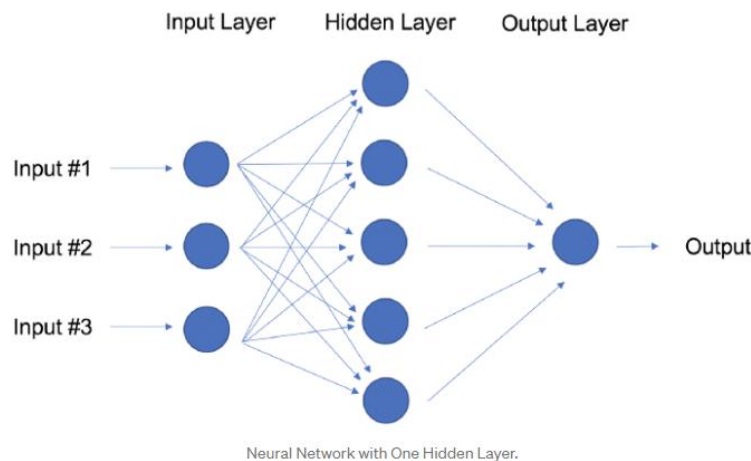


Рисунок 2.5 – Графічне представлення побудови нейронної сітки

Насправді структура нейронних мереж є досить гнучкою для того, щоб побудувати, вже знайомі, лінійну та логістичну регресію. Термін “Глибинне Навчання” походить від нейронної мережі з безліччю прихованих шарів і охоплює широкий спектр архітектур [24], [25].

Переважає більшість найкращих переможців конкурсів “Kaggle” – платформи для змагань з моделювання даних, використовує певні ансамблеві методи. Найпопулярнішими прикладами такого методу є Random Forest, XGBoost та LightGBM. Це важливо, оскільки будь-яка з оглянутих моделей може бути точною за певних умов, але неточною за інших. Зокрема, методи глибинного навчання

надзвичайно успішні у сферах візуальних даних (класифікація зображень), тексту, аудіо та відео. Найпоширенішими програмними рішеннями для застосування методу глибинного навчання є “Tensorflow” та “PyTorch”.

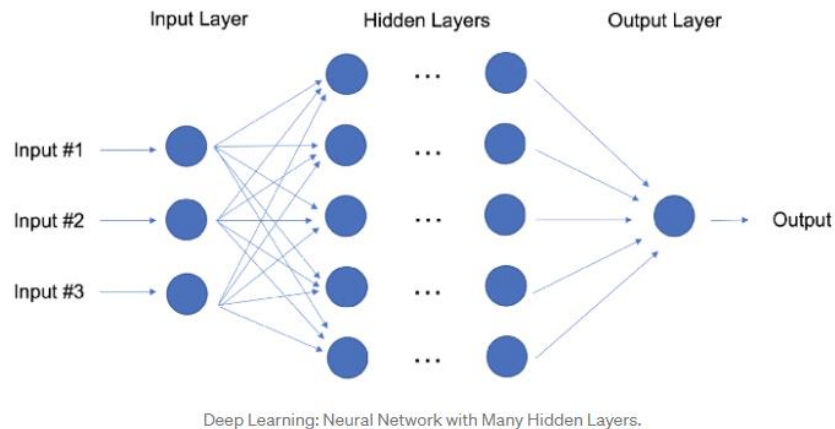


Рисунок 2.6 – Графічне представлення прихованих шарів нейронної сітки

Для найкращої продуктивності методи глибокого навчання вимагають багато даних - і багато обчислювальної потужності, оскільки метод самостійно налаштовує багато параметрів у величезних архітектурах. З наведеного стає зрозумілим, що фахівцям, що працюють із методом глибинного навчання потрібні дуже потужні комп'ютери із вдосконаленими графічними процесорами (GPU).

Особливо важко йти в ногу з розвитком технологій глибинного навчання, зокрема, тому, що дослідницькі та галузеві спільноти дуже швидко розвивають впровадження технологій глибинного навчання, щодня породжуючи цілі нові методології.

Трансферне навчання. Трансферне навчання означає повторне використання частини навченої раніше нейронної мережі та пристосування її до нового, але подібного завдання. Зокрема, як тільки ви тренуєте нейронну мережу, використовуючи дані для завдання, ви можете перенести частину тренуваних шарів нейронної мережі та об'єднати їх з кількома новими шарами, які ви можете тренувати,

використовуючи дані нового завдання. Додавши кілька шарів, нова нейронна мережа може швидко навчитися і швидко адаптуватися до нового завдання.

Головною перевагою методу трансферного навчання є те, що він потребує менше даних для тренування нейронної мережі, що особливо важливо, оскільки “тренування” алгоритмів глибокого навчання є доволі затратним як з точки зору часу, так і грошових ресурсів. Також, часто дуже важко знайти достатньо маркованих даних для навчання.

Наприклад, якщо для моделі сорочки використовується нейронна сітка з 20 прихованими шарами, то провівши кілька експериментів, дійдемо до висновку, що можливо перенести 18 шарів моделі сорочки та поєднати їх з одним новим шаром параметрів, щоб потренувати нейронну мережу на зображеннях штанів. У цьому випадку, модель штанів мала б 19 прихованих шарів. Вхідні та вихідні дані двох завдань різні, але багаторазові шари можуть узагальнювати інформацію, яка стосується обох, наприклад аспекти тканини.

Трансферне навчання стає дедалі популярнішим, і зараз існує безліч надійних попередньо навчених моделей для загальних завдань глибокого навчання, таких як класифікація зображень та тексту.

Навчання з підкріпленням (Reinforcement Learning). Для кращого розуміння даного методу розглянемо експеримент, де миша в лабіринті намагається знайти заховані шматочки сиру. Чим більше разів миша проходить лабіринт, тим краще вона стає у пошуку сиру. Спочатку миша може рухатися хаотично, але через деякий час досвід миші допомагає їй зрозуміти, які дії наближають її до сиру. Наведений приклад ілюструє, метод навчання з підкріпленням (Reinforcement Learning) для тренування певної системи. Загалом, навчання з підкріпленням – це метод машинного навчання, який допомагає агенту вчитися на досвіді. Записуючи дії та використовуючи підхід проб і помилок у заданому середовищі, такий метод може максимізувати сукупну винагороду. У нашому прикладі миша є агентом, а лабіринт - середовищем. Набір можливих дій для миші: рух вперед, назад, вліво або вправо. Нагорода - сир.

Однією з переваг RL – методу є те можливість застосовувати його за практичної відсутності історичних даних про певну проблему, оскільки цей метод не потребує введення інформації заздалегідь (на відміну від традиційних методів машинного навчання). Не дивно, що RL- метод є дуже успішним в ігровій індустрії, особливо в іграх «досконалої інформації», таких як, наприклад, шахи. Під час гри відгуки від агента та оточення надходять швидко, що дозволяє моделі швидко вчитися. Недоліком RL є те, що тренування може тривати дуже довго, якщо проблема складна.

Подібно до того, як ЕОМ (електронно обчислювальна машина) “Deer Blue” перемогла найкращого шахіста в 1997 році, “AlphaGo” - алгоритм, що базується на RL методі, переміг найкращого гравця у одну з найпопулярніших комп’ютерних ігор сучасності “Counter Strike: Go” у 2016 році. Нинішніми піонерами застосування RL- методів є команда “DeepMind” у Великобританії.

У квітні 2019 року команда “OpenAI Five” була першою системою штучного інтелекту (ШІ), яка перемогла команду чемпіонів світу з кіберспорту. Та сама команда ШІ, яка перемогла людську команду кіберспортсменів, також розробила робототехнічну руку, яка може виконувати різноманітні завдання.

З наведеного, можна стверджувати, що метод Навчання з підкріпленням (Reinforcement Learning) є особливо потужною формою ШІ, що стрімко розвивається і може знайти застосування практично у будь-якій сфері людської діяльності.

Обробка природної мови (Natural Language Processing). Величезний відсоток світових даних та знань передаються, в тій чи іншій формі, людською мовою. Очевидно, що комп’ютери ще не можуть повністю зрозуміти людський текст, але ми можемо навчити їх виконувати певні завдання. Наприклад, ми можемо навчити наші смартфони самозавершувати текстові повідомлення або виправляти неправильне написання слів. Ми навіть можемо навчити машину вести просту розмову з людиною.

Обробка природної мови (NLP) не є методом машинного навчання як такого, але є досить широко застосовуваною технікою підготовки тексту до машинного навчання. Наприклад, якщо розглянути велику кількість текстових документів у

різних форматах (word, Інтернет-блоги ...), більшість із них міститиме друкарські помилки, відсутність (пропуск) символів та інші слова, які потрібно відфільтрувати. На даний момент найпопулярнішим пакетом програмних рішень для обробки тексту є “NLTK” (Natural Language ToolKit), створений дослідниками зі Стенфорда.

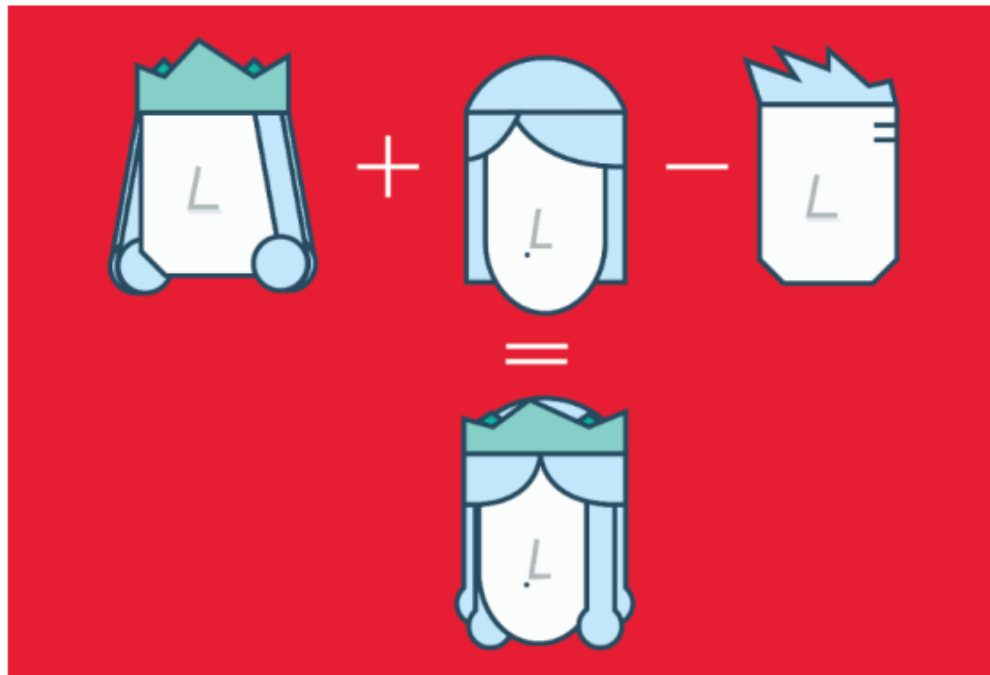
Найпростіший спосіб представлення тексту в числовому форматі подання - це обчислення частоту кожного слова в кожному текстовому документі. Уявімо матрицю цілих чисел, де кожен рядок представляє текстовий документ, а кожен стовпець - слово. Таке матричне подання частоти слів зазвичай називають частотною матрицею термінів (Term Frequency Matrix). Окрім цього, існує ще одне популярне матричне представлення текстового документа. Воно полягає у розділенні кожного вводу тексту на матриці, взявши до уваги те, наскільки важливим є кожне слово в усьому обсязі документів. Цей метод називається “Term Frequency Inverse Document Frequency” (TFIDF), і він, як правило, краще підходить для завдань машинного навчання.

