

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В. ДАЛЯ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

До захисту допускається
Т.в.о.завідувача кафедри
_____ Сафонова С.О.
« ____ » _____ 2020 р.

МАГІСТЕРСЬКА РОБОТА

НА ТЕМУ:

_____ Методи і алгоритми розпізнавання природнього мовлення _____

Освітньо-кваліфікаційний рівень “Магістр”
Спеціальність 122 – “Комп’ютерні науки”

Науковий керівник роботи:

_____ (підпис)

Білобородова Т.О.

_____ (ініціали, прізвище)

Консультант з охорони праці:

_____ (підпис)

Критська Я.О.

_____ (ініціали, прізвище)

Студент:

_____ (підпис)

Фурса П.С.

_____ (ініціали, прізвище)

Група:

_____ КН-18дм _____

Севєродонецьк 2020

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВОЛОДИМИРА ДАЛЯ

Факультет Інформаційних технологій та електроніки
Кафедра Комп'ютерних наук та інженерії
Освітньо-кваліфікаційний рівень Магістр
Напрямок підготовки _____
(шифр і назва)
Спеціальність 122 – “Комп'ютерні науки”
(шифр і назва)

ЗАТВЕРДЖУЮ:

Т.в.о. завідувача кафедри
_____ С.О. Сафонова
« _____ » _____ 2020 р.

**З А В Д А Н Н Я
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ**

Фурса Поліні Сергіївні
(прізвище, ім'я, по батькові)

1. Тема роботи Методи і алгоритми розпізнавання природнього мовлення
- керівник проекту (роботи) к.т.н. Білобородова Тетяна Олександрівна
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
- затверджені наказом вищого навчального закладу від " 11" 10 2019 р. № 35/15.15
2. Строк подання студентом роботи _____
3. Вихідні дані до роботи Матеріали переддипломної практики
4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) 1. Аналіз галузі застосування та технології обробки й алгоритмів розпізнавання природнього мовлення;
2. Аналіз використання глибоких нейронних мереж для розпізнавання природнього мовлення;
3. Практична реалізація обробки та аналізу природнього мовлення з використанням глибоких нейронних мереж;
4. Охорона праці в галузі;
5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	ст. викл. Критська Я.О.		

7. Дата видачі завдання _____

Керівник

_____ (підпис)

Завдання прийняв до виконання

_____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Огляд та аналіз вимог до роботи.	02.09.19-19.09.19	
2	Опис сучасних методів та технологій обробки природнього мовлення	22.09.19-10.10.19	
3	Практична реалізація методів та алгоритмів для розпізнавання природнього мовлення	11.10.19-15.11.19	
4	Дослідження та аналіз отриманих результатів	16.11.18-09.12.19	
5	Розробка заходів з охорони праці.	10.12.19-15.12.19	
6	Оформлення пояснювальної записки.	16.12.19-31.12.19	
7	Підготовка та подання магістерської роботи до захисту.	02.01.20-08.01.19	

Студент

_____ (підпис)

Фурса П. С.

_____ (прізвище та ініціали)

Науковий керівник

_____ (підпис)

Білобородова Т. О.

_____ (прізвище та ініціали)

АНОТАЦІЯ

Фурса П. С. Методи та алгоритми розпізнавання природнього мовлення.

Розглянуто та проаналізовано основні проблеми розпізнавання мовлення. Виконано аналіз методів та алгоритмів, що використовуються в галузі розпізнавання природнього мовлення. Розглянуто математичні основи роботи глибоких нейронних мереж. Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж проведена за двома сценаріями та включала наступні етапи: проведення експерименту за першим сценарієм, проведення експерименту за другим сценарієм, оцінка та порівняння результатів експериментів, проведених за першим та другим сценаріями, порівняння результатів експериментів, та винесення висновків по проведеній реалізації методів і алгоритмів розпізнавання природнього мовлення. Оцінка результатів проведена з використанням узагальненого коефіцієнту помилок в словах. Доведено залежність якості розпізнавання природнього мовлення від вихідних даних моделей: кількості файлів для навчання, якості запису аудіо файлів, акценту.

Ключові слова: розпізнавання мовлення, обробка природнього мовлення, мовленнєві дані, метод, рекурентні нейронні мережі, модель, швидкість розпізнавання, коефіцієнт помилок в словах.

АННОТАЦИЯ

Фурса П. С. Методы и алгоритмы распознавания естественного речи.

Рассмотрены и проанализированы основные проблемы распознавания речи. Выполнен анализ методов и алгоритмов, используемых в области распознавания естественного языка. Рассмотрены математические основы работы глубоких нейронных сетей. Практическая реализация методов и алгоритмов распознавания естественной речи реализована с использованием нейронных сетей проведена по двум сценариям и включала следующие этапы: проведение эксперимента по первому сценарию, проведение эксперимента по второму сценарию, оценка и сравнение результатов экспериментов, проведенных по первому и второму сценариям, сравнение результатов экспериментов и выводы по проведенной реализации методов и алгоритмов распознавания естественного языка. Оценка результатов проведена с использованием обобщенного коэффициента

ошибок в словах. Доказана зависимость качества распознавания естественного языка от исходных данных моделей: количества файлов для обучения, качества записи аудио файлов, акцента.

Ключевые слова: распознавание естественного языка, обработка естественного языка, текстовые данные, метод, рекуррентные нейронные сети, модель, скорость распознавания, коэффициент ошибок в словах.

ABSTRACT

Polina Fursa. Methods and algorithms of Nature Language Processing.

The main problems of Nature Language Processing are considered and analyzed. Analysis of methods and algorithms used in the field of natural language recognition was performed. Mathematical basics of deep neural networks were considered. Practical implementation of methods and natural speech recognition algorithms implemented using neural networks is performed in two scenarios, and included the following steps: the experiment in the first scenario, the experiment under the second scenario, evaluation and comparison of the results of experiments conducted on the first and second scenarios, the comparison of the experimental results and the conclusions drawn by the implementation of methods and natural language recognition algorithms. Evaluation of the results using a generalized WER. The dependence of the quality of natural language recognition on the initial data of the models: the number of files for training, the recording quality of audio files, the emphasis was proved.

Keywords: natural language recognition, natural language processing, speech-to-text, methods, recurrent neural network, word error rate.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧОК І СКОРОЧЕНЬ	8
ВСТУП.....	9
РОЗДІЛ 1 АНАЛІЗ ГАЛУЗІ ЗАСТОСУВАННЯ ТА ТЕХНОЛОГІЇ ОБРОБКИ ПРИРОДНЬОГО МОВЛЕННЯ	12
1.1 Основи обробки природнього мовлення	12
1.2 Галузі застосування технології обробки природнього мовлення	12
1.3 Аналіз проблем застосування обробки природнього мовлення	14
1.4 Аналіз методів процесу обробки природнього мовлення.....	15
1.5 Аналіз алгоритмів машинного навчання для обробки природнього мовлення	19
1.5.1 Особливості використання Марковських моделей	19
1.5.2 Аналіз глибоких нейронних мереж	22
1.6 Постановка наукової задачі та обґрунтування методики досліджень.....	23
1.7 Висновки до розділу 1.....	24
РОЗДІЛ 2 АНАЛІЗ ВИКОРИСТАННЯ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ ПРИРОДНЬОГО МОВЛЕННЯ.....	25
2.1 Особливості вхідних мовленнєвих даних	25
2.1.1 Прямо спостережувані властивості.....	25
2.1.2 Особливості n-грам.....	28
2.2 Архітектура мережі.....	30
2.2.1 Пакетна нормалізація для глибоких РНМ.....	31
2.2.2 Визначення архітектури мережі та гіперпараметрів.....	32
2.2.3 Частота згорток	32
2.2.4 Згортка Lookahead і односпрямовані моделі.....	33
2.3.1 Опис РНМ.....	34
2.3.2 РНМ навчання.....	36
2.4 Загальний принцип використання РНМ	37
2.4.1 Акцептор.....	37
2.4.2 Енкодер	38
2.4.3 Перетворювач.....	38
2.5 РНМ як особливості екстракторів	39
2.5.1 Частина мови тегування.....	39

2.6 Висновки до розділу 2.....	42
РОЗДІЛ 3 АНАЛІЗ ТА ОБРОБКА ТЕКСТУ З ВИКОРИСТАННЯМ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ.....	43
3.1 Mozilla Deep Speech	45
3.2 Проведення експерименту з розпізнавання природнього мовлення за першим сценарієм	48
3.2.1 Завантаження навченої моделі.....	48
3.2.2 Запис та підготовка аудіо файлів для тестування моделі.....	49
3.2.3 Тестування моделі.....	49
3.3 Проведення експерименту з розпізнавання природнього мовлення за другим сценарієм	51
3.3.1 Створення бібліотеки аудіо файлів	51
3.3.1.1 Запис аудіо файлів.....	51
3.3.1.2 Підготовка аудіо файлів для навчання DeepSpeech	52
3.3.2 Навчання моделі	54
3.3.3 Тестування моделі.....	55
3.4 Порівняння моделей розпізнавання	57
3.5 Висновки до розділу 3.....	57
РОЗДІЛ 4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ	59
4.1 Аналіз стану умов праці.....	59
4.2 Вимоги до приміщення.....	60
4.3 Вимоги до організації робочого місця	63
4.4 Рекомендації із пожежної профілактики	63
4.5 Мікроклімат	66
4.6 Охорона навколишнього природного середовища.....	67
4.7 Висновки до розділу 4.....	69
ВИСНОВКИ	70
ПЕРЕЛІК ПОСИЛАНЬ.....	71
Додаток А.....	73

ПЕРЕЛІК УМОВНИХ ПОЗНАЧОК І СКОРОЧЕНЬ

БШП	Багатошаровий перецептрон
КПС	Коефіцієнт помилок в словах
ОПМ	Обробка природнього мовлення
ПММ	Прихована Марковська модель
РНМ	Рекурентна нейронна мережа
ШІ	Штучний інтелект
NP	Noun phrase
TF -IDF	term frequency- inverse document frequency
VP	Verb phrase
WER	Word error rate

ВСТУП

Актуальність теми. Обробка природнього мовлення (англ. Natural Language Processing) - перетин машинного навчання й математичної лінгвістики, спрямований на вивчення методів аналізу й синтезу природнього мовлення. Ця область пов'язана з обробкою й аналізом великих обсягів даних природнього мовлення. Зародження цього напрямку пов'язують із появою перших обчислювальних машин і з ідеєю використання машини для вирішення корисних задач, пов'язаних із природнім мовленням, яким розмовляють і пишуть люди.

Сьогодні швидко росте кількість корисних додатків на основі використання NLP [1]:

- пошук (письмовий або усний);
- показ підходящої онлайн реклами;
- автоматичний (або за сприянням) переклад;
- аналіз настроїв для задач маркетингу;
- розпізнавання мовлення й чат-боти,
- голосові помічники (автоматизована допомога покупцеві, замовлення товарів і послуг).

Розуміння і створення мовлення з використанням комп'ютерів є дуже складним завданням. Крім проблем, пов'язаних із неоднозначними й змінними вхідними даними з погано визначеними й взагалі невизначеними наборами правил, природне мовлення проявляє додатковий набір властивостей, які зробили мовлення ще більш складним для обчислювальних підходів, включаючи машинне навчання: дискретне, композиційне й рідкісне.

Мовлення є символічним і дискретним. Основними елементами писемного мовлення є символи. Символи створюють слова, які в свою чергу позначають об'єкти, концепції, події, дії й ідеї. І символи, і слова є дискретними символами. Слова викликають певні ментальні уявлення, але вони також є окремими символами, значення яких є зовнішніми по відношенню до них. Не існує невід'ємного зв'язку між словами, який може бути виведеним із самих символів або з окремих букв, із яких вони складаються.

Також мовлення композиційне: літери утворюють слова, а слова утворюють фрази й речення. Значення фрази може бути більшим, ніж значення окремих слів, які її складають. Таким чином, щоб інтерпретувати текст, потрібно працювати над рівнями букв і слів, або подивитися на довгі послідовності слів, такі як речення, або навіть на весь документ.

Поєднання вищевказаних властивостей призводить до розрідженості даних. Спосіб, при якому слова (дискретні символи) можуть поєднуватися для сформування можливих значень

практично нескінченний. Кількість можливих речень величезна, їх ніколи не можна було перерахувати. Якщо відкрити випадкову книгу, то переважно більшість речень в ній ви не бачили й не чули раніше. Більш того, цілком ймовірно, що багато речень із чотирьох слів, які з'являються в книзі, також для вас є новими. Немає чіткого способу узагальнення одного речення з іншим або визначальної подібності між реченнями, яке не залежить від їх значення.

Сьогодні актуальність напрямку пов'язана перш за все з необхідністю обробляти великі масиви аудіо й текстової інформації, накопиченої людством за останні десятиліття. Коло завдань по обробці природнього мовлення стало набагато ширше [2]:

1. Завдання розпізнавання мовлення, переклад мовлення в текст, із яким зручніше працювати.
2. Зворотне завдання синтезу мовлення.
3. Завдання, пов'язані з інформаційним пошуком, із витяганням інформації із текстів.
4. Завдання класифікації текстів, коли необхідно віднести текст до одного зі заздалегідь відомих класів.
5. Завдання кластеризації текстів - угруповання множини текстів у групи схожих один на одного текстів.
6. Завдання резюмування текстів, коли нам із великого тексту необхідно отримати його стисле уявлення.
7. Завдання, пов'язані з аналізом соціальних мереж, витяганням думок користувачів соціальних мереж.
8. Завдання створення питально-відповідних систем, діалогових систем і завдання машинного перекладу сьогодні також залишаються актуальними.

Об'єкт дослідження: процеси перетворення мовленнєвого сигналу в цифрову інформацію.

Предмет дослідження: методи і алгоритми розпізнавання природнього мовлення та оцінки, що використовуються для визначення ефективності розпізнавання.

Мета і завдання дослідження. Метою дослідження є підвищення якості розпізнавання природнього мовлення за рахунок визначення вагомих факторів, що впливають на якість розпізнавання природнього мовлення з вживанням вузькоспеціалізованої специфічної термінології.

Для досягнення мети дослідження необхідно вирішити такі **завдання**:

- аналіз галузі застосування, методів обробки і алгоритмів розпізнавання природнього мовлення;
- аналіз використання глибоких нейронних мереж для розпізнаванні природнього мовлення;
- практична реалізація обробки та аналізу природнього мовлення з використанням глибоких

нейронних мереж;

- створення моделі з використанням глибоких нейронних мереж для розпізнавання мовленнєвої медичної термінології;

- оцінка ефективності навчання створеної моделі розпізнавання природнього мовлення.

Методи дослідження. Проведені в роботі дослідження засновані на методах глибокого навчання та інтелектуального аналізу даних, що використовувались для розпізнавання природнього мовлення, навчання власної моделі та оцінки ефективності навчання створеної моделі розпізнавання природнього мовлення.

Наукова новизна отриманих результатів:

Подальший розвиток методу розпізнавання природнього мовлення в умовах використання з метою розпізнавання мовленнєвої медичної термінології.

Особистий внесок здобувача полягає у застосуванні глибоких нейронних мереж з метою розпізнавання мовленнєвої медичної термінології.

- оцінка ефективності навчання створеної моделі розпізнавання природнього мовлення, що дозволяють вирішити поставлені задачі. Усі основні результати отримані автором особисто.

Практичне значення одержаних результатів полягає в реалізації обробки та аналізу природнього мовлення з використанням глибоких нейронних. Основні теоретичні результати роботи втілено у наступних реалізаціях та положеннях:

- обґрунтовано використання глибоких нейронних мереж для розпізнавання мовленнєвої медичної термінології.

- реалізовано розпізнавання природнього мовлення з використанням глибоких нейронних мереж для створення моделі розпізнавання мовленнєвої медичної термінології.

Апробація матеріалів дисертації. Основні положення, ідеї, висновки магістерської роботи доповідалися та обговорювалися на III міжнародній конференції «Theoretical and Applied Computer Science and Information Technology – 2019» та форумі «ІТ Ідея 2019».

Публікації. За темою магістерської роботи опубліковано статтю в науковому фаховому виданні України з наступними бібліографічними даними: Derevianchenko V.S., Biloborodova T.O., Skarga-Bandurova I.S., Fursa P.S., Koverha M.O. A survey of open-source speech recognition software for voice actuated control. *Наукові вісті Далівського університету*. Електронне видання. 2019.

Структура та обсяг роботи

Магістерська робота складається зі вступу, 4 розділів, висновків на 70 сторінках, списку використаних джерел з 24 найменувань на 2 сторінках, додатків на 10 сторінках. Загальний обсяг роботи складає 82 сторінок. В магістерській роботі міститься 3 таблиць, 15 рисунків.

РОЗДІЛ 1

АНАЛІЗ ГАЛУЗІ ЗАСТОСУВАННЯ ТА ТЕХНОЛОГІЇ ОБРОБКИ ПРИРОДНЬОГО МОВЛЕННЯ

1.1 Основи обробки природнього мовлення

Обробка природнього мовлення (ОПМ) є галуззю штучного інтелекту, яка займається взаємодією між комп'ютерами й людьми, які використовують природнє мовлення. Кінцева мета ОПМ полягає в тому, щоб читати, розшифровувати, розуміти й осмислювати людські мови. Більшість методів ОПМ засновані на машинному навчанні для вилучення інформації із людських мов [3].

ОПМ тягне за собою застосування алгоритмів для ідентифікації та вилучення правил природнього мовлення, так що неструктуровані мовні дані перетворюються в форму, зрозумілу для комп'ютерів. Коли текст наданий, комп'ютер використовує алгоритми для отримання значення, пов'язаного з кожним реченням, і збору з них необхідних даних. Іноді комп'ютер може погано розуміти сенс речення, що призводить до неясних результатів. Фактично, типова взаємодія між людьми й машинами з використанням обробки природнього мовлення може виглядати наступним чином:

1. Людина розмовляє з машиною
2. Машина захоплює звук.
3. Відбувається конвертація аудіо в текст.
4. Обробка текстових даних.
5. Конвертація даних в аудіо.
6. Машина реагує на людину, відтворюючи аудіофайл.

1.2 Галузі застосування технології обробки природнього мовлення

Для чого використовується ОПМ? Обробка природнього мовлення є рушійною силою наступних загальних додатків [4]:

- додатки для мовного перекладу, такі як Google Translate;

- текстові процесори, такі як Microsoft Word і Grammarly, які використовують ОПІ для перевірки граматичної правильності текстів;
- додатки інтерактивного голосового відповідача (IVR), що використовуються в центрах обробки викликів для відповіді на запити певних користувачів;
- додатки особистого помічника, такі як OK Google, Siri, Cortana і Alexa.
- Простіше кажучи, ОПМ представляє собою автоматичну обробку природного людського мовлення, такого як мова або текст, і, хоча сама концепція цікава, реальна цінність цієї технології виходить із варіантів використання. ОПМ може допомогти з великою кількістю завдань, і області застосування, здається, збільшуються щодня. Наведемо кілька прикладів:
 - ОПМ дозволяє розпізнавати й прогнозувати захворювання на основі електронних медичних карт і мовлення пацієнта. Ця можливість вивчається в умовах здоров'я, які переходять від серцево-судинних захворювань до депресії і навіть шизофренії. Наприклад, Amazon Comprehend Medical - це служба, яка використовує ОПМ для вилучення хворобливих станів, ліків і результатів лікування із заміток пацієнтів, звітів про клінічні випробування й інших електронних медичних карт;
 - організації можуть визначити, що клієнти говорять про послугу або продукт, ідентифікуючи й витягуючи інформацію з таких джерел, як соціальні мережі. Цей аналіз настроїв може надати багато інформації про клієнтів і їх рішення;
 - винахідник у IBM розробив когнітивного помічника, який працює як персоналізована пошукова система, дізнаючись все про вас, а потім нагадує вам ім'я, пісню або щось, що ви не можете згадати в той момент, коли вам це потрібно;
 - такі компанії, як Yahoo і Google, фільтрують і класифікують ваші електронні листи за допомогою ОПМ, аналізуючи текст в електронних листах, які проходять через їх сервери, і блокуючи спам ще до того, як вони потрапляють у вашу поштову скриньку;
 - щоб допомогти ідентифікувати фейкові новини, група ОПМ в Массачусетському технологічному інституті розробила нову систему для визначення, чи є джерело точним або політично необ'єктивним, визначаючи, чи можна довіряти джерелу новин чи ні;
 - Amazon Amazon і Apple Siri є прикладами інтелектуальних голосових інтерфейсів, які використовують ОПМ для відповіді на голосові запити й роблять все, наприклад, знаходять конкретний магазин, повідомляють нам прогноз погоди, пропонують кращий маршрут до офісу або вмикають світло в будинку;
 - розуміння того, що відбувається і про що говорять люди, може бути дуже цінним для фінансових трейдерів. ОПМ використовується для відстеження новин, звітів, коментарів про

можливе злиття між компаніями, потім все може бути включено в торговий алгоритм для отримання величезних прибутків;

– ОПМ також використовується як на етапі пошуку, так і при відборі талантів, виявляючи навички потенційних співробітників, а також виявляючи потенційних клієнтів, перш ніж вони стануть активними на ринку праці;

– за підтримки технології IBM Watson NLP компанія LegalMation розробила платформу для автоматизації рутинних судових завдань і допомагає юридичним групам економити час, скорочувати витрати й зміщати стратегічні цілі.

ОПМ особливо активно розвивається в сфері охорони здоров'я [5]. Ця технологія покращує надання медичної допомоги, діагностику захворювань і знижує витрати, в той час як медичні організації все частіше використовують електронні медичні картки. Той факт, що клінічна документація може бути поліпшена, означає, що пацієнти можуть бути краще зрозумілі й отримають користь завдяки кращій охороні здоров'я. Мета повинна полягати в тому, щоб оптимізувати їх досвід, і кілька організацій уже працюють над цим.

Такі компанії, як Winterlight Labs, роблять величезні поліпшення в лікуванні хвороби Альцгеймера шляхом моніторингу когнітивних порушень у мовленні й можуть також підтримувати клінічні випробування й дослідження для широкого спектра розладів центральної нервової системи [6]. Дотримуючись аналогічного підходу, Стенфордський університет розробив Woebot, терапевта-чатбота, для допомоги людям із тривогою й іншими розладами.

1.3 Аналіз проблем застосування обробки природнього мовлення

Обробка природнього мовлення вважається складною проблемою в інформатиці. Основні недоліки, з якими ми зустрічаємося в наші дні при використанні ОПМ, пов'язані з тим, що людське мовлення дуже складне [7]. Правила, які диктують передачу інформації із використанням природних мов, нелегко зрозуміти комп'ютерам. Деякі з цих правил можуть бути високорівневими й абстрактними; наприклад, коли людина використовує саркастичне зауваження для передачі інформації. Повне розуміння людської мови вимагає розуміння як слів, так і того, чим поняття зв'язані, щоб доставити намічене повідомлення. У той час як люди можуть легко оволодіти мовою, неоднозначність і неточні характеристики природних мов ускладнюють реалізацію ОПМ для машин.

Одна з головних проблем полягає у відсутності універсального рішення роботи з налаштуваннями з низьким обсягом даних (мови з низьким рівнем ресурсів, діалекти (включаючи

текстові «діалекти» в соціальних мережах), домени і т. д.). Існує теорія, що серед мов є універсальні спільності, які можуть бути використані універсальною мовною моделлю. Завдання полягає в тому, щоб отримати достатньо даних і обчислень для навчання такої мовної моделі. Це тісно пов'язано з недавніми зусиллями щодо навчання міжмовної моделі мови Transformer і вбудовуванню міжмовних речень. Міжмовне вбудовування слів ефективно з точки зору вибірки, оскільки для нього потрібний переклад тільки пари слів або навіть одномовні дані. Вбудовування досить добре згладжує місця для вбудовування слів, щоб виконувати важкі завдання, такі як класифікація тем, але не дозволяють виконувати більш легкі завдання, такі як машинний переклад. Проте недавні зусилля демонструють, що ці вбудовування утворюють блокування речень для неконтрольованого машинного перекладу. Більш складні моделі для задач більш високого рівня, таких як відповіді на питання вимагають тисячі прикладів для навчання. Перенесення завдань, що вимагають реального розуміння природнього мовлення, з мов із високими ресурсами на мови з низьким рівнем ресурсів все ще залишається дуже складним завданням.

Інша проблема полягає в роботі ОПМ із великими за об'ємом текстами або з декількома документами. Міркування з великим контекстом тісно пов'язане з розумінням природнього мовлення (natural language understanding) і вимагає значного розширення нинішніх систем, поки вони не зможуть читати цілі книги й сценарії фільмів. Сучасні моделі можуть багато зробити, якщо їх масштабувати для роботи з великою кількістю даних і великою кількістю обчислень. Таким чином, більш корисним напрямком є розробка методів, які можуть більш ефективно представляти контекст і краще відстежувати відповідну інформацію під час читання документа.

1.4 Аналіз методів процесу обробки природнього мовлення

Процес розуміння і маніпулювання мовою надзвичайно складний, і з цієї причини зазвичай використовують різні методи для вирішення різних завдань, перш ніж зв'язати все разом. Мови програмування, такі як Python або R, широко використовуються для виконання цих методів.

Обробка природнього мовлення включає в себе безліч різних методів інтерпретації людської мови, від статистичних та машинних методів навчання до заснованих на правилах і алгоритмічних підходів. Нам потрібен широкий спектр підходів, тому що текстові і голосові дані сильно відрізняються, як і практичне використання.

Основні завдання ОПМ включають у себе токенізацію й синтаксичний аналіз, лематизацію/визначення стемінгу, тегування частин мови, виявлення мови й ідентифікацію семантичних відношень [8]. Загалом, завдання ОПМ розбивають мову на більш короткі, елементарні

частини, намагаються зрозуміти відношення між частинами й дослідити, як частини працюють разом, щоб створити зміст. У всіх цих випадках головна мета полягає в тому, щоб взяти необроблений мовний ввід і використовувати лінгвістику й алгоритми для перетворення або збагачення тексту таким чином, щоб він приносив велику цінність. Робота любого компілятора ОПМ складається з основних трьох етапів (рис. 1.1).

На першому етапі виконується лексичний аналіз, у ході якого з тексту програми послідовно зчитуються лексеми, або токени - мінімальні неподільні одиниці мови. До таких неподільних одиниць відносяться слова, роздільники, знаки алгебраїчних операцій і т. д. Отриманий ланцюг токенів записується компілятором у спеціальну таблицю, в якій для кожного токена вказуються його код, тип (число, рядок, оператор, ідентифікатор, роздільник і т.д.), деякі інші характеристики. На другому етапі виконується синтаксичний аналіз, у ході якого проводиться розбір отриманої послідовності токенів. В результаті розбору формується абстрактне синтаксичне дерево, що є проміжною і, як випливає з назви, абстрактною формою представлення програми. Абстрактне синтаксичне дерево не є буквальним відображенням коду, яке воно представляє. Воно містить тільки необхідні дані для запису операцій. На третьому етапі виконується генерація машинного коду. Спочатку абстрактне синтаксичне дерево перетворюється в лінеаризоване абстрактне представлення, наприклад, у трьохадресний код, кожна команда якого складається не більше ніж із трьох операндів. Потім лінеаризоване представлення перекладається в машинний код.

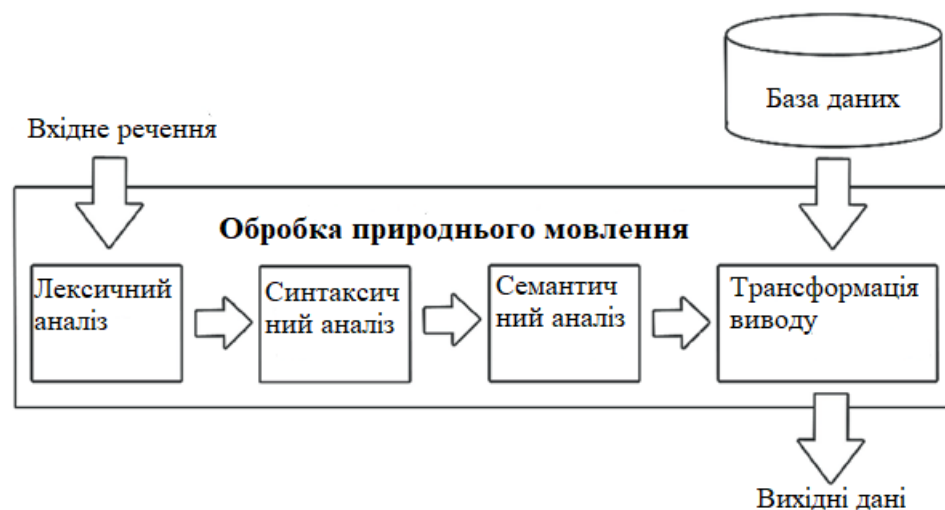


Рисунок 1.1 – Процес обробки природнього мовлення

Синтаксичний і семантичний аналізи є основними методами, використовуваними для виконання завдань обробки природньої мови. Далі наведені детальні описи методів ОПМ:

1. Попередня обробка тексту. Попередня обробка тексту переводить текст, написаний природнім мовленням, у формат зручний для подальшої роботи. Попередня обробка складається з різних етапів, які можуть відрізнятися в залежності від завдання й реалізації. Нижче наведено один набір етапів із усіх можливих:

- переклад усіх букв у тексті в нижній або верхній регістри;
- видалення цифр (чисел) або заміна на текстовий еквівалент (зазвичай використовуються регулярні вирази);
- видалення пунктуації. Зазвичай реалізується як видалення з тексту символів із заздалегідь заданого набору;
- видалення символів пробілів (whitespaces);
- токенизація (зазвичай реалізується на основі регулярних виразів);
- видалення стоп слів;
- стемінг;
- лематизації;
- векторизація.