Вкладення слів (Word Embeddings). Розглянуті вище методи “Term Frequency Matrix” (TFM) та “Term Frequency Inverse Document Frequency” (TFIDF) - це числові представлення текстових документів, які враховують лише частоту важливих слів для представлення текстових документів. З іншого боку, “вкладені” слова можуть фіксувати контекст слова в документі. За допомогою контексту вкладення слів може кількісно визначити подібність між словами, що, у свою чергу, дозволяє робити обчислення зі словами.

Word2Vec - це метод, заснований на нейронних мережах, який відображає сукупність слів у вигляді числового вектору. Потім ці вектори можуть використовуватися для пошуку синонімів, виконання арифметичних операцій зі словами або для представлення текстових документів (використовуючи середнє значення всіх векторів, побудованих на сукупності слів у документі).

Наприклад, використовується досить великий масив текстових документів для оцінки вбудовування слів. Припустимо також, що слова “король” (king), “королева” (queen), “чоловік” (man) і “жінка” (woman) є частиною такого масиву. Припустимо,

що вектор («слово») - це числовий вектор, який представляє слово «слово». Таким чином, для того, щоб оцінити вектор («жінка»), ми можемо виконати арифметичну операцію з векторами (рис. 2.6): $\text{vector}(\text{'king'}) + \text{vector}(\text{'woman'}) - \text{vector}(\text{'man'}) \sim \text{vector}(\text{'queen'})$.



Arithmetic with Word (Vectors) Embeddings.

Рисунок 2.6 – Графічне представлення методу Word2Vec

Окрім цього, існує ще одне популярне матричне представлення текстового документа. Воно полягає у розділенні кожного вводу тексту на матриці, взявши до уваги те, наскільки важливим є кожне слово в усьому обсязі документів.

Представлення слів дозволяє знаходити подібність між словами шляхом обчислення косинусної подібності між векторним представленнями двох слів. Подібність косинуса вимірює кут між двома векторами. Наприклад, припустимо, ми маємо доступ до твітів кількох тисяч користувачів Twitter. Також припустимо, що ми знаємо, хто з цих користувачів Twitter купив будинок. Щоб передбачити ймовірність придбання будинку новим користувачем Twitter, ми можемо поєднати метод

“Word2Vec” з логістичною регресією. Word2Vec - це метод, заснований на нейронних мережах, який відображає сукупність слів у вигляді числового вектору. Потім ці вектори можуть використовуватися для пошуку синонімів, виконання арифметичних операцій зі словами або для представлення текстових документів (використовуючи середнє значення всіх векторів, побудованих на сукупності слів у документі). Представлення слів дозволяє знаходити подібність між словами шляхом обчислення косинусної подібності між векторним представленнями двох слів. Подібність косинуса вимірює кут між двома векторами. Наприклад, припустимо, ми маємо доступ до твітів кількох тисяч користувачів Twitter.

Таке матричне подання частоти слів зазвичай називають частотною матрицею термінів (Term Frequency Matrix).

Окрім цього, існує ще одне популярне матричне представлення текстового документа. Воно полягає у розділенні кожного вводу тексту на матриці, взявши до уваги те, наскільки важливим є кожне слово в усьому обсязі документів.

Зазвичай вбудовування слів обчислюється за допомогою методів машинного навчання, але це часто є попереднім кроком до застосування алгоритму машинного навчання зверху.

Існує можливість як тренувати вбудовування слів самостійно, так і отримати заздалегідь підготовлений (за допомогою методу трансфертного навчання) набір векторів слів. Однією з найбільших бібліотек, що містить попередньо навчені вектори слів на 157 різних мовах є бібліотека “FastText”.

3 ЗАСТОСУВАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ПРИ РОЗРОБЦІ ДОДАТКУ ДЛЯ ЛЮДЕЙ З ОСОБЛИВИМИ ПОТРЕБАМИ

3.1 Розробка мобільного додатку - перекладача для людей з особливими потребами

У даному розділі буде описуватися створення мобільного додатку-перекладача для людей з особливими потребами та вдосконалення його інтерфейсу за допомогою методів машинного навчання.

Мобільний додаток - це комп'ютерна програма або програмне забезпечення, призначене для роботи на мобільних пристроях, таких як телефон, планшет або годинник. Від самого початку, такі програми були призначені для елементарних завдань, таких як робота з електронною поштою, календарем та базою контактних даних. Але суспільний попит на такі програми викликав швидке розширення і інтеграцію їх в інші області, такі як мобільні ігри, автоматизація виробництва, робота з GPS і послуги на основі місцезнаходження, відстеження замовлень і здійснення покупок. Таким чином сформувався величезний ринок мобільних додатків, що зараз налічує мільйони програм.

Ці програми, як правило, завантажуються з платформ розповсюдження додатків, якими керує власник певної мобільної ОС. Наприклад, магазин додатків "App Store" (ОС "iOS") або "Google Play Store" (ОС "Android"). Деякі програми є безкоштовними, а інші мають ціну, при цьому прибуток розподіляється між розробником програми та платформою її розповсюдження. Мобільні програми часто сильно відрізняються від десктопних додатків, призначених для роботи на настільних комп'ютерах, і веб-додатків, які запускаються у мобільних веб-браузерах, а не безпосередньо на мобільних пристроях. Люди з обмеженими можливостями також користуються смартфонами. Проте для них це не завжди легко і зручно, тому існують спеціалізовані додатки для людей з особливими потребами, аби полегшити їм роботу

із різноманітними програмами та сервісами, а також посприяти легшому контактуванню з навколишнім середовищем.

Мобільні програми часто сильно відрізняються від десктопних додатків, призначених для роботи на настільних комп'ютерах, і веб-додатків, які запускаються у мобільних веб-браузерах, а не безпосередньо на мобільних пристроях. Розроблюваний мною додаток має поєднати в собі ці задачі – полегшити людям з вадами зору користування сервісом перекладача, та спростити їх комунікацію із іншими людьми за допомогою систем голосового введення та речового синтезатору.

Також на меті є розробка адаптивного інтерфейсу із залученням методів машинного навчання. Отже, для того, аби вибрати оптимальне рішення щодо використання певних технічних засобів, та середовища розробки, необхідно оглянути наявні варіанти, і виявити їх переваги та недоліки.

Загалом розрізняють 3 типи мобільних додатків:

- “рідні” додатки (Native apps);
- гібридні додатки (Hybrid apps);
- веб-додатки (Web apps).

“Рідні” додатки (Native apps) додатки розроблені виключно для однієї мобільної ОС, тому вони є «рідними» для певної платформи або пристрою. Додаток, розроблений для таких систем, як iOS, Android, Windows, Symbian, Blackberry, не можна використовувати на жодній платформі, окрім власної. Іншими словами, неможливо використовувати додаток для Android на iPhone. Мови написання:

- iOS на Objective-C or Swift;
- android на Java;
- windows phone на Net.

Головною перевагою “рідних” додатків є їх висока продуктивність і зручність в користуванні, оскільки розробники використовують вже знайомий для власника певної ОС користувальницький інтерфейс. Крім того, доступ до широкого спектру Прикладних програмних інтерфейсів (Application Programming Interface, далі – API),

що робить використання програм максимально нелімітованим. “Рідні” додатки є легкодоступними на платформах розповсюдження (магазинах) додатків свого типу, що залежить від ОС і мають чітку тенденцію до задоволення цільових клієнтських запитів та очікувань.

Одним із мінусів “рідних” мобільних додатків є їх досить висока ціна порівняно з іншими типами додатків. Це обумовлено необхідністю створення дублікатів додатків для інших ОС, їх окремої підтримки та обслуговування, що призводить до більшої ціни продукту.

Веб-додатки (Web apps) поводять себе подібно до “рідних” додатків. Веб-додатки використовують веб-переглядач для запуску та зазвичай розроблені на HTML5, JavaScript або CSS. Ці додатки перенаправляють користувача на URL-адресу де пропонують опцію "встановити", просто створивши закладку на своїй сторінці.

Веб-додатки, як правило, потребують мінімального обсягу пам'яті пристрою. Оскільки всі персональні бази даних зберігаються на сервері, користувачі можуть отримати доступ до них з будь-якого пристрою, коли є підключення до інтернету. Ось чому використання веб-додатків з поганим підключенням часто призводить до некоректної роботи веб-додатків, та розчаруванні у них користувачів. Серйозним недоліком є доступ до дуже обмеженої кількості API для розробників, за винятком геолокації та декількох інших.

Гібридні додатки (Hybrid apps) розроблені з використанням мультиплатформних веб-технологій (наприклад, HTML5, CSS і Javascript). Так звані гібридні додатки є в основному, веб-додатками, замаскованими в обгортку “рідних” додатків [8]. Програми мають відповідні плюси і мінуси, які характерні для “рідних” та веб-додатків. Перевагою гібридних мультиплатформних програм є їх відносно швидка і не дуже складна розробка.

З іншого боку, гібридні додатки поступаються у продуктивності, швидкості і загальній оптимізованості в порівнянні, наприклад, із “рідними” програмами. Крім

того, існують певні недоліки у питаннях оформлення та дизайну, пов'язані з неможливістю додатку виглядати однаково на різних ОС-платформах.

Існує приблизно 33 категорії мобільних додатків у магазині “Google Play” та близько 24 категорій у “App Store”. Незважаючи на таку велику кількість і різноманіття, лише 7 категорій додатків привертають до себе увагу більше ніж 3% користувачів, за даними за даними дослідників з команди “Statistica”. На рисунку 3.1 наведена діаграма результатів їх дослідження.