2. Стемінг. Кількість коректних словоформ, значення яких схожі, але написання відрізняються суфіксами, приставками, закінченнями й іншим, дуже велика, що ускладнює створення словників і подальшу обробку. Стемінг дозволяє привести слово до його основної форми. Суть підходу в знаходженні основи слова, для цього з кінця й початку слова послідовно відрізають його частини. Правила відсікання для стемера створюються заздалегідь, і найчастіше являють собою регулярні вирази, що робить даний підхід трудомістким, так як при підключенні чергової мови потрібні нові лінгвістичні дослідження. Другим недоліком підходу є можлива втрата інформації при відрізанні частин, наприклад, можна втратити інформацію про частини мови.

3. Лематизація. Даний підхід є альтернативою стемінга. Основна ідея полягає в приведенні слова до словникової форми - леми. Наприклад для російської мови: для іменників - називний відмінок, однина; для прикметників - називний відмінок, однина, чоловічий рід; для дієслів, дієприкметників, дієприслівників - дієслово в інфінітиві недосконалого виду.

4. Векторизація. Більшість математичних моделей працюють у векторних просторах великих розмірностей, тому необхідно відобразити текст у векторному просторі. Основним підходом є мішок слів (bag-of-words): для документа формується вектор розмірності словника, для кожного слова виділяється своя розмірність, для документа записується ознака наскільки часто слово зустрічається в ньому, отримуємо вектор. Найбільш поширеним методом для обчислення ознаки є TF-IDF (TF -

частота слова, term frequency, IDF - зворотна частота документа, inverse document frequency). TF обчислюється, наприклад, лічильником входження слова. IDF зазвичай обчислюють як логарифм від числа документів у корпусі, розділений на кількість документів, де це слово представлено. Таким чином, якщо якесь слово зустрілося в усіх документах корпусу, то таке слово не буде нікуди додано. Плюсами мішка слів є проста реалізація, однак даний метод втрачає частину інформації, наприклад, порядок слів. Для зменшення втрати інформації можна використовувати мішок N-грам (додавати не тільки слова, а й словосполучення), або використовувати методи векторних уявлень слів це, наприклад, дозволяє знизити помилку на словах із однаковими написаннями, але різними значеннями.

5. Дедублікація. Так як кількість схожих документів може бути великою, необхідно позбавлятися від дублікатів. Так як кожен документ може бути представлений як вектор, то ми можемо визначити їх близькість, взявши косинус або іншу метрику. Мінусом є те, що для великих наборів повний перебір по всіх документах буде неможливий. Для оптимізації можна використовувати локально-чутливий хеш, який помістить близько схожі об'єкти.

6. Семантичний аналіз. Семантичний (смісловий) аналіз тексту - виділення семантичних відношень, формуванні семантичної мережі. У загальному випадку семантична мережа є графом, що відображає бінарні відношення між двома вузлами - смисловими одиницями тексту. Глибина семантичного аналізу може бути різною, а в реальних системах найчастіше будується тільки синтаксико-семантичне відношення тексту або окремих речень. Семантичний аналіз застосовується в задачах аналізу тональності тексту (sentiment analysis), Наприклад, для автоматизованого визначення позитивності відгуків.

7. Розпізнавання іменованих сутностей і витяг відношень. Іменовані сутності - об'єкти з тексту, які можуть бути віднесені до однієї із заздалегідь заявлених категорій (наприклад, організації, особи, адреси). Ідентифікація посилань на подібні сутності в тексті є завданням розпізнавання іменованих сутностей. Визначення семантичних відношень між іменованими сутностями або іншими об'єктами тексту, є завданням вилучення відношень. Ці два підходи застосовуються в багатьох задачах, наприклад, витяг синонімів із тексту, автоматичній побудові онтологій і реалізовані в багатьох працюючих системах, наприклад, NELL і Snowball.

8. Використання N-грам. N-грами - послідовність з n елементів. В ОПМ N-грами використовуються для побудови ймовірнісних моделей, задач схожості текстів, категоризації тексту й мови. Побудувавши N-грамну модель можна визначити ймовірність вживання заданої фрази в тексті. N-грамна модель розраховує ймовірність останнього слова N-грами, якщо відомі всі попередні, при цьому вважається, що ймовірність появи кожного слова залежить тільки від

попередніх слів. Використання N-грам застосовується в задачі виявлення плагіату. Текст розбивається на кілька фрагментів, представлених N-грамами. Порівняння N-грам один із одним дозволяє визначити ступінь подібності документів. Аналогічним способом можна вирішувати завдання виправлення орфографічних помилок, підбираючи слова кандидати для заміни.

1.5 Аналіз алгоритмів машинного навчання для обробки природнього мовлення

Процес обробки природнього мовлення включає декілька етапів з використанням методів, алгоритмів машинного навчання для вилучення інформації з людських мов. До цих методів відносяться метод Баєса, метод опорних векторів, Марковські моделі й нейронні мережі.

1.5.1 Особливості використання Марковських моделей

В багатьох джерелах вказується, що розпізнавання мовлення методом Маркова є найточнішим для розпізнавання слова або рядка, які можна розпізнати зі 100-відсотковим успіхом. Основні кроки, які виконуються в розпізнаванні слів/символів:

- сканування зображення тексту;
- виявлення рядків тексту;
- поділ на світлі й темні ділянки з метою виявлення кожної буквенної літери чи цифри;
- виявлення підключених компонентів
- переклад зображення символу в коди символів;
- порівняння символів у відсканованому файлі зображень із символами вивченого набору.

Існують певні фактори, які можуть вплинути на розпізнавання символів і можуть спричинити труднощі в процесі розпізнавання. Нижче наведено декілька факторів:

- морщинисті, рвані або іншим чином пошкоджені зображення;
- знебарвлений або спотворений текст
- надрукований низькоконтрастним або кольоровим чорнилом текст
- текст написаний нестандартними шрифтами або людським рукописом, або надруковано на певних типах паперу, які зменшують чіткість та контраст між фоном і переднім планом, у результаті сканування
- сканування з низьким пікселем.

Одним із основних факторів може бути сканування документа з невідповідним dpi. Документ, відсканований при 200 dpi, може призвести до поганих результатів, тоді як документ відсканований при dpi, що вище 600 може виділити більше місця для зберігання.

Процес розпізнавання символів повністю залежить від наявності вивченого набору. Навчений набір - це контейнер, у якому для певного символу доступна множина розмірів, стилів та шрифтів. Чим більше словник даних і чим ефективніший обраний метод, тим більше успіху, який можна досягти. Щоб зробити вивчений набір, текстовий файл потрібно створити з різними шрифтами, розмірами та стилями.

Приховані марковські моделі разом із програмуванням використовується для розпізнавання частин мови й символів. Існують дві фундаментальні методики, що демонструють розробки Маркова: для кожного символу й для кожного слова. Марковська модель - це статистичний інструмент, що використовується для моделювання генеративних послідовностей, який характеризується набором спостережуваних послідовностей. Прихована модель Маркова - це стохастична система, яка бере ланцюг Маркова з невідомими параметрами, й завдання полягає в пошуку прихованих параметрів від спостережуваних параметрів [9]. Безперервні моделі виконуються під час обробки сигналу, інакше дискретна ПММ виконується при обробці зображення.

У звичайній марковській моделі за станом можна спостерігати, тому ймовірності переходів - єдиний параметр. У прихованій марковській моделі можна спостерігати лише за змінними, на які впливає даний стан (рис. 1. 2). Кожний стан має ймовірносний розподіл серед усіх можливих вихідних значень. Тому послідовність символів, згенерована ПММ, дає інформацію про послідовність станів. На рис. 1.2 x позначає скриті стани, y - спостережувані результати, a - ймовірність переходів, b - ймовірність результатів

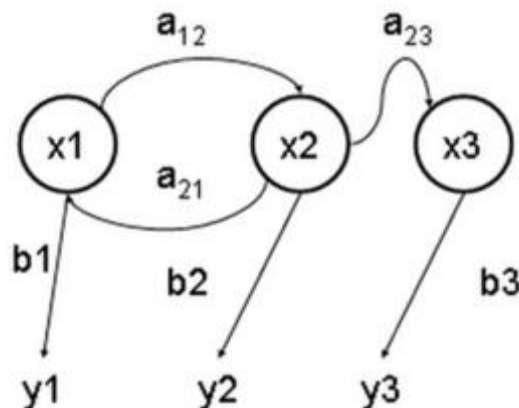


Рисунок 1.2 - Діаграма переходів у прихованій марковській моделі (приклад)

Діаграма (рис.1.3) показує загальну структуру ПММ. Овали є змінними з випадковим значенням. Випадкова змінна $x(t)$ являє собою значення прихованої змінної у момент часу t . Випадкова змінна $y(t)$ - це значення спостережуваної змінної у момент часу t . Стрілки на діаграмі символізують умовні залежності.

З діаграми видно, що значення прихованої змінної $x(t)$ у момент часу t залежить тільки від значення прихованої змінної $x(t-1)$ у момент $t-1$. Це називається властивістю Маркова. Хоча в той же час значення спостережуваної змінної $y(t)$ залежить тільки від значення прихованої змінної $x(t)$ обидві в момент часу t .

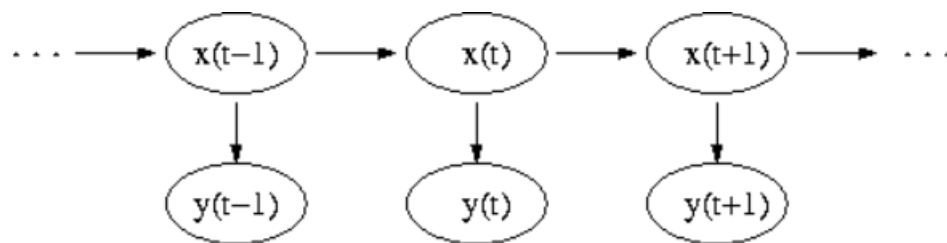


Рисунок 1.3 – Загальна структура ПММ

Імовірність побачити послідовність $Y = y(0), y(1), \dots, y(L-1)$ довжини L дорівнює:

$$P(Y) = \sum_x P(Y|X)P(X), \quad (1.1)$$

де сума пробігає по всім можливим послідовностям прихованих вузлів $X = x(0), x(1), \dots, x(L-1)$. Метод підрахунку повним перебором значень $P(Y)$ - дуже трудомісткий для багатьох завдань із реального життя в силу того, що кількість можливих послідовностей прихованих вузлів дуже велике.

Існують три основні завдання, пов'язані з ПММ:

- алгоритм прямого-зворотного ходу: надані параметри моделі й послідовність, потрібно обчислити ймовірність появи цієї послідовності (дозволяє вирішити задачу).
- алгоритм Вітербі: надані параметри моделі, потрібно визначити найбільш підходящу послідовність прихованих вузлів, що найбільш точно описує дану модель (допомагає при вирішенні даного завдання).
- алгоритм Баума-Велша: надана вихідна послідовність (або кілька) з дискретними значеннями, потрібно навчити ПММ на даному виході.

1.5.2 Аналіз глибоких нейронних мереж

Нейронні мережі забезпечують потужний механізм навчання, який потрібен для вирішення проблем використання ОПМ [10]. Основним компонентом у нейронних мережах для використання мовлення є вбудований шар і відображення дискретних символів у безперервних векторах простору відносно низької розмірності. При вбудовуванні слів вони перетворюються з ізольованих окремих символів в математичні об'єкти, якими можна оперувати. Зокрема, відстань між векторами можна прирівняти до відстані між словами, полегшуючи узагальнення відношення одного слова до іншого. Це представлення слів в якості векторів вивчається мережею як частина процесу навчання. Йдучи вгору по ієрархії, мережа також вчиться комбінувати вектори слів породжуючи прогнозування. Ця здатність у деякій мірі полегшує проблеми дискретності й нестачу даних. Існує два основних види архітектури нейронної мережі, які можна комбінувати в різних способах: мережі прямої передачі та рекурентні/рекурсивні мережі.

Нейромережі прямого розповсюдження, зокрема багат шарові перцептрони (БШП), дозволяють працювати з входами фіксованого розміру, або зі змінними входами довжин, в яких можна знехтувати порядок елементів. Подаючи мережу наборам вхідних компонентів, вона вчиться їх комбінувати за змістом. БШП можуть використовуватися, якщо раніше використовувалася лінійна модель. Нелінійність мережі, а також можливість легко інтегрувати заздалегідь навчені слова, призводять до вищої точності класифікації. Згорткові мережі прямого зв'язку - це спеціалізовані архітектури, які відмінно справляються з вилученням локальних зразків у даних: вони подають вхідні сигнали довільного розміру, й здатні витягувати сенс із локальних шаблонів, які чутливі до порядку слів, незалежно від того, де вони з'являються. Вони дуже добре працюють для визначення орієнтовних фраз або ідіом у довгих реченнях або документах.

Рекурентні нейронні мережі (RNN) є спеціалізованими моделями для послідовних даних. Ці мережеві компоненти приймають в якості вхідних даних послідовність елементів і створюють вектор фіксованого розміру, який підсумовує цю послідовність. Оскільки «підсумовування послідовності» означає різні речі для різних завдань (тобто, інформація, необхідна для відповіді на питання про сенс речення, відрізняється від інформації, необхідної для відповіді на питання про її грамотність), рекурентні нейромережі нечасто використовуються в якості автономного компонента, й їх сила полягає в тому, щоб бути компонентами, що навчаються, які можуть бути підключені до інших мережевих компонентів і навчені працювати з ними разом. Наприклад, вихідні дані рекурентної нейромережі можуть бути передані до нейромережі прямого поширення, яка буде намагатися передбачити якесь значення. Рекурентна мережа використовується в якості вхідного

трансформатора, який навчений виробляти змістовні представлення для нейромережі прямого поширення, яка буде працювати на ньому. Рекурентні мережі є дуже вражаючими моделями послідовностей і, можливо, є найбільш придатним варіантом нейронних мереж для обробки мовлення. Вони дозволяють відмовитися від припущення Маркова, яке було поширене в ОПМ протягом десятиліть, і розробки моделей, які можуть спиратися на цілі речення, беручи до уваги порядок слів, коли це необхідно, і не страждати від проблеми статистичної оцінки, що впливають з рідкісності даних. Ця здатність приводить до вражаючого успіху в мовному моделюванні, виконується завдання прогнозування ймовірності наступного слова в послідовності (або, що еквівалентно, ймовірність послідовності). Рекурсивні мережі розширюють рекурентні мережі від послідовностей до дерев.

Багато проблем природнього мовлення структуровані, що вимагає розробки комплексу вихідної структури або дерева. Моделі нейронної мережі можна пристосувати, що є необхідним, шляхом адаптації відомих алгоритмів структурованого прогнозування для лінійних моделей або з використанням нових архітектур, таких як моделі sequence-to-sequence (кодер-декодер). Такі моделі лежать в основі сучасного машинного перекладу. Нарешті, багато завдань прогнозування мовлення пов'язані один із одним у тому сенсі, що знання того, як виконати одне з них, допоможе навчитися виконувати інші. Крім того, в той час як у нас може бути нестача контрольованих (позначених) навчальних даних, у нас є достатній запас необробленого тексту (немаркованих даних). Чи можемо ми вчитися на пов'язаних завданнях і анотованих даних? Підходи нейронної мережі надають захоплюючі можливості для багатозадачного навчання та напівавтоматичне навчання.

1.6 Постановка наукової задачі та обґрунтування методики досліджень

Результати проведеного аналізу алгоритмів і методів, пов'язані з розпізнаванням природнього мовлення показали, що на сьогодні існують небагато подібних розробок, оскільки існує багато невирішених запитань у галузі штучного інтелекту. Аналоги мають кожен з поставлених завдань, як окреме вирішення, але не є оптимальним, оскільки рішення всіх задач не виконуються одночасно.

В опрацьованих публікаціях наведені алгоритми й методи використовуються в платних додатках і програмах. Кожна робота гарно описує мету й галузь застосування використаного метода. Сучасні дослідження зосереджені на застосуванні нейронних мереж для розпізнавання природнього мовлення, які при навчанні на достатніх об'ємах даних показують якісний результат розпізнавання природнього мовлення. Але процес отримання якісної моделі для розпізнавання з

використанням нейронних мереж вивчений недостатньо. Недослідженим є етап навчання моделі з використанням специфічної термінології та якість її подальшого використання для розпізнавання специфічних фраз.

Виходячи з вищевказаного, були поставлені наступні задачі:

1. аналіз галузі застосування, методів обробки і алгоритмів розпізнавання природнього мовлення;
2. аналіз використання глибоких нейронних мереж для розпізнаванні природнього мовлення;
3. практична реалізація обробки та аналізу природнього мовлення з використанням глибоких нейронних мереж;
4. створення моделі з використанням глибоких нейронних мереж для розпізнавання мовленнєвої медичної термінології;
5. оцінка ефективності навчання створеної моделі розпізнавання природнього мовлення.

Для вирішення поставлених задач доцільно використовувати глибокі рекурентні нейронні мережі.

1.7 Висновки до розділу 1

В процесі аналізу були знайдені статті на обрану тему, але кожна з робіт не мала на меті навчання моделі з використанням специфічної термінології та якість її подальшого використання для розпізнавання специфічних фраз.

Результатом аналізу було виявлено, що використання глибоких рекурентних нейронних мережа є більш ефективним для розпізнавання природнього мовлення.

РОЗДІЛ 2

АНАЛІЗ ВИКОРИСТАННЯ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ДЛЯ РОЗПІЗНАВАННЯ ПРИРОДНЬОГО МОВЛЕННЯ

2.1 Особливості вхідних мовленнєвих даних

Припустимо, що вектори x задані. При обробці мовлення вектори x отримані з мовленнєвих даних, щоб відобразити різні лінгвістичні властивості тексту. Зіставлення мовленнєвих даних із реальними векторами називається виділенням ознак або навчанням ознак, і виконується функцією об'єкта. Вибір правильних функцій є невід'ємною частиною успішного проекту машинного навчання. У той час як глибокі нейронні мережі зменшують велику потребу в проектуванні функцій, найбільш прийнятий набір основних функцій ще потрібно визначити. Це особливо важливо для мовленнєвих даних, які надходять у вигляді послідовності дискретних символів. Ця послідовність потребує в перетворенні будь-яким чином у числовий вектор, неочевидним способом.

2.1.1 Прямо спостережувані властивості

Особливості для окремих слів. Коли цільовою сутністю є слово поза контекстом, основним джерелом інформації є літери, що містять слово та їх порядок, також властивості, отримані з них, такі як довжина слова, орфографічна форма слова.

Також можна поглянути на слово стосовно зовнішніх джерел інформації.

Леми й основи. Лематизація - це лінгвістично визначений процес, який може не спрацювати для форм, які не в лексиконі лематизації, або неправильно написані. Лема слова може бути неоднозначною, і лематизація більш точна, коли слово дано в контексті. Більш грубий процес, ніж лематизація, який може працювати на будь-якій послідовності букв, називається стемінгом. Стемер відображає послідовності слів більш короткими послідовностями, заснованими на деяких специфічних для мови евристичних, таких як відображення в тій же послідовності. Важливо, що результатом визначення не обов'язково має бути правильне слово, можуть бути зображення й рисунки. Існують різні стемери з різними рівнями агресивності.

Лексичні ресурси [11]. Додатковим джерелом інформації про словоформи є лексичні ресурси. По суті, це словники, які призначені для програмного доступу машин, а не для зчитування людьми.

Лексичний ресурс зазвичай буде містити інформацію про слова, пов'язуючи їх із іншими словами й/або надаючи додаткову інформацію.

Розподільна інформація. Можливості використання лексичних ресурсів для введення знань у вектори розподілу, які отримані з алгоритмів нейронної мережі.

Особливості мовлення – це процес, коли розглядаються речення, абзац або документ, ознаками, що спостерігаються є кількість і порядок букв та слів.

Вага (значимість) [12]. Як і раніше, можна інтегрувати статистику на основі зовнішньої інформації, орієнтуючись на прикладі слів, які зустрічаються в документі багато разів, але нечасто зустрічаються щодо зовнішнього набору документів. При використанні моделі «торба слів», зазвичай використовується статистичний показник TF-IDF, використовується для оцінки важливості слів у контексті документа. Розглянемо документ d , який є частиною великого набору D . Замість того, щоб представляти кожне слово w в d по його нормованій кількості в документі $\frac{\#_d(w)}{\sum_{w \in d} \#_d(w)}$ (частота слова), значимість оберненої частоти документа $\frac{\#_d(w)}{\sum_{w \in d} \#_d(w)} * \log \frac{|D|}{|\{d \in D: w \in d\}|}$. Другий показник – це обернена частота документа: обернене з числа окремих документів у наборі, до якого це слово походить. На цьому наголошують слова, які є відмінними від поточного тексту. Крім слів, можна також подивитися на послідовні пари або трійки слів. Вони називаються n-грамами.

Особливості слів у контексті. При розгляді слова в реченні або документі, безпосередньо спостерігаються особливості слова такі як його позиція в реченні, а також слова або літери, що оточують його. Слова, які ближчі до цільового слова, часто більш інформативні про це слово, ніж ті, що знаходяться далі один від одного.

Вікна. Часто, прийнято фокусуватися на безпосередньому контексті слова, розглядаючи вікно, що оточує його (тобто слова з кожного боку, з типовими значеннями k , рівними 2, 5 і 10), а також прийняти особливості, щоб бути тотожними словами в межах вікна.

Фіксований розмір вікна дає можливість розслабитися, в моделі «торба слів» цей порядок не має значення, і враховує відносне положення слів у вікні. Це призводить до відносного положення, такого як «слово X з'явилося два слова зліва від цільового слова».

Позиція. Крім контексту слова, може цікавити його абсолютне положення в реченні. Можуть бути такі функції, як «цільове слово - 5-е слово в реченні» або бін-версія, яка вказує на грубіші категорії: з'являється воно в перших 10 словах, між словами 10 і 20 і т. д.

Особливості словесних відношень. При розгляді двох слів у контексті, крім кожної позиції слова, що оточують їх, також можна подивитися на відстань між словами й ідентичності слів, які з'являються між ними.

Хоча мовленнєві властивості тексту не спостерігаються безпосередньо з поверхневих форм слів у реченнях і їх порядку, вони можуть бути виведені з рядка речення з різним ступенем точності. Існують спеціалізовані системи для передбачення частин мови, синтаксичних дерев, семантичних ролей, дискурсивних відносин і інших мовленнєвих властивостей із різним ступенем точності, і ці передбачення часто служать хорошими характеристиками для подальших завдань класифікації.

Розглянемо деякі форми мовних анотацій. Розглянемо речення *follow the instructions here*. Один рівень анотації призначає до кожного слова його частину мови:

follow	the	instructions	here
Verb	Det	Noun	Adverb

Піднімаючись далі по ланцюжку, позначаємо синтаксичні кордони, вказуючи на те, що *follow* це дієслово (Verb).

[VP follow] [NP the instructions] [NP here]

Більш глобальна синтаксична структура - це фразово-структурне дерево (рис. 2.1):

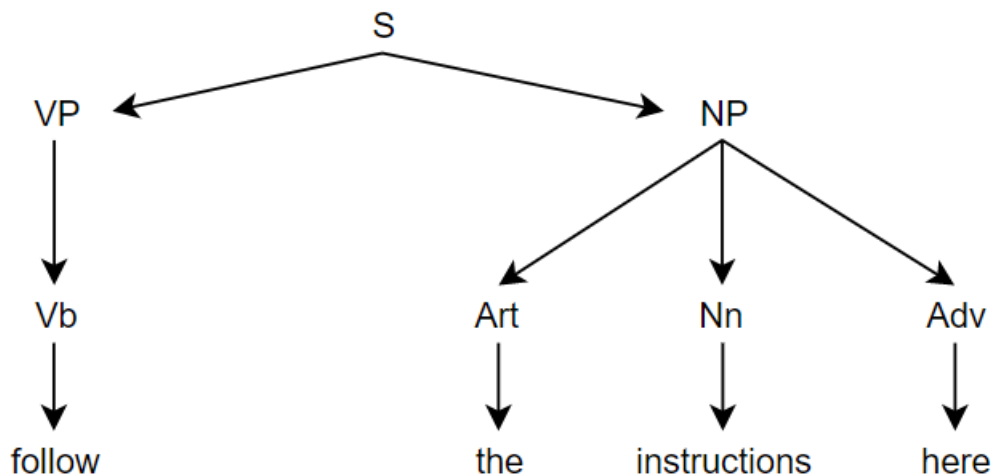


Рисунок 2.1 - Фразово-структурне дерево

Дерева виборчих груп є вкладеними, позначені дугою над реченням, що вказує на ієрархію синтаксичних одиниць: іменникова фраза *the instructions here* складається з іменникової фрази *the instructions* та прийменникової фрази (англ. PP) *here*. Маючи *the instructions*, вкладений під VP, а не під NP, прийменник *here* сигналізує, що за допомогою прийменника модифікується дієслово *follow*.

Інший вид синтаксичної анотації - це дерево залежності (рис. 2.2). За синтаксисом залежності кожне слово в реченні є модифікатором іншого слова, яке називається його *головою*. Кожне слово в реченні очолює інше словосполучення, крім основного слова, як правило, дієслова, яке є коренем речення і очолює спеціальний вузол «корінь».

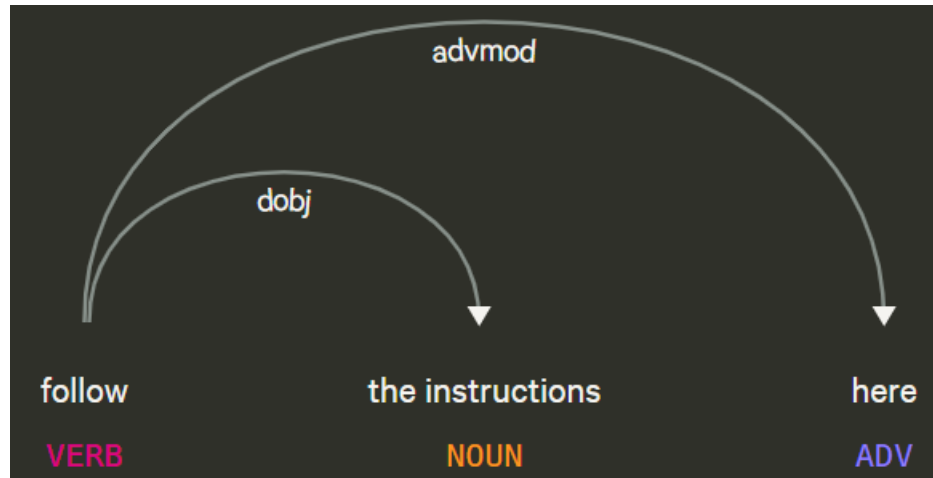


Рисунок 2.2 – Дерево залежності речення

У той час як дерева вибраних груп роблять явне групування слів у словосполучення, дерева залежностей явно змінюють відношення модифікації та зв'язки між словами. Слова, які знаходяться далеко за поверхневою формою речення, можуть бути близькими за своїм деревом залежності.

Відношення залежності синтаксичні: вони стосуються структури речення. Інші види відношень більш семантичні. Наприклад, модифікатор дієслова *follow* також називається аргументом дієслова. Синтаксичне дерево чітко позначає, що *the instructions* є аргументом відкритості, але дерево не показує смислово-рольові аргументи відносно дієслова *follow*.

2.1.2 Особливості n-грам

Особливий випадок комбінації ознак є те, що n-грами - послідовності слів заданої довжини. Слово-біграми, а також триграми (послідовності з трьох букв або слів) також поширені. Крім того, 4 грами і 5 грами іноді використовуються для букв, але рідко для слів через невирішеність існуючих задач.

Двонаправлені РНМ ще більш узагальнюють концепцію n-грам і можуть бути чутливими до інформативних n-грам різної довжини, а також до n-грам із пробілами в них.

Використання представлення слів за допомогою векторів. Вважається, що кожному слову присвоєно цілочисельний індекс, і використовуються символи такі як w або w_i для позначення слова, та його індексу. $E_{[w]}$ - це рядок в E , відповідний слову w . Іноді використовується w , w_i для позначення векторів, відповідних w і w_i .

Подібність слів. Нейронна мережа повинна обчислити схожість між двома словами, враховуючи заздалегідь навчені вектори вбудовування слів і використовуючи функцію подібності по векторам $\text{sim}(u, v)$. Поширений та ефективний вибір для розрахунку подібності векторів є *косинус подібності*, відповідає косинусу кута між векторами:

$$\text{sim}_{\cos}(u, v) = \frac{u \cdot v}{\|u\| \|v\|} \quad (2.1)$$

Коли вектори u і v мають одиничну довжину, схожість косинусів зменшується до скалярного добутку. Працювати зі скалярними добутками дуже зручно, в обчислювальному відношенні, і це є загальним для нормалізації вкладення матриці таким чином, що кожен рядок має одиничну довжину.