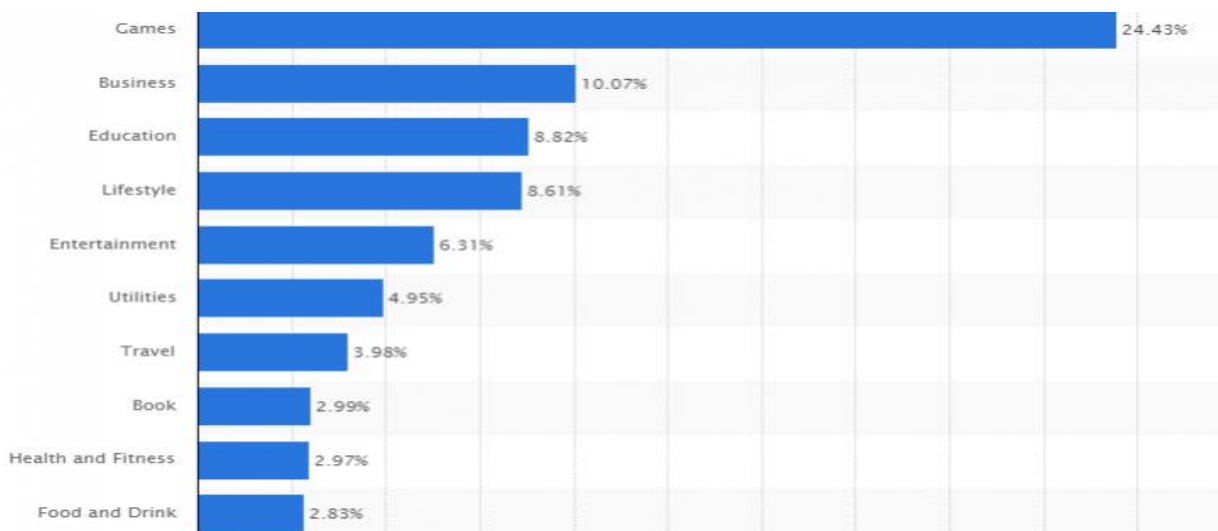


Рисунок 3.1 – Співвідношення популярності мобільних додатків

Фреймворк (Framework) — інфраструктура програмних рішень, що полегшує розробку складних систем. Спрощено дану інфраструктуру можна вважати своєрідною комплексною бібліотекою, але при цьому вона має ряд обмежень, що задають правила створення структури проекту та написання коду.

Програмний фреймворк (software framework) — це готовий до використання комплекс програмних рішень, включаючи дизайн, логіку та базову функціональність системи або підсистеми. Відповідно — програмний фреймворк може містити в собі також допоміжні програми, деякі бібліотеки коду, скрипти та загалом все, що полегшує створення та поєднання різних компонентів великого програмного

забезпечення чи швидке створення готового і не обов'язково об'ємного програмного продукту. Побудова кінцевого продукту відбувається, зазвичай, на базі єдиного API.

Java. Офіційна мова програмування, що підтримується середовищем розробки Android Studio. За даними щорічного опитування ресурсу Stackoverflow, в 2018 році Java увійшла до п'ятірки найпопулярніших мов програмування.

При розробці додатків на Java під ОС “Android” використовуються не тільки Java-класи, що містять код, але також файли маніфесту на мові XML, що надають системі основну інформацію про програму, і системи автоматичного складання Gradle, Maven або Ant, команди в яких пишуться на мовах Groovy, POM і XML відповідно; за замовчуванням в проектах використовується Gradle. Для верстки UI-частини зазвичай також використовується мова XML.

У грудні 2014 року Google визнала Android Studio офіційним середовищем розробки під ОС Android. Ця платформа вдосконалюється з року в рік, чим полегшує життя Android-розробникам. Такі її можливості, як візуальний UI-редактор і автодоповнення коду, допомагають зробити процес розробки більш комфортним.

C та C++. Більш низькорівневі мови, також підтримуються Android Studio з використанням Java NDK(Native Development Kit). Це дозволяє писати “рідні” додатки, що може стати в нагоді для створення ігор або інших ресурсоємних програм. Android Studio пропонує підтримку C / C ++ через Android NDK (Native Development Kit).

Це означає, що код буде запускатися не через Java Virtual Machine, а безпосередньо через смартфон, що дасть вам більше контролю над такими елементами системи, як пам'ять, сенсори, жести і т. Д., А також можливість вичавити з Android-пристроїв максимум ресурсів . Це також означає, що користуватися доведеться бібліотеками, написаними на C або C ++.

PhoneGap. PhoneGap дозволяє розробляти Android-додатки силами веб-розробки. Під час розробки, програма буде відображатися через WebView, але як би в обгортці мобільного додатка. Для розробників PhoneGap - це щось на зразок моста для

доступу до вбудованих функцій смартфона або планшета на кшталт акселерометра або камери. Під час розробки, програма буде відображатися через WebView, але як би в обгортці мобільного додатка Основними причинами використовувати PhoneGap для розробки додатків є:

- скорочення термінів і вартості розробки програми для декількох платформ відразу;
- можливість використовувати вже наявний сайт клієнта і його фрагменти.

Unity. Мультиплатформенний інструмент для розробки 2D і 3D додатків та ігор Unity. Вважається один з кращих інструментів для демонстрації 3D контенту. Створені за допомогою Unity програми працюють на операційних системах Windows, OS X, Linux, Android, Apple iOS, Windows Phone, BlackBerry, а також на ігрових приставках Wii, PlayStation 3 і Xbox 360 [13].

Переваги:

- відмінний варіант для створення мобільних ігор для цілого ряду пристроїв;
- 3D-двигжок дає високоякісні результати без будь-яких складних конфігурацій;
- присутня велика кількість хороших безкоштовних плагінів;
- “Unity” дозволяє розробнику зробити свої власні шейдери і змінити шлях, яким Unity візуалізує гру.

App Inventor. App Inventor для Android - це веб-додаток з відкритим вихідним кодом, що спочатку першопочатково був розроблений Google, а тепер підтримується Массачусетським технологічним інститутом (MIT). Компілятор, що транслює візуальну блокову мову App Inventor в байт-код Android, заснований на фреймворку GNU для реалізації динамічних мов Kawa.[14]

Дослідивши різноманітні підходи до розробки мобільних додатків, я вирішив, що для реалізації мого програмного продукту, найкращим чином підійде платформа

“App Inventor”. Завдяки відкритому коду та зручній конструкції даної платформи, можна UI-інтерфейсу для людей з обмеженими можливостями.

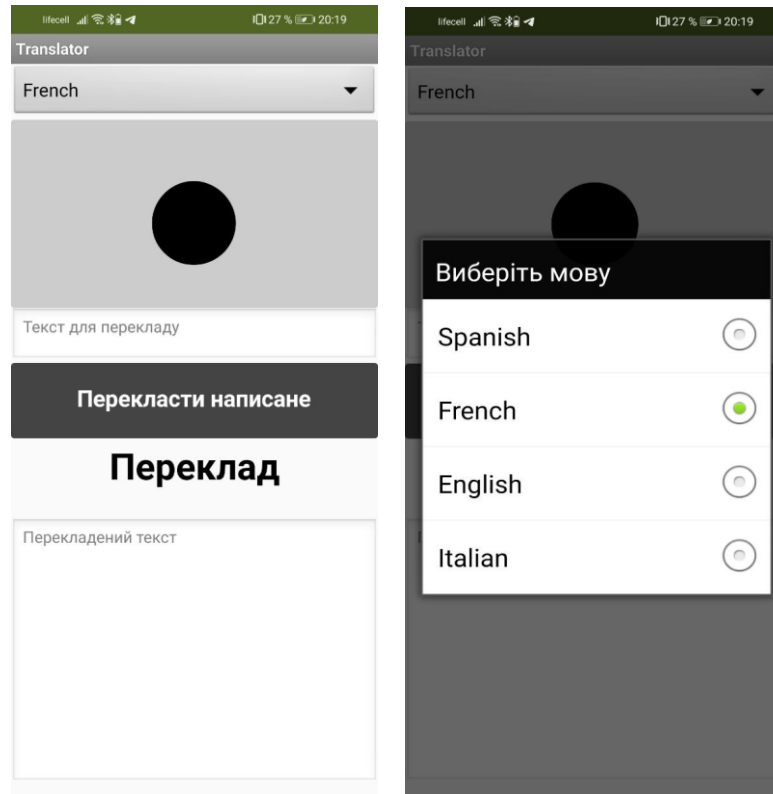


Рисунок 3.2 – Вигляд інтерфейсу та меню вибору мови

Згідно зі специфікою мого проекту, вимоги, яким має відповідати мобільний додаток мають наступний вигляд:

- наявність мінімалістичного адаптивного користувацького інтерфейсу, розробленого із застосуванням методів машинного навчання;
- наявність підлаштування контрастності кольорів інтерфейсу згідно з користувацькими вподобаннями;
- використання шрифтів великого розміру;
- наявність технології розпізнавання голосу;
- наявність перекладача тексту;
- відтворення перекладеного тексту речовим синтезатором;

— озвучування пунктів меню вибору мови.

Завдяки зручній технології попереднього перегляду App Inventor, яка реалізується за допомогою підключення утиліти “AI Companion”, наявна можливість одразу переглянути попередньо розроблений інтерфейс на екрані смартфона. В подальшому для вдосконалення та адаптивності ми застосуємо алгоритми машинного навчання. Він виглядає наступним чином, як зображено на рисунку 3.2.

Розроблюваний мобільний додаток має доволі вузьку направленість. Він пристосований під потреби людей з проблемами зору, а також, дає можливість спілкуватись між собою людям зі слуховими вадами.

Опис інтерфейсу. Попередній інтерфейс умовно можна розділити на 3 частини. У верхній частині знаходиться випадаюче меню вибору мови (рис. 3.5), а також велика кнопка, що активує функцію розпізнавання голосу. Даний додаток підтримує переклад на 4 мови – англійську, французьку, італійську та іспанську.

В середній частині інтерфейсу знаходиться текстове поле для альтернативного введення тексту (за допомогою екранної клавіатури). Це поле необхідне для надання можливості людям з розладами мовлення користуватися даним додатком. Нижче знаходиться кнопка, за допомогою натискання якої, активується одразу дві функції – перекладу написаного тексту, та озвучення його за допомогою речового синтезатору. Ця кнопка також знадобиться в разі, якщо користувач захоче повторно прослухати перекладений текст.

В нижній частині вікна знаходиться велике текстове поле, в якому відобразатиметься перекладений текст.