Пошук подібних слів. При нормалізованій матриці вкладень косинус подібності між двома словами w_1 і w_2 визначається як:

$$\text{sim}_{\cos}(w_1, w_2) = E_{\{w_1\}} * E_{\{w_2\}}, \quad (2.2)$$

де k - це число найбільш відмінних одне від одного слів.

Нехай $w = E_{[w]}$ вектор відповідний слову w . Подібність із усіма іншими словами можна обчислити за допомогою множення матриці на вектор $s = E * w$. Результатом s є вектор подібності, де $s_{[i]}$ - це подібність w з i -м словом е словнику (i -й рядок в E). k , найбільш схожі слова, можна витягти знаходячи індекси, відповідні k старшим значенням у s .

Подібність із групою слів. Можливо, потрібно буде знайти найбільш подібне слово для групи слів. Ця необхідність виникає, коли є список пов'язаних слів, і хочеться розширити його (наприклад, є список імен генів, і є бажання знайти імена додаткових генів). Інший випадок використання - це коли потрібно, щоб подібність була в даному сенсі слова. Створивши список слів, пов'язаних із цим сенсом, можна направити запит подібності в цьому сенсі. Є багато способів визначити подібність елемента з групою, тут береться визначення середньої подібності з елементами в групі, тобто заданої

групою слів $w_{1:k}$ визначається його подібність зі словом w як: $\text{sim}(w, w_{1:k}) = \frac{1}{k} \sum_i^k \text{sim}_{\cos}(w, w_i)$.

Завдяки лінійності, обчислення середнього косинуса подібності, від групи слів до всіх інших, слова можуть бути знайдені з використанням одного множення матриці на вектор, на цей раз між вкладеною матрицею й середнім вектором слів у групі. Вектор, в якому $s_{[w]} = \text{sim}(w, w_{1:k})$ обчислюється за:

$$s = E(w_1 + w_2 + \dots + w_k) / k. \quad (2.3)$$

2.2 Архітектура мережі

Архітектура глибокої рекурентної нейронної мережі може бути представлена наступним чином (рис. 2.3)



Рисунок 2.3 - Архітектура глибокої рекурентної нейронної мережі

Представлена архітектура, ~~що~~ складається зі змінних компонентів. Подана система - це рекурентна нейронна мережа (RNN) з одним або декількома згортковими вхідними шарами, за ним слідує кілька повторюваних (односпрямованих або двонаправлених) шарів і один повністю зв'язаний шар перед softmax шаром. Мережа навчається наскрізним способом із використанням функції втрати нейромережевої часової класифікації (НЧК, Connectionist temporal classification), яка дозволяє безпосередньо прогнозувати послідовності символів із вхідного аудіо.

Входи в мережу представляють собою послідовність лог-спектрограм нормалізованих за потужністю звукових аудіокліпів на 20 мс вікна. Виходи - алфавіт кожної мови. На кожному виході тимчасовий крок t . Рекурентна нейронна мережа (РНМ) робить прогноз, $p(l_t/x)$, де l_t це або символ в алфавіті, або порожній символ. Англійською у нас $l_t \in \{a, b, c, \dots, z, \text{пробіл}, \text{апостроф}, \text{пропуск}\}$, де додали символ пробілу для позначення границь слова. Під час виводу моделі, НЧК з'єднуються з

мовою, модель навчається на великому корпусі мовлення. Використовується спеціалізований пошук променя, щоб знайти транскрипцію y , яка максимізує

$$Q(y) = \log(p_{\text{RNN}}(y/x)) + \alpha \log(p_{\text{LM}}(y)) + \beta \text{wc}(y), \quad (2.4)$$

де $\text{wc}(y)$ - це кількість слів (англійською мовою) в транскрипції y .

Вага α контролює відносні вклади мовленнєвої моделі й мережу НЧК.

Вага β заохочує більше слів у транскрипції. Ці параметри налаштовані на тривалий набір розробки.

2.2.1 Пакемна нормалізація для глибоких РНМ

Для ефективного поглинання даних при масштабуванні навчального набору збільшується глибина мереж, шляхом додавання більшої кількості повторюваних шарів. Проте, стає все складніше навчати мережі, що використовують градієнтний спуск у міру збільшення розміру й глибини. При застосуванні глибоких мереж РНМ, на великих наборах даних, варіант використання пакетної нормалізації, істотно покращує остаточну помилку узагальнення на додаток до прискорення навчання. При застосуванні до глибоких мереж РНМ, на великих наборах даних, варіант використання пакетної нормалізації істотно покращує остаточну помилку узагальнення на додаток до прискорення навчання. Рекурентний шар реалізований як:

$$h_t^l = f(W^l h_t^{l-1} + U^l h_{t-1}^l + b), \quad (2.5)$$

де обчислюються активація шару l на часовому кроці t шляхом об'єднання активацій із попереднього шару h_{t-1}^l , в той же час крок t і активація з поточного шару на попередньому кроці часу h_t^l .

Є два способи застосування пакетної нормалізації для періодичної операції. Це вставити перетворення пакетної нормалізації, $B()$, перед кожною нелінійністю наступним чином:

$$h_t^l = f\left(B(W^l h_t^{l-1} + U^l h_{t-1}^l)\right). \quad (2.6)$$

В цьому випадку статистика середнього й дисперсії накопичується за один часовий інтервал міні-пакета. Це не є ефективним. Альтернативою (sequence-wise нормалізація) є нормалізувати пакетно тільки вертикальні зв'язки. Періодичне обчислення задається

$$h_t^l = f(V(W^l h_t^{l-1}) + U^l h_{t-1}^l). \quad (2.7)$$

Для кожної прихованої одиниці обчислюється середнє значення й дисперсія статистики по всім пунктам у міні-пакеті по довжині послідовності.

2.2.2 Визначення архітектури мережі та гіперпараметрів

Мережа складається з 3-х щільних шарів двонапрямлених осередків рекуррентної нейронної мережі (2 шари) і вихідного шару (щільного). Кожен із цих шарів використовує 2048 прихованих юнітів.

Архітектура виглядає наступним чином:

26 MFCC features →→ Dense Layer (2048 units) →→ Dropout 0.05 →→ Dense Layer (2048 units) →→ Dropout 0.05 →→ Dense Layer (2048 units) →→ Dropout 0.05 →→ Bi-Directional LSTM (2048 units per direction) →→ Dense Layer (2048 units) →→ Output.

MFCC - це мел-частотні кепстральні коефіцієнти, які виступають в ролі набору ознак, які б однозначно класифікували звукову хвилю.

Dropout (виняток) - метод регуляризації штучних нейронних мереж, призначений для запобігання перенавчання мережі.

LSTM складається з 3 вентилів: вхідний вентиль, вентиль забування і вихідний вентиль. В теорії можна використовувати індивідуальні функції активації (σ) для кожного вентиля. Проте, сигмоїдні функції активації зазвичай використовуються для всіх трьох воріт.

2.2.3 Частота згорток

Часова згортка зазвичай використовується в розпізнаванні мовлення для ефективного моделювання незмінності перекладу в часі для висловлювань змінної довжини. Згортка в частоті моделює спектральну дисперсію через мінливість. Динаміка лаконічніше, якщо моделювання відбувається в великій повністю підключеній мережі. Згортка в спробах моделювання спектральної

дисперсії, більш стиснута, через мінливість акустичних систем, ніж у великих повністю підключених мережах. Архітектура згортки з майбутнім розміром контексту 2 представлена на рис. 2.4.

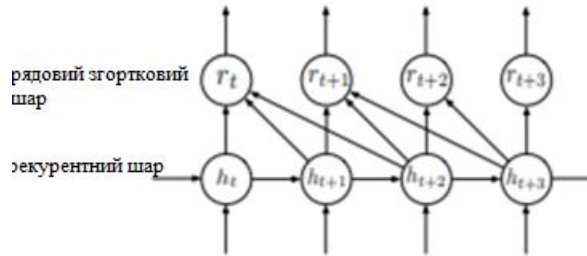


Рисунок 2.4 - Архітектура згортки з майбутнім розміром контексту 2

2.2.4 Згортка Lookahead і односпрямовані моделі

Двонаправлені моделі РНМ, із низькою затримкою, є складними для розгортки в онлайні, тому що вони не можуть передавати процес транскрипції, як висловлювання, яке надходить від користувача. Проте, моделі з тільки прямими рекурентними співвідношеннями зазвичай працюють гірше, ніж аналогічні двонаправлені моделі, маючи на увазі, що певний обсяг майбутнього контексту життєво важливий для гарної продуктивності. Одним із можливих рішень є затримка системи до тих пір, поки в неї не буде більше контексту, але викликати таку поведінку в моделі важко. Щоб побудувати односпрямовану модель, без втрати точності, розроблений спеціальний шар, який викликається в вигляді згорткового перегляду, показаний на рис.2.4. Шар запам'ятовує вагові коефіцієнти для лінійного об'єднання активацій кожного нейрона τ і тимчасові кроки в майбутнє, таким чином, дозволяючи контролювати кількість необхідного майбутнього контексту. Випереджаючий шар визначається матрицею параметрів $W \in R^{(d,\tau)}$, де d відповідає кількості нейронів у попередньому шарі. Активації r_t для нового шару на часовому кроці t :

$$r_{t,i} = \sum_{j=1}^{\tau+1} W_{i,j} h_{t+j-1,i}, \text{ for } 1 \leq i \leq d. \quad (2.8)$$

Прогнозована згортка ставиться вище всіх поточних шарів. Це дозволяє транслювати всі обчислення нижче прогнозованої згортки на більш тонкій деталізації.

2.3 Рекурентні нейронні мережі: послідовність моделювання і стеки

При роботі з мовленнєвими даними робота пов'язана з послідовностями, такими як слова (послідовності букв), речення (послідовності слів) і документи.

Рекурентні нейронні мережі дозволяють перетворювати входи послідовних даних довільного розміру в вектори фіксованого розміру, звертаючи увагу на структуровані властивості входів. РНМ, особливо з закритою архітектурою, такі як LSTM і GRU, дуже потужні в захопленні статистичних закономірностей у послідовних входах. Вони внесли найсильніший внесок глибокого навчання до набору інструментів статистичної обробки природнього мовлення.

2.3.1 Опис РНМ

Використовується $x:j$ щоб визначити послідовність векторів x_1, \dots, x_j . На високому рівні, РНМ є функцією, яка приймає в якості вхідних даних довільну довжину впорядкованої послідовності n d_{in} -мірні вектори $x_{1:n} = x_1, x_2, \dots, x_n$, ($x_i \in R^{d_{in}}$) і повертає в якості вихідного вектора одиночний d_{out} -мірний вектор $y_n \in R^{d_{out}}$:

$$\begin{aligned} y_n &= RNN(x_{1:n}) \\ x_i &\in R^{d_{in}} \quad y_n \in R^{d_{out}} \end{aligned} \quad (2.9)$$

Це неявно визначає вихідний вектор y_i для кожного префікса $x_{1:i}$ послідовності $x_{1:n}$. Позначимо через RNN^* функцію, яка повертає цю послідовність:

$$\begin{aligned} y_{1:n} &= RNN^*(x_{1:n}) \\ y_i &= RNN(x_{1:i}) \\ x_i &\in R^{d_{in}} \quad y_n \in R^{d_{out}} \end{aligned} \quad (2.10)$$

Вихідний вектор y_n використовується для подальшого прогнозування. Наприклад, модель для прогнозування умовної імовірності події e , з урахуванням послідовності $x_{1:n}$, може бути визначеною як $p(e = j | x_{1:n}) = \text{softmax}(RNN(x_{1:n}) * W + b)_{[j]}$, j -й елемент, у вихідному векторі, в результаті softmax операції над лінійним перетворенням кодування РНМ $y_n = RNN(x_{1:n})$. Функція РНМ забезпечує основу для формування всієї історії x_1, \dots, x_i без використання Марковських моделей, які традиційно

використовується для моделювання послідовностей. Дійсно, засновані на РНМ мовленнєві моделі призводять до дуже високих показників, у порівнянні з моделями на основі N-грам.

Якщо трохи докладніше, то РНМ визначається рекурсивно за допомогою функції R , приймаючи в якості вхідних даних стан вектору s_{i-1} і вхідний вектор x_i , повертаючи новий вектор стану s_i . Потім вектор стану s_i відображається у вихідному векторі y_i за допомогою простої детермінованої функції $O(\cdot)$ ². Основою рекурсії є вектор початкового стану s_0 , який також є входом до РНМ. Для стислості часто опускається початковий вектор s_0 або припускається, що це нульовий вектор. При побудові РНМ, так само, як при побудові мережі прямого зв'язку, один повинен показувати розмір входів x_i , а також розміри виходів y_i . Розміри станів s_i є функцією вихідного розміру.

$$\begin{aligned} \text{RNN}^*(x_{1:n}; s_0) &= y_{1:n} \\ y_i &= O(s_i) \\ s_i &= R(s_{i-1}, x_i) \\ x_i \in R^{d_{in}} \quad y_n \in R^{d_{out}}, \quad s_i \in R^{(d_{out})} \end{aligned} \quad (2.11)$$

Функції R і O однакові у всіх позиціях послідовності, але РНМ продовжує відстежувати стан обчислень через вектор стану s_i , який зберігається й передається через виклики R . Графічно РНМ традиційно представлений на рисунку 2.5.

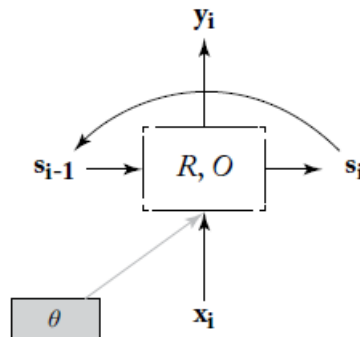


Рисунок 2.5 - Графічне зображення РНМ (рекурсивної)

Це представлення дорівнює рекурсивному визначенню та є правильним для довільно довгих послідовностей. Однак для вхідної послідовності кінцевого розміру можна розгорнути рекурсію, що призведе до структури на рисунку 2.6.

Хоча це не відображається в візуалізації, тут додаються параметри, щоб виділити той факт, що одні й ті ж параметри є загальними для всіх тимчасових кроків. Різні екземпляри R і O приведуть до різних мережових структур, і демонструватимуть різні властивості з точки зору часу їх виконання та здатності до ефективного навчання, з використанням градієнтних методів. Однак вони не всі дотримуються такого абстрактного інтерфейсу.

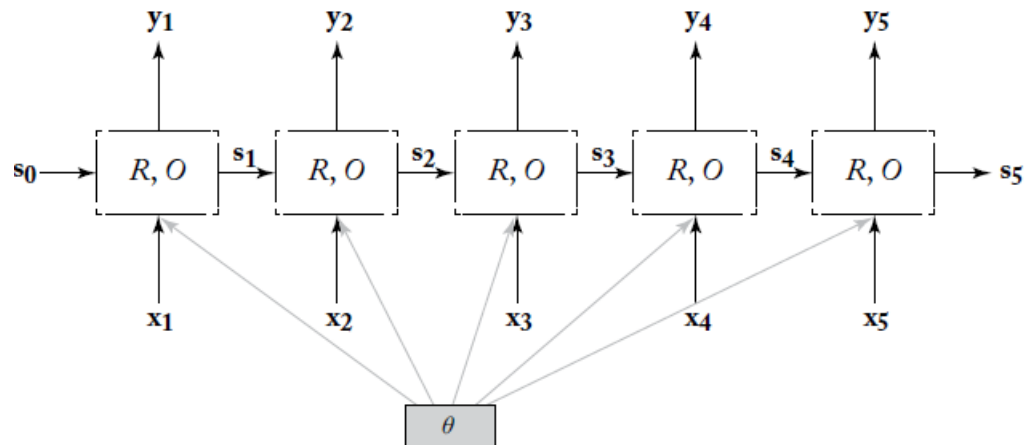


Рисунок 2.6 - Графічне представлення РНМ (у розгорнутому виді)

По-перше, значення s_i (отже, y_i) базується на всьому вході x_1, \dots, x_i . Наприклад, розширюючи рекурсію для $i = 4$, отримуємо:

$$\begin{aligned}
 s_4 &= R(s_3; x_4) \\
 &= R(R(s_2, x_3), x_4) \\
 &= R(R(R(s_1, x_2), x_3), x_4) \\
 &= R(R(R(R(s_0, x_1), x_2), x_3), x_4)
 \end{aligned} \tag{2.12}$$

Таким чином, s_n і y_n можна розглядати як елементи, що кодують всю вхідну послідовність.. Робота мережевого навчання полягає в тому, щоб встановити параметри R і O такими, що стан передає корисну інформацію для завдання, яке потрібно вирішувати.

2.3.2 РНМ навчання

Виходячи з рис. 2.6, розгорнута РНМ - це дуже глибока нейронна мережа (точніше, дуже великий обчислювальний граф із досить складними вузлами), в якому ті ж параметри

розподіляються між багатьма частинами обчислень, а додаткові входи додаються до різних шарів [13]. Для навчання мережі РНМ все, що потрібно, це створити розгорнутий граф обчислень для заданої вхідної послідовності, додати вузол втрат до розгорнутого графа, а потім використовувати зворотний алгоритм (зворотного поширення) для розрахунку градієнтів щодо цієї втрати. Ця процедура в літературі РНМ згадується як зворотне поширення в часі.

Яка мета навчання? Важливо розуміти, що РНМ самостійно багато не працює, але служить компонентом, що навчають у більшій мережі. Остаточне прогнозування й обчислення втрат виконуються цією більш великою мережею, а помилка поширюється назад, через РНМ. Таким чином, РНМ вчиться кодувати властивості вхідних послідовностей, які корисні для подальшого завдання прогнозування. Контрольний сигнал подається не на РНМ, а через велику мережу.

2.4 Загальний принцип використання РНМ

2.4.1 Акцептор

Одним із варіантів є базування контрольного сигналу тільки на кінцевому вихідному векторі, y_n [14]. З цієї точки зору, РНМ навчається в якості *акцептора*. Спостереження йде за остаточним станом, а потім приймається рішення про результат. Наприклад, можливість навчання РНМ читати символи слова по одному, а потім використовувати кінцевий стан, щоб передбачити частини мови цього слова. РНМ, яка читає в реченні і на основі кінцевого стану вирішує, чи буде він передавати позитивний чи негативний настрій або РНМ, яка читає послідовності слів і вирішує, чи є це дійсною іменною фразою. Втрата в таких випадках визначається функцією $y_n = O(s_n)$. Як правило, вихідний вектор РНМ y_n подається в повністю пов'язаний шар або MLP (багатошаровий перцептрон Румельхарта), який виробляє прогноз. Потім градієнти помилок поширюються через іншу частину послідовності (рис. 2.7). Втрата може приймати будь-яку знайому форму: перехресна ентропія, шарнір, маржа і т. д.

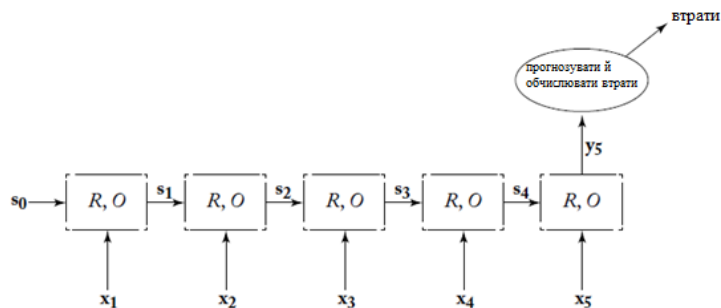


Рисунок 2.7 - Граф навчання акцептора РНМ

2.4.2 Енкодер

Як і у випадку з акцептором, контроль енкодера використовує тільки кінцевий вихідний вектор u_n . Однак, на відміну від акцептора, де передбачення виконується виключно на основі кінцевого вектора, тут кінцевий вектор розглядається як кодування інформації у послідовності і використовується в якості додаткової інформації разом із іншими сигналами. Наприклад, система підсумовування витяжного документа може спочатку виконати документ із РНМ, що призведе до вектору u_n , об'єднуючи весь документ. Потім u_n буде використовуватися разом із іншими функціями для вибору речень, які будуть включені до узагальнення.

2.4.3 Перетворювач

Іншим варіантом обробки РНМ є перетворювач, що виробляє вихід t_i для кожного входу, який він зчитує [15]. Змодельований таким чином, можна обчислити локальний сигнал втрат $L_{local}(t_i, t_i)$ для кожного з виходів t_i на основі істинної позначки t_i . Втрата для розгорнутої послідовності буде тоді: $L(t_{1:n}, t_{1:n}) = \sum_i^n L_{local}(t_i, t_i)$ або використовуючи іншу комбінацію, а не суму, таку як середнє або середньозважене значення (рис. 2.8). Одним із прикладів такого перетворювача є мітка послідовності, в якій використовується $x_{i:n}$, щоб бути характерним представленням для n слів речення, і t_i як вхід для прогнозування призначення тега слова i на основі слів $1:i$. Супер-тегер ССГ базується на архітектурі, що забезпечує дуже високі результати супер-тегування ССГ, хоча в багатьох випадках перетворювач, заснований на двонаправленій РНМ краще підходить для таких завдань тегування. Ще одним варіантом використання установки перетворення є мовне моделювання, в якому послідовність слів $x_{1:l}$ використовується для прогнозування розподілу $(i+1)$ слова [16]. Мовні моделі, засновані на РНМ, забезпечують набагато краще вирішення складностей, ніж традиційні мовні моделі.

Використання РНМ, в якості перетворювачів, дозволяє послабити використання прихованих марковських моделей, які традиційно використовується в мовних моделях, і встановити умови для всієї історії передбачень.

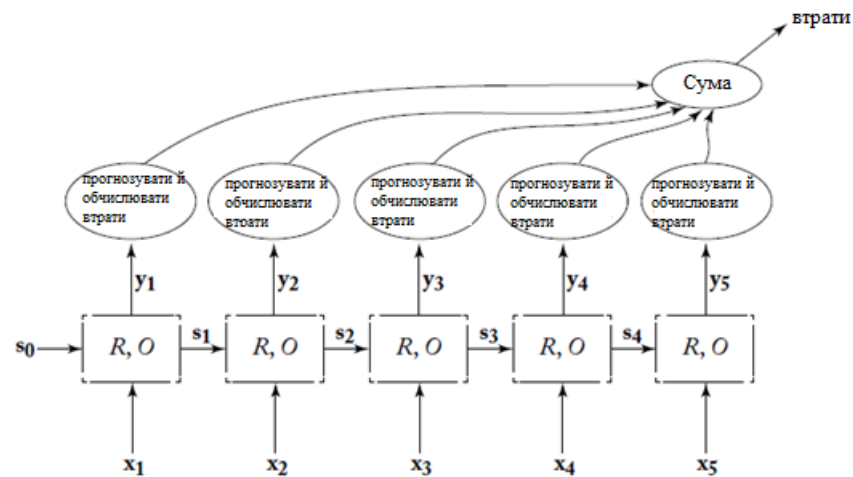


Рисунок 2.8 – Учбовий граф перетворювача РНМ

Окремими випадками перетворювача РНМ є генератор РНМ і відповідна умовна генерація (також звана кодером-декодером) та умовна генерація з архітектурами.

2.5 РНМ як особливості екстракторів

Основний варіант використання РНМ - це гнучкі, екстрактори функцій, яких навчають, що можуть замінити частини більш традиційної функції вилучення конвесрів при роботі з послідовностями. Зокрема, РНМ є добрими замінами для екстракторів на основі вікон.

2.5.1 Частина мови тегування

Корпус: глибока biRNN розмічування частин мови є окремим випадком завдання тегування послідовності, присвоюючи вихідний тег для кожного з n вхідних слів. Це робить biRNN ідеальним кандидатом на основу структури.

Враховуючи речення зі словами $s = w_{1:n}$, слова переводяться до вхідних векторів x , використовуючи особливість функції $x_i = \varphi(s, i)$. Вхідні вектори будуть подаватися до глибокої biRNN, виробляючи вихідні дані векторів $y_{1:n} = \text{biRNN}^*(x_{1:n})$. Кожен із векторів y_i потім буде поданий до багат шарового перцептрона Румельхарта, який буде передбачати один із k можливих вихідних тегів для слова. Кожний вектор y_i фокусується на позиції i в послідовності, але також містить інформацію щодо всієї послідовності, що оточує цю позицію («нескінченне вікно»). Завдяки процедурі навчання, biRNN навчиться фокусуватися на послідовних аспектах, які є інформативними для прогнозування мітки для w_i , і кодувати їх до векторів y_i .

Від слів до входів із символічними рівнями РНМ. Як можна зіставити слово w_i із вектором введення x_i ? Одна можливість полягає в тому, щоб використовувати матрицю введення, яка може бути або ініціалізованою, або попередньо навчена з використанням такої техніки, як W_{ORD2VEC} , з контекстами позиційного вікна. Таке відображення буде виконуватися через матрицю вкладення E , яка буде показувати відображення слова для вбудовування векторів $e_i = E[w_i]$. Хоча відображення працює добре, воно також може страждати від проблем покриття словникових предметів, не помічених під час навчання або без попередньої підготовки.

Слова складаються з символів і певних суфіксів та префіксів, а також інших орфографічних підказок, такі як наявність великої літери, дефісів або цифр, можуть дати чіткі підказки щодо класу неоднозначності слова [17].

Далі визначені функції заміняться на функції-екстрактори з РНМ. Для слова w , яке складається з символів c_1, \dots, c_l , відображається кожен символ у відповідне вбудовування вектора c_i . Потім слово буде закодовано з використанням прямої РНМ і зворотної РНМ над символами. Ці РНМ можуть потім замінити вектор вбудовування слова або, що ще краще, приєднатися до нього:

$$x_i = \varphi(s, i) = [E[w_i]; \text{RNN}^f(c_{1:t}); \text{RNN}^b(c_{1:t})]. \quad (2.13)$$

Важливо, що РНМ, яка працює в прямому напрямку, фокусується на захопленні суфіксів, у зворотному напрямку. РНМ, що фокусується на префіксах, і обидві РНМ можуть бути чутливі до заголовних букв, дефісів і навіть до довжини слів. Остаточною моделлю тегування моделей стає:

$$p(t_i = j | w_1, \dots, w_n) = \text{softmax}(\text{MLP}(\text{biRNN}(x_{1:n}, i)))_{[j]}. \quad (2.14)$$

$$x_i = \varphi(s, i) = [E[w_i]; \text{RNN}^f(c_{1:t}); \text{RNN}^b(c_{1:t})]. \quad (2.15)$$

Модель навчається з використанням крос-ентропійної втрати [18]. Використання випадання слів для вбудовування слів корисно. Ілюстрація архітектури приведена на рис. 2.9

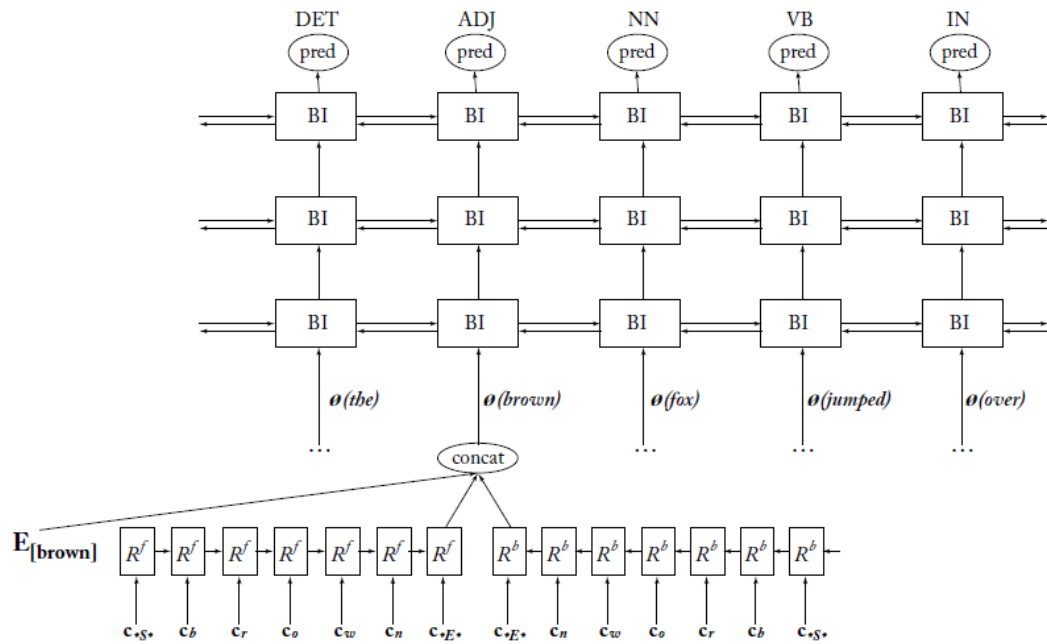


Рисунок 2.9 - Ілюстрація архітектури тегування РНМ

Кожне слово w_i перетворюється у вектор $\varphi(w_i)$, який представляє собою конкатенацію вектора вкладання і кінцевих станів рівня символів рекурентних нейромереж, які прямують вперед і назад. Потім вектори слова потрапляють до глибокої biRNN. Далі вихід кожної із станів biRNN зовнішнього рівня приєднується до мережі прогнозування (багатошаровий перцептрон Румельхарта, за яким слідує softmax), що призводить до прогнозування тега. Важливо, що кожне передбачення тегування може обумовлювати всі вхідні речення

Особливості згорткової нейронної мережі. В наведеній вище архітектурі слова зіставляються з векторами використання рухомих вперед і назад РНМ над символами слова. Альтернатива полягає в тому, щоб представляти слова, використовуючи згортку на рівні символів та об'єднуючи нейронні мережі.