Найбільшим компонентом екрану користувача (40%) є кнопка початку голосового введення. Таке рішення зумовлене тим, що саме ця кнопка запускає основну функцію додатку, а тому, вона має виділятися на фоні інших компоненті, задля того, щоб користувачеві було легше її знайти та майже неможливо промазати, людям, які бачать зовсім слабко.

Також, виходячи з того, що потенційними користувачами даного додатку є особи з вадами зору, одразу після запуску мобільного додатку, користувачеві зачитується коротка інструкція по користуванню. Дана функція реалізована за допомогою використання технології речевого синтезатору. Для того, щоб отримати більш детальну аудіо-інструкцію з користування додатком, користувачеві достатньо потрусити своїм смартфоном.

Опис деяких програмних циклів. Для реалізації випадяючого меню вибору мови, з подальшим її озвученням, я використав програмний цикл, що зображено на рисунку 3.3.

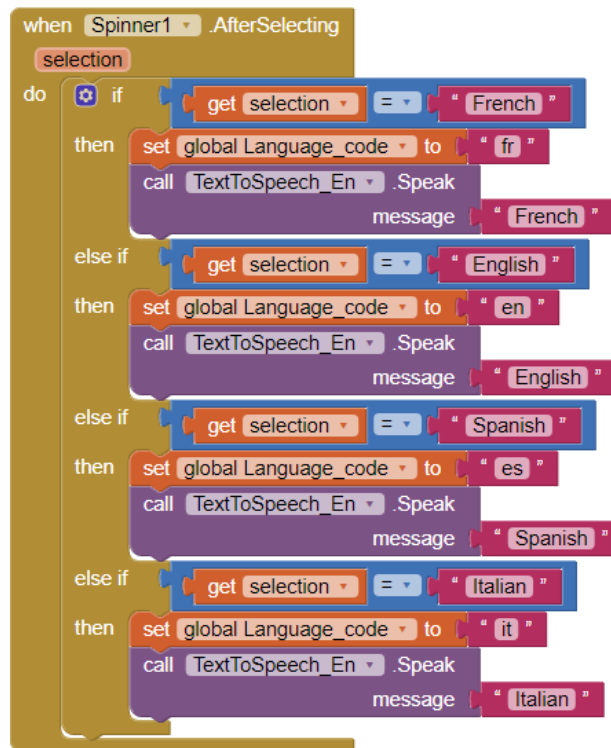


Рисунок 3.3 – Цикл роботи меню вибору мови

Наступний логічний цикл, що зображено на рисунку 3.4 виконується після отримання перекладу.

Цикл працює за наступним алгоритмом. При виборі одного із запропонованих варіантів мовного меню, попередньо ініціалізований глобальній змінній,

присвоюється значення певного мовного коду. Наприклад, якщо вибрано англійську мову, то глобальній змінній присвоюється значення “en”, якщо іспанську – “es” і так далі за аналогією. Після цього, системою викликається речовий синтезатор, що в залежності від вибраного пункту меню, озвучує його назву.

В ньому, одразу після отримання перекладу, до текстового поля, заноситься значення змінної, отриманої внаслідок виконання перекладу тексту.

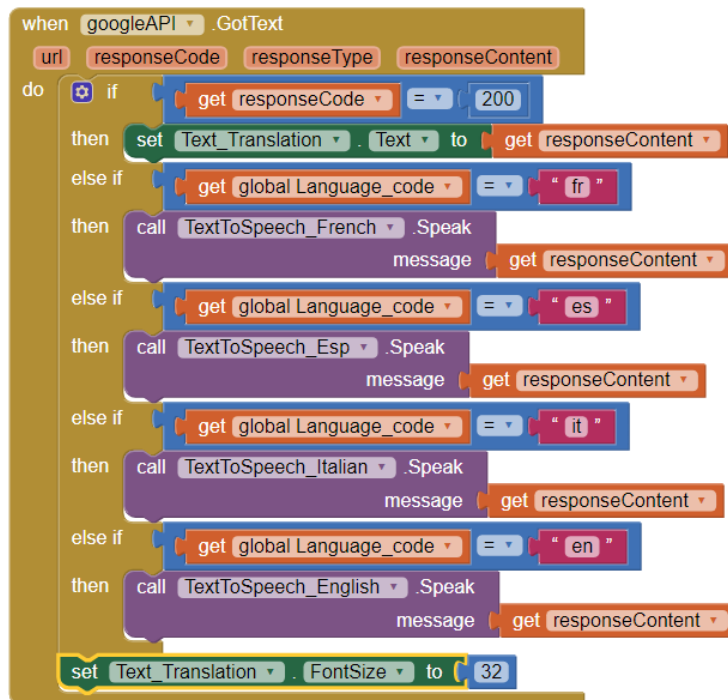


Рисунок 3.4 – Цикл отримання та озвучення перекладу

Після цього в залежності від того, яке значення присвоєне глобальній змінній мовного коду, викликається певний голосовий синтезатор, з налаштуваннями на коректне озвучування певної мови.

Наприклад, якщо глобальній змінній мови призначено мовний код “fr” – це означає, що виконувався переклад на французьку мову, тому буде викликаний речовий синтезатор з відповідними налаштуваннями, що коректно озвучить занесене до текстового поля повідомлення. Після цього розмір шрифту в текстовому полі буде збільшено до 32 пунктів.

Можна відзначити, що розмір всіх основних компонентів вікна додатку, особливо текстової частини є досить великим. Таке рішення було прийнято мною з міркувань зручності для людей з вадами зору.

3.2 Огляд та застосування нейронної мережі GNMT

Google Translate – безкоштовний онлайн-сервіс машинного перекладу, який дозволяє автоматично перекладати слова, тексти, електронні документи, веб-сторінки багатьма мовами світу. Для певних мов користувачам система пропонує варіанти перекладів, наприклад для термінів і багатозначних слів. Спочатку GT використовував перекладач SYSTRAN, яким і зараз користуються такі онлайн-сервіси як Vabel Fish, AOL та Yahoo.

Однак починаючи з жовтня 2007 року, Google запустила власне програмне забезпечення. Переклад здійснюється з використанням статистичного підходу до машинного перекладу, що дозволяє системі «навчатися» і самовдосконалюватися: коли Google Translate створює переклад, він шукає зразки в сотнях мільйонів документів, щоб надати найкращий переклад. Знайшовши зразки в текстах, перекладених іншими людьми, Google Translate може робити інтелектуальні припущення щодо правильного перекладу. Цей процес пошуку зразків у великих обсягах тексту і називається "статистичним машинним перекладом". Оскільки переклади генеруються машинами, не всі вони ідеальні. Чим більше текстів певною мовою перекладають люди, а Google Translate їх аналізує, тим кращою буде якість перекладу. Саме тому точність перекладу різними мовами іноді відрізняється.

Google Translate успішно інтегрується в браузері, завдяки спеціальним розширенням, що встановлюються додатково. Для Mozilla Firefox існує дуже багато плагінів, щодо Google Chrome, то дані компоненти вже встановлені у нових версіях. Існує також спеціальний клієнт для ОС Windows, що реалізований у безплатній та комерційній версії, яка відрізняється кількома компонентами. Сервіс також працює з

популярними зараз мобільними операційними системами, такими як Android та iOS. Для мобільних пристроїв була створена спеціальна функція голосового вводу, яка підтримує 15 мов (української, на жаль, немає), а система може «прочитати» переклад на 23 мовах за допомогою синтезу мовлення.

Google Translate – перспективна система для перекладу, тому що збільшення пам'яті перекладів і поповнення двомовних корпусів забезпечує кращу якість перекладів. Її перевагою є також те, що вона може самовдосконалюватися за рахунок збільшення баз даних. Однак статистичний метод не є ідеальним, тому система такого типу не зможе надати повноцінного перекладу.

В 2016 році Google представила систему нейронного машинного перекладу Google Neural Machine Translation (далі - GNMT), яка використовує штучну нейронну систему для покращення якості перекладу. Нейронна модель машинного перекладу використовує інші принципи роботи з текстом ніж стандартний статистичний метод перекладу [8].

Чим більше текстів певною мовою перекладають люди, а Google Translate їх аналізує, тим кращою буде якість перекладу. Саме тому точність перекладу різними мовами іноді відрізняється.

Оскільки переклади генеруються машинами, не всі вони ідеальні. Чим більше текстів певною мовою перекладають люди, а Google Translate їх аналізує, тим кращою буде якість перекладу. Саме тому точність перекладу різними мовами іноді відрізняється.

GNMT перекладає ціле речення, враховуючи контекст. Графічно алгоритм перекладу представлений на рисунку 3.5.

Система не запам'ятовує сотні варіантів перекладу фраз – вона працює з семантичною забарвленістю тексту. Оскільки переклади генеруються машинами, не всі вони ідеальні. Чим більше текстів певною мовою перекладають люди, а Google Translate їх аналізує, тим кращою буде якість перекладу. Під час перекладу, речення розбивається на словарні сегменти. Далі, за допомогою спеціальних декодерів,

система визначає “вагу” кожного сегменту в тексті. Далі визначаються максимально вірогідні значення і переклад сегментів. Останній етап – з’єднати перекладені сегменти з урахуванням правил граматики.

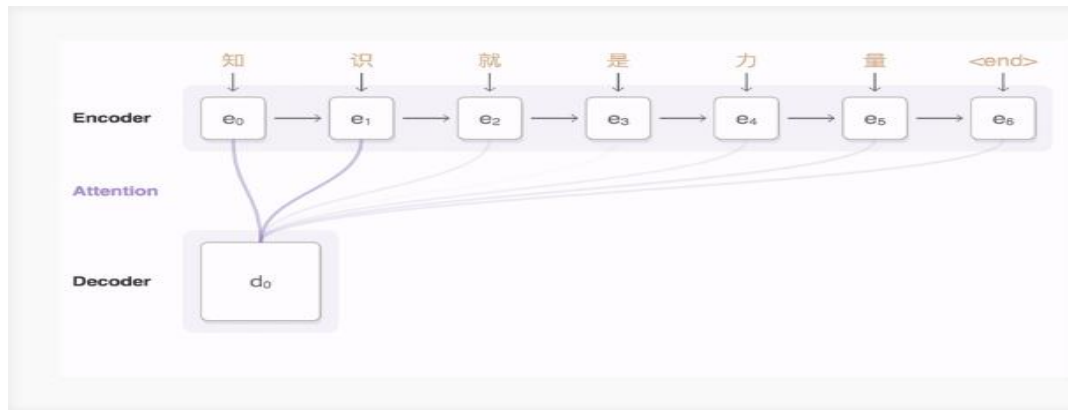


Рисунок 3.5 – Графічне представлення роботи Google Neural Machine Translation

Для того, щоб зрозуміти принцип роботи нейронного перекладу від Google, необхідно розібратися у внутрішньому алгоритмі його роботи.

В основі технології GNMT знаходиться принцип роботи рекурентних двонаправлених нейронних мереж (Bidirectional Recurrent Neural Networks), що опрацьовують матричні вираховування ймовірностей.

В нейронній системі, найменшим елементом є не слово, а його фрагменти. Це дозволяє зконцентрувати обчислювальні потужності не на словоформах, а на контексті і змісті речення. GNMT використовує близько 32000 таких фрагментів. За словами розробників, це дозволяє забезпечити високу швидкість і точність перекладу без використання надмірних обчислювальних потужностей.

Термін “рекурентний” означає, що система вираховує значення слова чи фрази на основі попередніх значень певної послідовності. Саме такий підхід дозволяє нейронній мережі враховувати контекст і правильно вибирати серед різних варіантів перекладу.

Двонаправленість означає те, що нейронна мережа розділена на два потоки – аналізуючий та синтезуючий(схема зображена на рис. 2.7). Кожен потік складається з восьми слоїв, які і проводять векторний аналіз. Перший потік розбиває речення на змістові елементи і аналізує їх, а другий – вираховує найбільш ймовірний варіант перекладу, виходячи з контексту та модулю уваги.

Звернемо увагу на те, що аналізуюча нейромережа зчитує речення як з ліва на право, так і з права на ліво. Це дозволяє їй в повній мірі розпізнати контекст. Окремо вона формує модуль уваги, за допомогою якого, другий потік виокремлює цінність окремих смислових фрагментів.

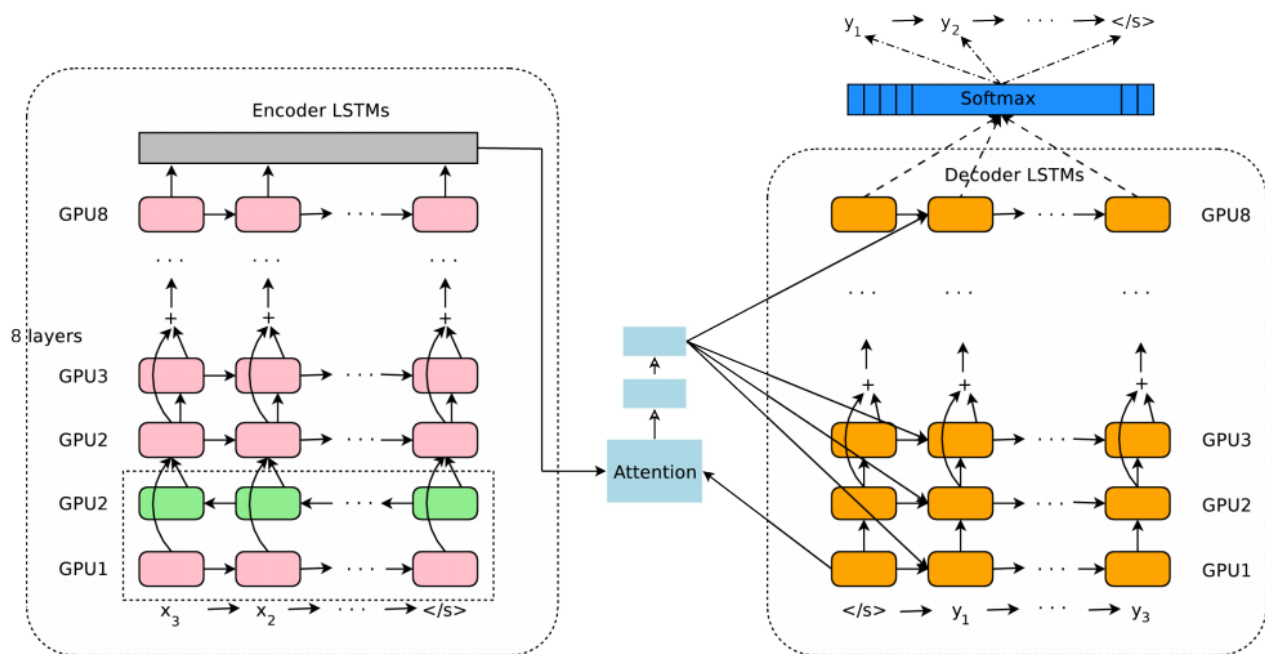


Рисунок 3.6 Графічне представлення двонаправлених рекурентних систем перекладу

Модель архітектури GNMT(рисунок 3.6), системи нейронного машинного перекладу Google. Зліва - мережа кодера, праворуч - мережа декодера, посередині - модуль уваги. Нижній шар кодера є двонаправленим: рожеві вузли збирають інформацію зліва направо, тоді як зелені вузли збирають інформацію справа наліво. Інші шари кодера є односпрямованими. Залишкові з'єднання починаються з рівня третього знизу в кодері та декодері. Модель розділена на кілька графічних процесорів,

щоб пришвидшити навчання. У нашій установці ми маємо 8 шарів LSTM(Long Short-Term Memory) кодера (1 двонаправлений шарі 7 односпрямованих шарів), і 8 шарів декодера. За допомогою цього налаштування одна репліка моделі розділена на 8 сторінок і розміщується на 8 різних графічних процесорах, що зазвичай належать одній хост-машині. Під час навчання нижній двонаправлений шар кодера обчислюється спочатку паралельно. Після закінчення обох шарів однонаправленого кодера можна розпочати обчислення, кожен на окремому графічному процесорі.

Аналіз фрагментів значно знижує ризики некоректного перекладу слів і словосполучень з різноманітними суфіксами, префіксами і закінченнями. Система самонавчання дозволяє нейронній мережі з високою точністю перекласти навіть ті поняття, які відсутні в загальноживаних словниках – сленг, жаргонізми та неологізми. Але нейромережа може працювати і інакше – виконувати переклад по літерам. Цей сценарій її роботи є необхідним при транслітерації – перекладі власних назв та імен з однієї мови на іншу. Система самонавчання дозволяє нейронній мережі з високою точністю перекласти навіть ті поняття, які відсутні в загальноживаних словниках – сленг, жаргонізми та неологізми.

Досліджувана модель уважно слідує загальним рамкам навчання послідовності до послідовності. Вона складається з трьох компонентів: мережі кодера, мережі декодера та мережі уваги. Кодер перетворює вихідне речення у список векторів, по одному вектору на кожен вхідний символ. Враховуючи цей список векторів, декодер видає по одному символу за раз, доки не буде створений спеціальний символ кінця речення (End-of-Sentence Symbol - EOS).

Щоб зберегти якомога більше паралелізму під час запуску шарів декодера, ми використовуємо вихідний шар нижнього декодера лише для отримання контексту періодичної уваги, який надсилається безпосередньо до всіх інших шарів декодера. Шар softmax також розділений і розміщений на декількох графічних процесорах. Під час навчання нижній двонаправлений шар кодера обчислюється спочатку паралельно.

Після закінчення обох шарів однонаправленого кодера можна розпочати обчислення, кожен на окремому графічному процесорі.

Залежно від розміру вихідного словника ми або маємо їх працювати на тих самих графічних процесорах, що і кодерів та декодерів мереж, або ж вони повинні працювати на окремому наборі спеціальних графічних процесорів.

При розрахунках ми використовуємо жирні маленькі літери для позначення векторів (наприклад, v , \mathbf{o}_i), жирні великі літери для позначення матриць (наприклад, U , W), курсивні великі літери позначення множин (наприклад, V , T), великі літери позначатимуть послідовність (наприклад, X , Y), а маленькі - окремі символи в послідовності (наприклад, x_1 , x_2).

Представимо, що (X, Y) - пара вихідного та цільового речень. Нехай $X = x_1, x_2, x_3 \dots, x_M$ - послідовність символів M у вихідному реченні, а $Y = y_1, y_2, y_3 \dots, y_N$ - послідовність N символів у цільовому реченні. Кодер - це просто функція наступної форми (1.1):

$$x_1, x_2, x_3, \dots, x_M = \text{EncoderRNN}(x_1, x_2, x_3, \dots, x_M) \quad (3.1)$$

У цьому рівнянні $x_1, x_2, x_3 \dots, x_M$ - це список векторів фіксованого розміру. Кількість членів у списку така ж, як і кількість символів у вихідному реченні (M у цьому прикладі). За допомогою правила ланцюга умовну ймовірність послідовності $P(Y | X)$ можна розкласти як:

$$P(Y | X) = P(Y | x_1, x_2, x_3, \dots, x_M) = \prod_{i=1}^N P(y_i | y_0, y_1, y_2, \dots, y_{i-1}; x_1, x_2, x_3, \dots, x_M) \quad (3.2)$$

де y_0 - спеціальний символ "початку речення", який додається до кожного цільового речення.

Під час умовиводу ми обчислюємо ймовірність наступного символу, враховуючи кодування вихідного речення і декодовану цільову послідовність до цього часу (3.3):

$$P(y_i | y_0, y_1, y_2, \dots, y_{i-1}; x_1, x_2, x_3, \dots, x_M) \quad (3.3)$$

Наш декодер реалізований як поєднання мережі RNN та рівня softmax. Мережа RNN декодера виробляє прихований стан y_i для наступного передбачуваного символу, який потім проходить через шар softmax, щоб генерувати розподіл ймовірності за вихідними символами-кандидатами. У наших експериментах ми виявили, що для досягнення хорошої точності систем NMT як кодер, так і декодер RNN повинні бути достатньо глибокими, щоб зафіксувати тонкі порушення в мовах джерела та цільовій мові. Це спостереження подібне до попередніх спостережень, що глибокі LSTM значно перевершують дрібні LSTM.

У модулі уваги (3.4), нехай y_{i-1} є виводом декодера-RNN з минулого кроку часу декодування (у нашій реалізації ми використовуємо вихід з нижнього шару декодера). Коефіцієнт уваги a_i для поточного кроку часу обчислюється за такими формулами:

$$s_t = \text{AttentionFunction}(y_{i-1}, x_t) \quad \forall t, 1 \leq t \leq M$$

$$p_t = \exp(s_t) / \sum_{t=1}^M \exp(s_t) \quad \forall t, 1 \leq t \leq M \quad (3.4)$$

$$a_i = \sum_{t=1}^M p_t \cdot x_t$$

Де функція уваги у нашій реалізації - це мережа прямого пересилання з одним прихованим шаром. У цій роботі кожен додатковий шар зменшував дисперсію майже на 10%. Ми використовуємо глибоко складену мережу довгострокової короткочасної

пам'яті (LSTM) як для кодера RNN, так і для декодера RNN. Кодер і декодер з'єднані за допомогою модуля уваги, який дозволяє декодеру фокусуватись на різних областях вихідного речення під час декодування.