Структуровані моделі. В наведеній вище моделі прогнозування тегів для слова i виконується незалежно від інших тегів. Це може добре працювати, але можна також поставити i -й канал на прогнози попередньої моделі. Умовою можуть бути або попередні k теги (слідуючи припущенням Маркова), і в цьому випадку використовуються вкладення тегів E , що призводить до:

$$p(t_i = j | w_1, \dots, w_n, t_{i-1}, \dots, t_{i-k}) = \text{softmax} (\text{MLP} ([\text{biRNN}(x_{1:n}, i); E[t_{i-1}]; \dots; E[t_{i-k}]]))_j \quad (2.16)$$

або на всій послідовності попередніх прогнозів $t_{1:i-1}$, і в цьому випадку РНМ використовується для кодування послідовності тегів:

$$p(t_i = j | w_1, \dots, w_n, t_{1:i-1}) = \text{softmax}(\text{MLP}([\text{biRNN}(x_{1:n}, i); \text{RNN}^t(t_{1:i-1})]))_{[j]}. \quad (2.17)$$

В обох випадках модель можна запустити в режимі, прогножуючи теги t_i в послідовності або променевий пошук, щоб знайти послідовність тегування з високою оцінкою.

2.6 Висновки до розділу 2

Впровадження методів нейронних мереж було трансформаційним для ОПМ. Це спонукало перейти від лінійних моделей з важкою інженерною характеристикою (і, зокрема, інженерії баккоффа та комбінованих особливостей), до багат шарових перцептронів, які вивчають поєднання функцій; до таких архітектур, як конволюційні нейронні мережі, які можуть ідентифікувати узагальнюючі n -грами: до таких архітектур, як РНМ та двонаправлені РНМ, які можуть ідентифікувати тонкі структури та закономірності в послідовностях довільних довжин. Також створені методи кодування слів як векторів на основі подібності розподілу, які можуть бути ефективними для напіваавтоматичного навчання.

Загалом, галузь швидко розвивається й складно передбачити, що буде далі. Хоча РНМ є набагато досконалішою за попередні методики ОПМ, основна проблема залишається невирішеною: мова є дискретною та неоднозначною, тому без ретельного керівництва людини нейронна мережа малоймовірно навчиться всім тонкощам самостійно. Впровадження нейронних методів ОПМ є важливою дослідницькою проблемою.

РОЗДІЛ 3

АНАЛІЗ ТА ОБРОБКА ТЕКСТУ З ВИКОРИСТАННЯМ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech. Mozilla Deep Speech - це реалізація Mozilla архітектури нейронної мережі Baidu Deep Speech [18]. Запропонований Mozilla набір крім моделі розпізнавання мовлення й прикладів вимови включає готові до застосування модулі для Python і NodeJS, що дозволяють без зайвих ускладнень вбудувати в свої програми функції розпізнавання мовлення. Також поставляється інструментарій для розпізнавання з командного рядка. Програмний рушій розпізнавання мовлення досить швидкий і не вимогливий до ресурсів.

Апаратне забезпечення, що використане для проведення експерименту з практичної реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech, має наступні характеристики:

- 1) Операційна система – 64-розрядна, Ubuntu 18.04.
- 2) Оперативна пам'ять – 24 Гб.
- 3) Процесор – AMD Ryzen 5 1600(3,8 ГГц).
- 4) Відеокарта Nvideo GeForce GTX 1060 3GB

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech проведена за двома сценаріями.

Перший сценарій включає наступні умови та дії:

- використання навченої моделі, що пропонується Mozilla Deep Speech;
- запис аудіо файлів для тестування моделі за участі 15 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

Другий сценарій оснований на створенні власної моделі розпізнавання мовлення з використанням медичних мовленнєвих елементів та включає наступні умови та дії:

- підготовка даних для створення моделі, що включає створення бібліотеки аудіо файлів, які містять мовленнєві елементи медичного напрямку за участі 45 добровольців;

- створення моделі, що включає визначення архітектури та конфігурації мережі для навчання моделі та безпосередньо навчання моделі з використанням підготовлених даних, що складаються з записаних аудіо файлів, які містять мовленнєві елементи медичного напрямку;

- запис аудіо файлів для тестування моделі за участі 5 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech включає наступні етапи.

1. Інсталяція Mozilla Deep Speech.
2. Проведення експерименту за першим сценарієм.
3. Проведення експерименту за другим сценарієм.
4. Оцінка та порівняння результатів експериментів, проведених за першим та другим сценаріями.
5. Узагальнення знань, отриманих на підставі оцінки та порівняння результатів експериментів, та винесення висновків по проведеній реалізації методів і алгоритмів розпізнавання природнього мовлення.

Оцінка результатів експериментів з розпізнавання природнього мовлення, проведених за першим та другим сценаріями проводилась з використанням WER та зафіксованої швидкості розпізнавання в секундах.

Розпізнавання мовлення оцінюється за двома критеріями: коефіцієнт помилок в словах (англ. word error rate -WER) та затримкою (англ. Latency).

1) Коефіцієнт помилок в словах (КПС) вимірює невідповідність рівня слів: він порівнює слова, видані розпізнавачем, з тими, які користувач насправді вимовив. Кожна помилка (заміна, вставка або видалення) враховується в розпізнавальнику. КПС можна обчислити так:

$$\text{КПС} = \frac{\text{кількість замін} + \text{кількість вставок} + \text{кількість видалень}}{\text{загальна кількість слів}} * 100\% \quad (3.1)$$

2) Затримка

Затримка визначається як загальний час (у секундах), необхідний для розпізнавання мовлення. Точніше, затримка є часом з моменту закінчення запису до моменту появи розпізнавання на екрані.

3.1 Mozilla Deep Speech

DeepSpeech значно простіше традиційних систем і при цьому забезпечує більш високу якість розпізнавання при наявності стороннього шуму. В розробці не використовуються традиційні акустичні моделі й концепція фонем, замість них застосовується добре оптимізована система машинного навчання на основі нейронної мережі, що дозволяє обійтися без розробки окремих компонентів для моделювання різних відхилень, таких як шум, відлуння й особливості мови. Особливістю DeepSpeech є те, що для отримання якісного розпізнавання дана архітектура вимагає великого обсягу різнорідних даних для здійснення навчання, надиктованих у реальних умовах різними голосами й при наявності природних шумів. DeepSpeech розроблений із різних причин:

- відкрита сучасна альтернатива пропрієтарним рішенням, таким як Amazon Alexa, Google Assistant або Baidu;
- автономна підтримка;
- приватність;
- низька затримка;
- працює у «вимкнених» налаштуваннях, наприклад, промислові системи управління, які відключені з поважних причин або мобільні пристрої із низьким сигналом мережі;
- масштабованість.

Інсталяція DeepSpeech проводиться наступним чином. Для установки й використання DeepSpeech за допомогою CPU потрібно зробити наступне:

1) створити і активувати віртуальне оточення (virtualenv):

```
virtualenv -p python3 $HOME/tmp/deepspeech-venv/  
source $HOME/tmp/deepspeech-venv/bin/activate
```

2) встановити DeepSpeech

```
pip3 install deepspeech
```

3) завантажити попередньо навчену модель і розпакувати

```
curl -LO
```

```
https://github.com/mozilla/DeepSpeech/releases/download/v0.5.1/deepspeech-0.5.1-models.tar.gz
```

```
tar xvf deepspeech-0.5.1-models.tar.gz
```

4) використати

```
deepspeech --model шлях/до/моделі --alphabet шлях/до/алфавіта --lm
шлях/до/файлу/мовленнєвої/моделі --trie шлях/до/дерева--audio шлях/до/аудіо/файлу
```

Наприклад

```
deepspeech --model deepspeech-0.5.1-models/output_graph.pbmm --alphabet deepspeech-
0.5.1-models/alphabet.txt --lm deepspeech-0.5.1-models/lm.binary --trie deepspeech-
0.5.1-models/trie --audio audio/2830-3980-0043.wav
```

Для того, щоб використовувати DeepSpeech за допомогою GPU необхідно:

1) встановити пакет TensorFlow

```
pip install tensorflow-gpu
```

2) додати репозиторії пакетів

```
wget
```

```
https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86_64/cu
da-repo-ubuntu1804_10.0.130-1_amd64.deb
```

```
sudo dpkg -i cuda-repo-ubuntu1804_10.0.130-1_amd64.deb
```

```
sudo apt-key adv --fetch-keys
```

```
https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86_64/7f
a2af80.pub
```

```
sudo apt-get update
```

```
wget http://developer.download.nvidia.com/compute/machine-
```

```
learning/repos/ubuntu1804/x86_64/nvidia-machine-learning-repo-
```

```
ubuntu1804_1.0.0-1_amd64.deb
```

```
sudo apt install ./nvidia-machine-learning-repo-ubuntu1804_1.0.0-
```

```
1_amd64.deb
```

```
sudo apt-get update
```

3) встановити драйвер Nvidia

```
sudo apt-get install --no-install-recommends nvidia-driver-418
```

4) перезавантажити ПК, перевірити що графічний процесор видимий командою - nvidia-smi

5) встановити бібліотеки розробки та середовища виконання:

```
sudo apt-get install --no-install-recommends \
  cuda-10-0 \
  libcudnn7=7.6.2.24-1+cuda10.0 \
  libcudnn7-dev=7.6.2.24-1+cuda10.0
```

6) встановити TensorRT. Потрібно, щоб libcudnn7 був встановленим

```
sudo apt-get install -y --no-install-recommends libnvinfer5=5.1.5-1+cuda10.0 \
  libnvinfer-dev=5.1.5-1+cuda10.0
```

7) створити і активувати віртуальне оточення (virtualenv):

```
virtualenv -p python3 $HOME/tmp/deepspeech-venv-gpu/
source $HOME/tmp/deepspeech-venv-gpu/bin/activate
```

8) встановити DeepSpeech з підтримкою GPU:

```
pip3 install deepspeech-gpu
```

9) завантажити попередньо навчену модель і розпакувати

```
curl -LO
https://github.com/mozilla/DeepSpeech/releases/download/v0.5.1/deepspeech
0.5.1-models.tar.gz
tar xvf deepspeech-0.5.1-models.tar.gz
```

10) використати

```
deepspeech-model шлях/до/моделі-alphabet шлях/до/алфавіта -lm
шлях/до/файлу/мовленої/моделі-trie шлях/до/дерева-audio шлях/до/аудіо/файлу
```

Наприклад:

```
deepspeech --model models/output_graph.pbmm -alphabet
models/alphabet.txt --lm models/lm.binary --trie models/trie -
audio audio/test.wav
```

Папка моделі містить 7 файлів:

```

├── alphabet.txt
├── lm.binary
├── output_graph.pb
├── output_graph.pbmm
├── output_graph.rounded.pb
├── output_graph.rounded.pbmm
└── trie

```

- alphabet.txt не вимагає пояснень. Містить список усіх символів англійського алфавіту.
- output_graph.pb - модель у вигляді замороженого тензорного графа.
- rounded - моделі, які використовують округлені ваги для збільшення обчислень.
- pbmm - це графіки з відображенням у пам'яті, щоб зробити їх більш ефективними.
- lm.binary - це двійкова модель мови.
- trie - це мовленнєва модель у вигляді дерева.

3.2 Проведення експерименту з розпізнавання природнього мовлення за першим сценарієм

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech за першим сценарієм включає наступні умови та дії:

- використання навченої моделі, що пропонується Mozilla Deep Speech;
- запис аудіо файлів для тестування моделі за участі 15 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

3.2.1 Завантаження навченої моделі

Запис аудіо файлів для навчання моделі за участі 39577 добровольців. З них 67 % осіб чоловічої статі та 33 % осіб жіночої статі.

Дані, які використовувались для навчання стандартної моделі:

сумарний об'єм аудіофайлів становить - 30 ГБ;

загальна сума годин – 1087;

акценти: 23% - США, 9% - Великобританія, 4% - Індія та Південна Азія (Індія, Пакистан, Шрі-Ланка), 3% - Австрія, 3% - Канада, 1% - Шотландія, 1% - Південна Африка (Південна Африка, Зімбабве, Намібія), 1% - Нова Зеландія.

3.2.2 Запис та підготовка аудіо файлів для тестування моделі

Запис аудіо файлів для тестування моделі за участі 15 добровольців. З них 5 осіб жіночої статі та 9 осіб чоловічої статі.

Дані, які використовувались для навчання стандартної моделі:

сумарний об'єм аудіофайлів становить – 2,07 МБ;

загальна сума годин – 93 с;

Підготовка записаних аудіо файлів включала конвертування у формат wav, змінення частоти дискретизації до 16 кГц та перетворення аудіодоріжки у моно-формат.

3.2.3 Тестування моделі

У таблиці 3.1 наведені результати тестування стандартної моделі. Її можна встановити відразу після інсталяції Deep Speech.

Таблиця 3.1 – результати розпізнавання попередньо навченої моделі

№	Текст	Вік (роки)	Стать	Акцент	Текст на виході	Коефіцієнт помилок слова (%)
1	down below in the darkness were hundreds of people sleeping in peace	20	чоловіча	США	then below and the darkness were hundreds of people sleeping in peace	17
2	this was the strangest of all things that ever came to earth from outer space	30	чоловіча	Великобританія	this was the strangest of all things that ever came to earth from out a space	13
3	follow the instructions here	20	чоловіча	Шотландія	full of these structions here	100
4	he didn't need to seek out the old woman for this	20	жіноча	США	you didn't need to seek out the old woman for this	9
5	don't point the gun at me	20	жіноча	Індія	don't bind the gun at me	0
6	i told you to have the ice box fixed	50	жіноча	Австралія	i don't do it to him the eye both ex	78

Продовження табл. 3.1

№	Текст	Вік (роки)	Стать	Акцент	Текст на виході	Коефіцієнт помилок слова (%)
7	but finally the merchant appeared and asked the boy to shear four sheep	30	чоловіча	Індія	but finally the master appeared and as the boy was shared for ship	46
8	most meteorites are more or less rounded	20	інша	Великобританія	not me i am more less rounded	71
9	from here on you will be alone the alchemist said	30	чоловіча	Канада	from heron you will be alone the alkanet said	30
10	my tax dollars pay for those public school proms	40	чоловіча	Нова Зеландія	my text dollars pay for those public school prone	22
11	an army is coming the boy said	20	жіноча	Філіппіни	an army calling the voice that	57
12	did she discuss it with you	30	жіноча	Ірландія	did she discuss it with you	0
13	i'm crazy about adrian	30	чоловіча	Гонконг	i am crazy about angering	25
14	he decided to spend the night there	підліток	чоловіча	Африка	he decided to spend the night there	0
15	he's in the insurance business	30	чоловіча	Південна Атлантика	he's in the insurance business	0
						29

Загальний коефіцієнт помилок слова при використанні стандартної моделі становить 29 %.