Щоб мати найкращий можливий контекст у кожній точці мережі кодера, має сенс використовувати двонаправлений RNN для кодера, який також використовувався.

Більш конкретно, нехай $LSTM_i$ та $LSTM_{i+1}$ будуть i -ми та $(i+1)$ -ми слоями LSTM у стеку, параметри яких W_i і W_{i+1} відповідно. На t -му кроці часу для накопиченого LSTM без залишкових з'єднань ми маємо наступне рівняння (1.5):

$$\begin{aligned} c_t^i, m_t^i &= LSTM_i(c_{t-1}^i, m_{t-1}^i, x_t^{i-1}; W^i) \\ x_t^i &= m_t^i \\ c_t^{i+1}, m_t^{i+1} &= LSTM_{i+1}(c_{t-1}^{i+1}, m_{t-1}^{i+1}, x_t^i; W^{i+1}) \end{aligned} \quad (3.5)$$

У даному випадку, x_t^i – це введення даних в момент часу t , а m_t^i і c_t^i – приховані стани та стани пам'яті LSTM в момент часу t відповідно.

З урахуванням залишкових зав'язків між $LSTM_i$ та $LSTM_{i+1}$, наведені вище рівняння приймають наступну форму (1.6):

$$\begin{aligned} c_t^i, m_t^i &= LSTM_i(c_{t-1}^i, m_{t-1}^i, x_t^{i-1}; W^i) \\ x_t^i &= m_t^i + x_t^{i-1} \\ c_t^{i+1}, m_t^{i+1} &= LSTM_{i+1}(c_{t-1}^{i+1}, m_{t-1}^{i+1}, x_t^i; W^{i+1}) \end{aligned} \quad (3.6)$$

Слідуючи концепції моделювання відмінностей між вихідним рівнем проміжного шару та цільовими показниками, яка в минулому показала свою

ефективність для багатьох проектів, ми вводимо залишкові зв'язки між шарами LSTM у стеку.

Як зазначалося вище, глибокоскладені LSTM часто дають кращу точність порівняно з більш дрібними моделями. Однак просто укладання більшої кількості шарів LSTM працює лише на певну кількість шарів, за межами яких мережа стає занадто повільною та важкою для навчання, ймовірно через проблеми вибуху та зникнення градієнта. Щоб мати найкращий можливий контекст у кожній точці мережі кодера, має сенс використовувати двонаправлений RNN для кодера, який також використовувався.

За нашим досвідом роботи з масштабними завданнями перекладу, прості шари LSTM з накопиченням працюють добре до 4 шарів, ледве з 6 шарами, і дуже погано перевищують 8 шарів.

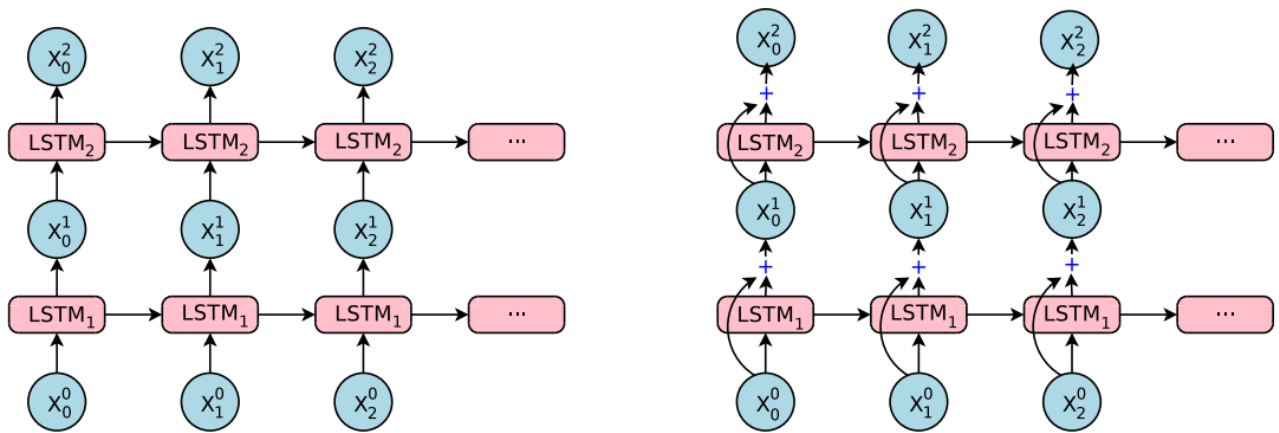


Рисунок 3.7 Графічна схема роботи з'єднання LSTM систем

Різниця між звичайним LSTM з накопиченням та нашим LSTM з накопиченням із залишковими з'єднаннями представлена на рисунку 3.7. Зліва: прості складені шари LSTM. Справа: наша реалізація складених шарів LSTM із залишковими з'єднаннями. За залишкових з'єднань, вхід до нижнього рівня LSTM (x_i^0 до $LSTM_i$) елементарно додається до виводу з нижнього (x_i^1) го шару. Потім ця сума подається

на верхній рівень LSTM ($LSTM_2$) як новий тип вхідних даних. Як зазначалося вище, глибокоскладені LSTM часто дають кращу точність порівняно з більш дрібними моделями. Однак просто укладання більшої кількості шарів LSTM працює лише на певну кількість шарів, за межами яких мережа стає занадто повільною та важкою для навчання, ймовірно через проблеми вибуху та зникнення градієнта. За нашим досвідом роботи з масштабними завданнями перекладу, прості шари LSTM з накопиченням працюють добре до 4 шарів, ледве з 6 шарами, і дуже погано перевищують 8 шарів.

Залишкові з'єднання значно покращують градієнтний потік при зворотньому потоці, що дозволяє нам навчити дуже глибокі мережі кодування та декодування. У більшості наших експериментів ми використовуємо 8 шарів LSTM для кодера та декодера, хоча залишкові з'єднання можуть дозволити нам навчити значно глибші.

Для систем перекладу інформація, необхідна для перекладу певних слів на вихідній стороні, може відображатися де завгодно на вихідній стороні. Часто інформація на вихідній стороні приблизно зліва направо, подібно до 5 цільової, але залежно від мовної пари інформація для конкретного вихідного слова може розподілятися і навіть розподілятися в певних регіонах вхідної сторони.

Щоб забезпечити максимально можливу розпаралелювання під час обчислення, двонаправлені з'єднання використовуються лише для нижнього шару кодера - всі інші шари кодера є односпрямованими. Рисунок 3.8 ілюструє використання двонаправлених LSTM на нижньому шарі кодера. Шар $LSTM_f$ обробляє вихідне речення зліва направо, тоді як шар $LSTM_b$ обробляє вихідне речення справа наліво.

На додаток до паралельності даних, для покращення швидкості обчислення градієнта на кожній репліці використовується модельний паралелізм. Мережі кодера та декодера розділені вздовж виміру глибини та розміщені на декількох графічних процесорах, ефективно виконуючи кожен шар на іншому графічному процесорі.

Модельний паралелізм накладає певні обмеження на модельні архітектури, якими ми можемо користуватися. Наприклад, ми не можемо дозволити собі

двонаправлені шари LSTM для всіх шарів кодера, оскільки це призведе до зменшення паралельності між наступними шарами, оскільки кожному шару доведеться чекати, поки не закінчаться напрямки попереднього шару вперед і назад.

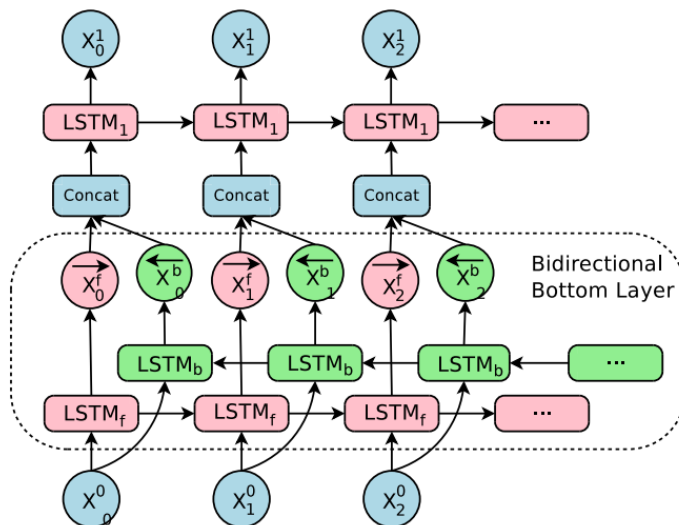


Рисунок 3.8 Робота двонаправлених LSTM на нижньому шарі кодера

Через складність нашої моделі, ми використовуємо як паралельність моделі, так і паралельність даних, щоб пришвидшити навчання. Паралелізм даних є простим: ми одночасно навчаємо n реплік моделі, використовуючи алгоритм Downpour SGD [12]. Усі n реплік мають спільну копію параметрів моделі, причому кожна репліка асинхронно оновлює параметри за допомогою комбінації алгоритмів Адама [25] та SGD. У наших експериментах n часто становить близько 10. Кожна копія одночасно працює на міні-партії з m пар речень, що часто становить 128 в наших експериментах. На додаток до паралельності даних, для покращення швидкості обчислення градієнта на кожній репліці використовується модельний паралелізм.

Мережі кодера та декодера розділені вздовж виміру глибини та розміщені на декількох графічних процесорах, ефективно виконуючи кожен шар на іншому графічному процесорі. Оскільки всі, крім першого рівня кодера, є односпрямований, рівень $i + 1$ може розпочати обчислення до того, як рівень i буде повністю закінчений,

що покращує швидкість навчання. Шар softmax також розділений, кожен розділ відповідає за підмножину символів у вихідному словнику. На рисунку 1 показано більше деталей про те, як робиться розділення. Модельний паралелізм накладає певні обмеження на модельні архітектури, якими ми можемо користуватися. Наприклад, ми не можемо дозволити собі двонаправлені шари LSTM для всіх шарів кодера, оскільки це призведе до зменшення паралельності між наступними шарами, оскільки кожному шару доведеться чекати, поки не закінчаться напрямки попереднього шару вперед і назад. Це фактично змусило б нас використовувати лише 2 графічні процесори б паралельно (один для прямого напрямку, а другий для зворотного). До уваги частина модель, ми вирішили вирівняти вихідні дані нижнього декодера з висновками верхнього кодера, щоб максимізувати паралельність під час роботи мережі декодера. Якби ми вирівняли верхній шар декодера до верхнього рівня кодера, ми б видалили весь паралелізм у мережі декодера і не отримали б користь від використання більше одного графічного процесора для декодування.

Моделі нейронного машинного перекладу часто працюють із словниками із фіксованими словами, хоча переклад є принципово відкритою проблемою словникового запасу (імена, числа, дати тощо). Існує дві широкі категорії підходів до вирішення питань перекладу слів, що не містять словникового запасу. Один із підходів полягає в простому копіюванні рідкісних слів з джерела в ціль (оскільки більшість рідкісних слів - це імена чи цифри, де правильним перекладом є просто копія), або на основі моделі уваги, використовуючи зовнішню модель вирівнювання, або навіть з використанням більш складної сітчастої вказівної мережі. Ще однією широкою категорією підходів є використання підслівних одиниць, наприклад, символів, змішаних слів / символів або більш розумних підслів.

Наш найуспішніший підхід відноситься до другої категорії (підслівні одиниці), і ми застосовуємо модель введення тексту (WPM), спочатку розроблену для вирішення японсько-корейської проблеми сегментації для системи розпізнавання мовлення Google. Цей підхід повністю керується даними та гарантовано генерує

детерміновану сегментацію для будь-якої можливої послідовності символів. Це схоже на метод, який використовується для боротьби з рідкісними словами в нейронному машинному перекладі. Для обробки довільних слів ми спочатку розбиваємо слова на заготівлі, отримавши навчену модель тексту.

Спеціальні символи меж слова додаються перед навчанням моделі, так що вихідну послідовність слів можна відновити з послідовності заготівлі без двозначності. Під час декодування модель спочатку створює послідовність заготівлі, яка потім перетворюється у відповідну послідовність слів

Модель текстового заголовка генерується з використанням підходу, керованого даними, щоб максимізувати ймовірність мовної моделі навчальних даних, враховуючи еволюційне визначення слова.

Враховуючи навчальний корпус і кількість бажаних лексем D , завдання оптимізації полягає у виборі D -заголовків таким чином, щоб отриманий корпус був мінімальним за кількістю наборів при сегментації відповідно до обраної моделі заготівлі. Наш жадібний алгоритм цієї задачі оптимізації. Порівняно з початковою реалізацією, ми використовуємо спеціальний символ лише на початку слів, а не на обох кінцях.

Ми також скорочуємо кількість основних символів до керованої кількості залежно від даних (приблизно 500 для західних мов, більше для азійських мов), а решту призначаємо спеціальному невідомому символу, щоб уникнути забруднення даного словника слів дуже рідкісними символами. Ми виявили, що використання загального словникового запасу від 8 до 32 тис. Заголовків досягає як високої точності (оцінки BLEU), так і швидкої швидкості декодування у всіх парах мовних пар, які ми пробували. Як уже згадувалося вище, при перекладі часто має сенс копіювати рідкісні імена або числа безпосередньо від джерела до цілі. Щоб полегшити цей тип прямого копіювання, ми завжди використовуємо спільну модель заголовка як для вихідної, так і для цільової мови.

Використовуючи такий підхід, гарантується, що однаковий рядок у вихідному та цільовому реченні буде сегментований точно так само, що полегшить системі навчання скопіювати ці маркери. Письменники забезпечують баланс між гнучкістю символів та ефективністю слів. Ми також виявляємо, що наші моделі отримують кращі загальні показники BLEU при використанні заголовків - можливо, завдяки тому, що наші моделі зараз ефективно працюють із практично нескінченним словниковим запасом, не вдаючись лише до символів. Останні 7 зробили б середні довжини вхідних та вихідних послідовностей набагато довшими, а тому потребували б більших обчислень.

3.3 Впровадження інтерфейсу, заснованого на технології машинного навчання

Для задоволення потреб людей з обмеженими можливостями в минулому фахівці розробляли різні інтерфейси, але ці підходи були зроблені для задоволення конкретних потреб групи користувачів з конкретною інвалідністю. Наприклад, використовували спеціальну систему виявлення обличчя для моніторингу фізичної активності людей, які страждають на церебральний параліч. розробив недорогу систему відстеження очей для створення людського комп'ютерного інтерфейсу для людей, які демонструють обмежений рух своїх тіл. У розпізнаванні жестів використовується як метод взаємодії для спілкування та забезпечує систему взаємодії з комп'ютером для людей, які страждають на ДЦП, на основі системи відстеження облич.

У мовленні та під час відстеження голови використовуються підходи, щоб забезпечити вільне керування комп'ютером та іншими пристроями, і розширив цей підхід за допомогою розпізнавання обличчя. Запровадив подібний підхід, але замість відстеження голови використовували жести обличчя та рук. використовується

сенсорний та мовленнєві підходи в мобільному пульті дистанційного керування, що дозволяють інвалідам керувати побутовою технікою.

Формулювання проблеми: Вивчаючи та аналізуючи вищезазначені підходи та декілька інших, було виявлено такі проблеми: Конкретна цільова група користувачів - У більшості випадків програма орієнтована лише на один тип інвалідності. Це може спричинити проблеми, як тільки цільова група користувачів не зможе скористатися програмою через вторинну інвалідність. Наприклад, додаток, що надається для німецьких людей похилого віку, які страждають погіршенням зору, напевно, не може бути ефективно використаний, оскільки використання програми стосується рук або верхніх кінцівок користувачів. Навіть якщо для цього не знадобляться кінцівки, а слухові можливості, ці конкретні старші користувачі не можуть використовувати додаток із очікуваною продуктивністю. Спеціалізований дизайн інтерфейсу користувача - Надані інтерфейси користувача спрямовані на окремі конкретні завдання, такі як написання, розмова або рухливість (управління інвалідним кріслом). Однак це не ефективно для підготовки людей з інвалідністю для вищої освіти або професій, оскільки це в кінцевому підсумку надасть окремі заявки на кожну окрему діяльність. Проблема взаємозалежності - Вона виникає, коли цільові групи користувачів занадто загальні. Наприклад, використовувати систему письма / читання, засновану на дотику, для людей із вадами зору (шрифт Брайля) надзвичайно важко використовувати, якщо користувач також паралегічний і одночасно потрібно керувати інвалідним візком.

Одним з найважливіших критеріїв для розробки ефективного користувальницького інтерфейсу (UI) є фактор масштабованості користувача. Як обговорювалося в попередньому розділі, майже всі ці інтерфейси намагалися компенсувати фізіологічні обмеження, додавши деякі спеціальні функції доступності. Однак такий підхід різко зменшує масштабованість інтерфейсу. Отже, далі ми представляємо новий орієнтований на користувача підхід до розробки інтерфейсу для людей з обмеженими можливостями.

Принцип полягає у тому, щоб зосередити увагу на користувачі та його потребах, проаналізувавши три взаємозалежні простори: користувач, потреби та простір додатків. Користувацький простір: ВООЗ каже, що існує більше одного мільярда людей з різними типами інвалідності у всьому світі, з яких приблизно 150 мільйонів відчувають значні труднощі. Близько 285 мільйонів людей із вадами зору та 360 мільйонів мають вади слуху. Різні дослідники надали різні точки зору на класифікацію інвалідності, наступним чином : Медична перспектива визначає інвалідів як людей із проблемами здоров'я, які спричинені захворюваннями, травмами, вроджені або отримані в наслідок нещасного випадку.

Соціально-політична перспектива зосереджена на філософії незалежного життя, а не на складному наборі умов, незалежно від того, як люди втрачали свою функціональність. Перспектива Спектру стосується ступеня функціональності певних органів чуття, таких як зоровий, слуховий або тактильний сенс. Економічна перспектива визначає інвалідів з точки зору працевлаштування на ринку. На думку Всесвітньої організації охорони здоров'я (ВООЗ), інвалідність не слід розглядати ні як суто медичну, ні як суто соціальну. Необхідний зважений підхід, який надає належну вагу різним аспектам інвалідності. Інвалідність виникає внаслідок взаємодії умов здоров'я з контекстними, екологічними та особистими факторами. Міжнародний

Класифікація Асоціації функціонування, інвалідності та здоров'я (ICF) приймає нейтральну мову і не робить різниці між типом та причиною інвалідності. Вони описують інвалідність простим та зрозумілим способом з урахуванням функціонування людини та її обмежень та структурованою інформацією у значущий спосіб. Ми віримо, що незалежно від причини та від обмежень, які страждають інваліди, їх слід розглядати як людей з тими ж потребами, що й інші. Різниця лише в способі задоволення їхніх потреб. Для цього розробникам програм та взаємодії потрібно вивчити відповідне різномаяття потреб.

Потрібний простір: У цій роботі я пропоную нове визначення поняття "потреба", таке як "основні цілі та види діяльності, пов'язані із способами взаємодії

людей". Іншими словами, щоб визначити основні цілі та заходи для людей із вадами зору, необхідно вказати дані, які надає спосіб зору. Їх можна перерахувати наступним чином: виявлення об'єкта, ідентифікація об'єкта, відстань між об'єктами, відстань до об'єктів. Для завершення аналізу потреби важливо визначити роль кожної потреби в різних сферах життя. Простір програми. Попередні розділи показали, що інвалідність можна описати як обмеження ступеня здібностей (DOA).

Color Contrast Checker

[Home](#) > [Resources](#) > Color Contrast Checker

Foreground Color
#000000
Lightness

Background Color
#969696
Lightness

Contrast Ratio
7.09:1
[permalink](#)

Normal Text
WCAG AA: **Pass**
WCAG AAA: **Pass**
The five boxing wizards jump quickly.

Large Text
WCAG AA: **Pass**
WCAG AAA: **Pass**
The five boxing wizards jump quickly.

Graphical Objects and User Interface Components
WCAG AA: **Pass**
Text Input

Рисунок 3.9 – Перевірка контрастності за допомогою сервісу “Webiam”

Для комфортного користування мобільним додатком людям із проблемами зору, користувацький інтерфейс має містити як можна менше функціональних можливостей, зібраних на одному екрані.

Світова практика у цій сфері каже – якщо можеш зменшити кількість дій, що має виконувати користувач – зроби це. Такий підхід зумовлено тим, що додаток має бути “невидимим”, тобто не перетягувати на себе увагу, а лише якісно виконувати мінімальну кількість завдань.

Крім того, він показав, як аналізувати основні цілі та потреби, що відповідають обмеженням, спричиненим порушення модальності. Тепер питання полягає в тому, як вибрати відповідні замітники, щоб вирішити або зменшити наслідки таких обмежень.

Для завершення аналізу потреби важливо визначити роль кожної потреби в різних сферах життя. Наприклад, виявлення об'єкта може бути як виявленням слів, речень, так і символом або метою читання. Або це може забезпечити виявлення столів, дверей або коридору з метою пересування.

Окрім явища природного заміщення, технологічний замітник також має значний вплив на поліпшення умов життя осіб з обмеженими можливостями. У таблиці 1 узагальнено ряд органів чуття, виконавчих механізмів та технологій, які можуть бути замінені один одним. З іншого боку, бувають ситуації, що обмежують технологічний замітник конкретними умовами. Наприклад, відстеження погляду очей можна використовувати як вказівник миші, якщо користувач не переглядає певну область, а функція перетворення тексту в текст працює досить добре, якщо користувач не перебуває в галасливому середовищі. На додаток до згаданих недоліків, слід вирішити випадок, коли особа страждає від кількох інвалідностей одночасно, оскільки в таких випадках, крім проблеми адаптації, виникає також проблема взаємозалежності.

Також при розробці інтерфейсу слід враховувати необхідність відповідності контрасту кольорів міжнародному стандарту WCAG (Web Content Accessibility Guidelines). Згідно цього стандарту, мінімальний рівень контрастності основного кольору і фону в інтерфейсах для незрячих – 4.5 до 1. Контраст кольору шрифту має складати 3 до 1.

Отже, як видно з даних зображень, при розробці інтерфейсу, нейронна мережа має контролювати, аби параметри інтерфейсу задовольняли наведені вимоги WCAG (Web Content Accessibility Guidelines).

Для перевірки цих параметрів використовуються такі сервіси як “Webiam” та “Leaverou”. При розробці UI-інтерфейсу, я використовував сервіс “Webiam”(рис. 3.9 та 3.10).

Color Contrast Checker

[Home](#) > [Resources](#) > Color Contrast Checker

The screenshot displays the 'Color Contrast Checker' interface. At the top, it shows the breadcrumb 'Home > Resources > Color Contrast Checker'. Below this, there are two color selection panels. The 'Foreground Color' panel shows '#FFFFFF' and a 'Lightness' slider. The 'Background Color' panel shows '#3B3B3B' and a 'Lightness' slider. To the right, a green-bordered box displays 'Contrast Ratio 11.2:1' with a 'permalink' link below it. The main content area is divided into three sections, each with a red underline:

- Normal Text:** WCAG AA: **Pass**, WCAG AAA: **Pass**. The sample text is 'The five boxing wizards jump quickly.' on a dark background.
- Large Text:** WCAG AA: **Pass**, WCAG AAA: **Pass**. The sample text is 'The five boxing wizards jump quickly.' on a dark background.
- Graphical Objects and User Interface Components:** WCAG AA: **Pass**. The sample is a 'Text Input' field with a checkmark above it on a dark background.

Рисунок 3.10 – Перевірка контрастності за допомогою сервісу “Webiam”

Для комфортного користування мобільним додатком людям із проблемами зору, користувацький інтерфейс має містити як можна менше функціональних можливостей, зібраних на одному екрані.

Світова практика у цій сфері каже – якщо можеш зменшити кількість дій, що має виконувати користувач – зроби це. Такий підхід зумовлено тим, що додаток має бути “невидимим”, тобто не перетягувати на себе увагу, а лише якісно виконувати мінімальну кількість завдань. Одним із рішень у такому випадку є користувацький інтерфейс, який надзвичайно пристосовується та розумний, що

може збільшити час зміни фокусу та часу вибору шляхом вивчення вподобань користувачів та інтелектуальної адаптації.

Принцип полягає у тому, щоб зосередити увагу на користувачі та його потребах, проаналізувавши три взаємозалежні простори: користувач, потреби та простір додатків. Користувацький простір: ВООЗ каже, що існує більше одного мільярда людей з різними типами інвалідності у всьому світі, з яких приблизно 150 мільйонів відчують значні труднощі. Близько 285 мільйонів людей із вадами зору та 360 мільйонів мають вади слуху. Різні дослідники надали різні точки зору на класифікацію інвалідності, наступним чином : Медична перспектива визначає інвалідів як людей із проблемами здоров'я, які спричинені захворюваннями, травмами, вроджені або отримані в наслідок нещасного випадку.

Одним з найважливіших критеріїв для розробки ефективного користувацького інтерфейсу (UI) є фактор масштабованості користувача. Як обговорювалося в попередньому розділі, майже всі ці інтерфейси намагалися компенсувати фізіологічні обмеження, додавши деякі спеціальні функції доступності. Однак такий підхід різко зменшує масштабованість інтерфейсу. Отже, далі ми представляємо новий орієнтований на користувача підхід до розробки інтерфейсу для людей з обмеженими можливостями.

Отже, як видно з даних зображень, при розробці інтерфейсу, нейронна мережа має контролювати, аби параметри інтерфейсу задовольняли наведені вимоги WCAG (Web Content Accessibility Guidelines).

ВИСНОВКИ

Під час створення магістерської роботи було проведено дослідження використання сучасних інформаційних технологій людьми з обмеженими можливостями. З нього було зроблено висновок про актуальність даної роботи.

В рамках даної дипломної роботи, було проведено дослідження різноманітних методів машинного навчання та оцінка їх зручності при розробці додатків для людей з особливими потребами. За результатами такого дослідження стало зрозумілим, що технологія штучного інтелекту є однією з найперспективніших галузей у світі IT-технологій. Її використання допомагає розробити адаптивний інтерфейс, спеціалізований для людей з обмеженими можливостями.

Було розроблено мобільний додаток онлайн перекладу для людей з обмеженими можливостями. Для його конструювання використовувалася платформа “App Inventor”, а також - різноманітні Android та Google API, що наповнювали його необхідним функціоналом, як то розпізнавання мови, голосовий синтезатор чи онлайн-переклад. Також, при конструюванні UI-інтерфейсу мною були задіяні методи машинного навчання для дотримання міжнародних стандартів WCAG (Web Content Accessibility Guidelines) щодо контрастності кольорів інтерфейсу та величини шрифтів.

Мета роботи — дослідити ефективність впровадження технологій машинного навчання при розробці додатку для людей з особливими потребами була досягнута. За допомогою технологій обробки природної мови, методу кластеризації, голосового синтезатору, онлайн-перекладу та голосових інструкцій для користувача було виявлено, що впровадження технології штучного інтелекту дає можливість розробити адаптивний інтерфейс, що буде підлаштовуватися під конкретні потреби людини з обмеженими можливостями.

ПЕРЕЛІК ПОСИЛАНЬ

1. T. Üstün, N. Kostanjsek, S. Chatterji, J. Rehm. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0) / - World Health Organization 2010. — 152 с.
2. Assistive technology (AT): products, equipment, and systems that enhance learning, working, and daily living for persons with disabilities [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://www.atia.org/at-resources/what-is-at/> (дата звернення 30.09.2020). – Назва з екрана.
3. Types of Assistive technology [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://webaccess.berkeley.edu/resources/assistive-technology/> (дата звернення 12.10.2020). – Назва з екрана.
4. Tech Advancements For People With Disabilities [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://tech.co/news/5-tech-advancements-people-with-disabilities-2015-10> (дата звернення 12.10.2020). – Назва з екрана.
5. Useful IT technologies for disabled [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://disit.uk/technology/tekhnlolgy-for-disabled/> (дата звернення 12.10.2020). – Назва з екрана.
6. Speech synthesis [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://techcrunch.com/speech-synthesis/> (дата звернення 15.10.2020). – Назва з екрана.
7. Comparison of machine translation applications [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://it-tutor.com/machine-translation/>
8. Neural Machine translation [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://www.cossa.us/trends/196086/> (дата звернення 15.10.2020). – Назва з екрана.
9. The first technology for deaf and hard of hearing who accompanied free calls [Електронний ресурс]. Електронні дані. – Режим доступу - <https://rogervoice.com/en/> (дата звернення 20.10.2020). – Назва з екрана.

10. Ava gives the deaf and hard-of-hearing a more present voice in group conversations [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://techcrunch.com/2016/11/21/ava-gives-the-hearing-impaired-a-more-present-voice-in-group-conversations/> (дата звернення 20.10.2020). – Назва з екрана.

11 Improvement on speech recognition and synthesis for disabled individuals using fuzzy neural net retrofits [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://ieeexplore.ieee.org/document/23936/authors#authors> (дата звернення 08.11.2020). – Назва з екрана.

12. Data Science and Disability [Електронний ресурс]. – Електронні дані. – Режим доступу <https://www.kdnuggets.com/2016/03/data-science-disability.html> (дата звернення 08.11.2020). – Назва з екрана.

13. Mobility Aids Detection using Convolution Neural Network (CNN) [Електронний ресурс]. – Електронні дані. – Режим доступу <https://core.ac.uk/download/pdf/190046504.pdf> (дата звернення 08.11.2020). – Назва з екрана.

14. Felicia Kamriani Krishnendu Roy. App Inventor 2 Essentials / - Packt Publishing 2016 — 225 с.

15. The MIT App Inventor Library: Documentation & Support [Електронний ресурс]. – Електронні дані. – Режим доступу - <http://appinventor.mit.edu/explore/library.html> (дата звернення 17.11.2020). – Назва з екрана.

16. Ethics, Machine Learning and Disabilities [Електронний ресурс]. – Електронні дані. – Режим доступу - <https://abilitynet.org.uk/news-blogs/ethics-machine-learning-and-disabilities> (дата звернення 17.11.2020). – Назва з екрана.

17. Millan, J.D.R., et al.: Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience* 4, с. 1–15 (2010)

18. Ong, C., Lu, M.V., Lau, B.: A face based real time communication for physically and speech disabled people. In: Assistive and Augmentive Communication for the Disabled, c. 70–102 (2011)
19. Vazquez, L.J.G., Minor, M.A., Sossa, A.J.H.: Low cost human computer interface voluntary eye movement as communication system for disabled people with limited movements. In: 2011 Pan American Health Care Exchanges, c. 165–170. IEEE (2011)
20. Murata, Y., Yoshida, K., Suzuki, K., Takahashi, D.: Proposal of an automobile driving interface using gesture operation for disabled people. In: 6th International Conf. on Advances in Computer-Human Interactions, ACHI 2013, c. 472–478 (2013)
21. Karper, A., Rogers, A.: Icando: Low cost multimodal interface for hand disabled people. *Journal on Multimodal User Interfaces* 1(2), 21–29 (2007)
22. Varona, J., Manresa-Yee, C., Perales, F.J.: Hands-free vision-based interface for computer accessibility. *Journal of Network and Computer Applications* 31(4), c. 357– 374 (2008)
23. Valles, M., et al.: Multimodal environmental control system for elderly and disabled people. In: Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 2, c. 516–517. IEEE (1996)
24. WHO: Draft action plan for the prevention of avoidable blindness and visual impairment 2014-2019. Sixty-Sixth World Health Assembly (2013)
25. ICF: International classification of functioning, disability and health / World Health Organization. World Health Organization, Geneva (2011)