Найбільший коефіцієнт помилок слова, записаних із шотландським акцентом, на такий показник могла вплинути погана вимова людини.

Найменший коефіцієнт помилок слова, записаний у людей із індійським, ірландським та південно-атлантичним акцентом.

3.3 Проведення експерименту з розпізнавання природнього мовлення за другим сценарієм

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech за другим сценарієм має свої особливості. Другий сценарій оснований на створенні власної моделі розпізнавання мовлення з використанням медичної термінології та включає наступні умови та дії:

- підготовка даних для створення моделі;
- визначення архітектури та конфігурації мережі для навчання моделі та, безпосередньо, навчання моделі;
- запис аудіо файлів для навчання моделі за участі 45 добровольців;
- запис аудіо файлів для тестування моделі за участі 5 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

3.3.1 Створення бібліотеки аудіо файлів

Бібліотеку аудіо файлів створено за участі 5 добровольців.

Використане апаратне забезпечення: комп'ютерна гарнітура SteelSeries Arctis 3.

3.3.1.1 Запис аудіо файлів

Проведено запис аудіо файлів, що містять фрази з мовленнєвими елементами медичного напрямку, що промовлялись добровольцями. Перелік фраз для запису аудіо файлів представлений на рис. 3.1.

Assistant left close
 Assistant left open
 Assistant left grasp needle
 Assistant left grasp prostate
 Assistant left stay here
 Assistant right cut here
 Assistant left grasp here
 Assistant left move out
 Assistant right move out
 Assistant left move down bladder
 Assistant left move out prostate
 Assistant left move up here
 Assistant left traction down here
 Assistant left traction up here
 Assistant left traction up seminal vesicle
 Assistant right move down bladder
 Assistant right move up here
 Assistant right traction down here
 Assistant right traction up here
 Assistant left move toward Manipulator left
 Assistant left move toward Manipulator right
 Assistant right move toward Manipulator left
 Assistant right move toward Manipulator right
 Assistant left move anterior above bladder

Рисунок 3.1 – Перелік використаних фраз для створення експериментальної моделі

Кожен доброволець записував по 5 разів кожен фразу.

3.3.1.2 Підготовка аудіо файлів для навчання DeepSpeech

Перш за все необхідно привести файли до виду, який буде зрозумілий DeepSpeech. Для цього файли повинні бути в форматі .wav з частотою дискретизації 16кГц і в моноканалі (рис. 3.2).

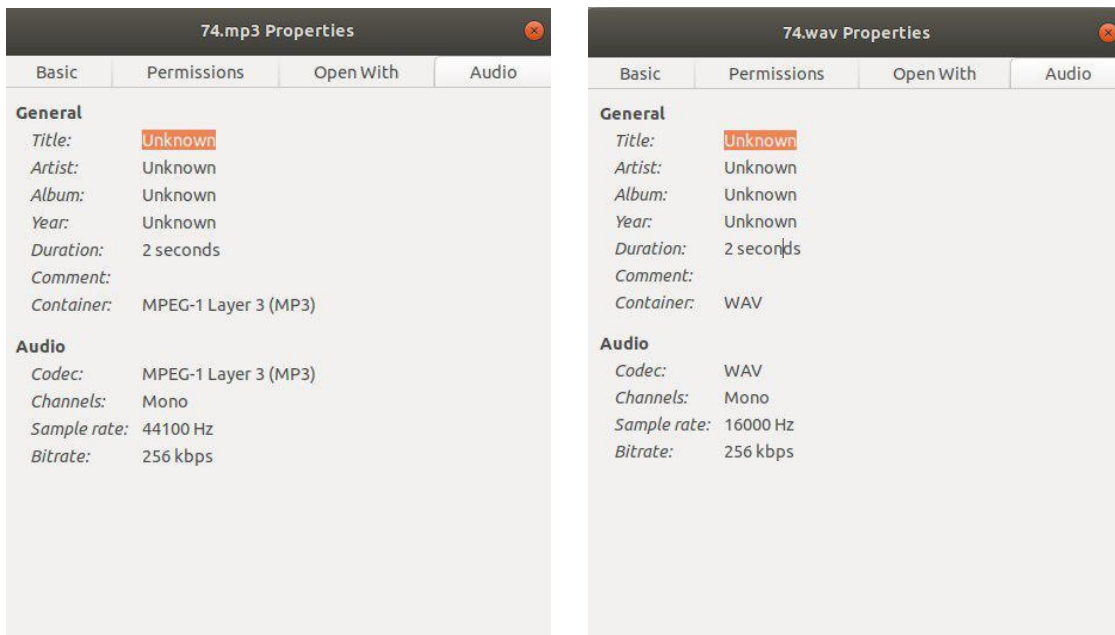


Рисунок 3.2 – Процес конвертування аудіофайлів

Для того щоб DeepSpeech зміг навчатися, необхідно створити 3 csv-файла dev.csv, test.csv і train.csv із наступною структурою:

```
wav_filename,wav_filesize,transcript
```

де

wav_filename – шлях до аудіофайлу,

wav_filesize – розмір файлу в байтах,

transcript – вміст файлу.

Фрагмент csv-файлу, підготовленого для навчання моделі, представлений на рис. 3.3.

	A	B	C
1	wav_filename	wav_filesize	transcript
2	1.wav	211342	down below in the darkness were hundreds of people sleeping in peace
3	2.wav	224554	this was the strangest of all things that ever came to earth from outer space
4	3.wav	112442	follow the instructions here
5	4.wav	189365	he didn't need to seek out the old woman for this
6	5.wav	99874	don't point the gun at me
7	6.wav	114734	i told you to have the ice box fixed
8	7.wav	200347	but finally the merchant appeared and asked the boy to shear four sheep
9	8.wav	167598	most meteorites are more or less rounded
10	9.wav	172409	from here on you will be alone the alchemist said
11	10.wav	171298	my tax dollars pay for those public school proms
12	11.wav	100320	an army is coming the boy said
13	12.wav	95456	did she discuss it with you
14	13.wav	93800	i'm crazy about adrian
15	14.wav	110755	he decided to spend the night there
16	15.wav	107823	he's in the insurance business

Рисунок 3.3 – Фрагмент csv-файлу для навчання моделі

3.3.2 Навчання моделі

Для навчання власної моделі з використання записаної бібліотеки аудіофайлів необхідно виконати наступне:

1) встановити Git Large File Storage вручну або через менеджер пакетів, якщо він доступний у вашій системі;

2) потім клонувати репозиторій DeepSpeech:

```
git clone https://github.com/mozilla/DeepSpeech
```

3) створити віртуальне середовище:

```
virtualenv -p python3 $HOME/tmp/deepspeech-train-venv/
```

При створенні віртуального середовища з'явиться каталог, що містить двійковий файл python3 і все необхідне для запуску deepspeech. Можна використовувати будь-який каталог.

4) активуйте віртуальне середовище:

```
source $HOME/tmp/deepspeech-train-venv/bin/activate
```

Кожен раз, коли потрібно працювати з DeepSpeech, потрібно активувати цю віртуальне середовище.

5) встановити залежності Python:

```
cd DeepSpeech
```

```
pip3 install -r requirements.txt
```

Також знадобиться пакет `ds_ctcdecoder` Python, який необхідний для декодування вихідних даних акустичної моделі DeepSpeech у текст.

```
pip3 install $(python3 util/taskcluster.py --decoder)
```

б) Тепер можна приступати до навчання

```
python3 DeepSpeech.py --n_hidden 2048 --checkpoint_dir
path/to/checkpoint/folder --epochs 3 --train_files my-train.csv --dev_files my-
dev.csv --test_files my_dev.csv --learning_rate 0.0001 --export_dir my_models
--checkpoint_dir – шлях до папки де будуть зберігатися контрольні точки навчання,
--epochs – кількість епох навчання,
--train_files – шлях до csv-файлу з аудіофайлами для навчання моделі,
--dev_files - шлях до csv-файлу з перевірочним набором аудіофайлів,
--test_files – шлях до csv-файлу з файлами для тестування моделі,
--export_dir – шлях до папки, де буде зберігатися модель навчання.
```

3.3.3 Тестування моделі

Процес тестування власної моделі включав запис аудіо файлів для тестування моделі за участі 5 добровольців, підготовку та обробку записаних для тестування аудіо файлів, тестування моделі з використанням підготовлених аудіо файлів.

За результатами тестування власно створеної моделі отримані наступні результати WER для розпізнавання кожної тестової фрази, представлені в табл. 3.2.

Таблиця 3.2 - Результати тестування створеної моделі

№	Фраза	Розпізнана фраза	трозп, с	КПС, %	№	Фраза	Розпізнана фраза	трозп, с	КПС, %
1	Left traction up here	Left catching up here	0.675	25	11	Left move anterior above bladder	left must endure above blade	0.765	80
		left direction up he	0.687	50			let me and darrel about la	0.713	100
		a less traced oppose	0.695	100			oh lattimore and dare a ball blore	1.011	100
		left decomposure	0.603	100			left a very orebo later	0.862	80
		miss curran up her	0.730	75			love to it and far about lat	1.093	100
2	Right move toward Manipulator left	reduction down here	0.815	100	12	Right move up	right mowat	0.492	66
		rise rection dan here	0.726	83			right milt	0.614	66
		right dragon down to her	0.819	50			written out	0.585	100
		etched absinthe	0.675	100			right mobile	0.582	66
		right action down here	0.814	83			ride a	0.662	100
3	Right traction down here	reduction down here	0.815	50	13	Right move down the bladder	right don't said winter	0.778	80
		rise rection dan here	0.638	75			right moved down a blade	0.714	60
		right dragon down to her	0.759	75			writing down the bl	0.628	80
		etched absinthe	0.610	100			all right move down to blot her	0.786	60
		right action down here	0.755	50			at docible	0.594	100
4	Left close	let lose	0.552	100	14	Left move toward Manipulator left	left more to god do inchie	1.146	83
		left close	0.618	0			her eyes moved over in chief	0.815	100
		the left close	0.591	33			left more tired the reef	0.718	83
		left	0.499	50			left move toward the we clashed	0.807	50
		left close	0.561	0			ah to attain sure	0.657	100
5	Left grasp here	left grasped her	0.611	66	15	Assistance Left traction up seminal vesicle	left action up seminaverit	0.815	60
		left dress here	0.616	33			s left direction up simon with thick mist	0.851	87.5
		the lad grasped her	0.636	100			less direction up seven over ice	0.808	83
		let earth pear	0.643	100			let the extra up seven recite	0.952	100
		last trap here	0.727	66			this traction up singing out there beside	1.190	71
6	Left traction down here	left direction down hale	0.699	50	16	Right move down the bladder	right moved down a blade	0.699	60
		left direction down here	0.754	25			writing down the bl	0.603	80
		the last train down here	0.704	50			all right move down to blot her	0.802	57
		laughed to condone her	0.737	100			at docible	0.609	100
		lived traction down here	0.767	25			right on to the platter	0.850	80
7	Right move up here	her eyes moved up here	0.670	50	17	Left grasp prostate	lived grass rose	1.033	100
		writing of her	0.597	100			left gasped prostate	0.691	33
		or i too up her	0.694	100			alas to grasp prostate	0.735	50
		at both apperture	0.625	100			left it for state	0.873	75
		right more up here	0.670	25			left great thrust at	0.740	75
8	Left open	left open	0.606	0	18	Left move out the prostate	left me all the pro	0.806	60
		left open	0.603	0			left mobile sir state	0.713	80
		let open	0.540	50			last move out the prostate	0.782	20
		left open	0.585	0			left more at a petate	0.770	80
		left open	0.584	0			he moved more houseboat	0.814	100
9	Right move toward da Vinci right	right moved toward the inter it	0.785	33	19	Left grasp the needle	lived up in de	0.711	100
		right moved toward the insure it	0.804	33			let dropped the need	0.715	100
		bride i dread the winter it	0.723	100			the left grasps any do	0.627	80
		or i more dread the winter i	0.707	100			left gasps and he do	0.700	80
		or i to move to the red right	0.800	100			we grasped in no	0.687	100
10	Left stay here	he left that here	0.626	33	20	Right cut here	her right got here	0.647	50
		left they hear	0.642	33			right god he	0.674	33
		the left they here	0.603	33			it caught the her	0.558	100
		let they he	0.557	100			at that her	0.736	100
		well i say	0.600	100			it got her	0.696	100

Продовження табл. 3.2

№	Фраза	Розпізнана фраза	t _{розп} , с	КПС,%
21	Left move out	left wilt	0.591s	66
		let me	0.592s	100
		the last move out	0.623s	50
		left me out	0.647s	33
		tis more out	0.615s	66
				69.2

Загальний коефіцієнт помилок при використанні власно навченої моделі становить 69.2 %.

Найбільший коефіцієнт помилок на фразах: Right move up here, Right move toward da Vinci right, Left move anterior above bladder, Right move up, Left grasp the needle, Right cut here.

Найменший коефіцієнт помилок на фразі Left open.

Затримка є найбільшою на фразі Left move anterior above bladder

Затримка є найменшою на фразі Left close.

3.4 Порівняння моделей розпізнавання

Проведений порівняльний аналіз коефіцієнту помилок тестування стандартної моделі та моделі, власно навченої з використанням створеної бібліотеки аудіо файлів із фразами медичної галузі. Результати представлені в табл. 3.3

Таблиця 3.3 – Порівняльний аналіз узагальненого коефіцієнту помилок в словах за результатами тестування моделей

Модель	Узагальнений коефіцієнт помилок в словах (WER)
Стандартна модель	29%
Створена модель	69.2%

Як видно з результатів, коефіцієнт помилок слова створеної моделі вищий за коефіцієнт помилок слова стандартної моделі на 40.2 %. Беручи до уваги вихідні дані моделей: кількість файлів для навчання, якість запису аудіо файлів, акцент можна зробити висновок про вплив цих факторів на якість моделі та в майбутніх дослідженнях прийняти заходи по їх усуненню.

3.5 Висновки до розділу 3

В третьому розділі представлено практична реалізація та проведення експерименту з розпізнавання природнього мовлення. Практична реалізація методів і алгоритмів розпізнавання

природнього мовлення реалізована з використанням нейронних мереж за допомогою Mozilla Deep Speech проведена за двома сценаріями.

Перший сценарій включав використання навченої моделі, що пропонується Mozilla Deep Speech, запис аудіо файлів для тестування моделі за участі, підготовку та обробку записаних аудіо файлів, тестування моделі з використанням підготовлених аудіо файлів.

Другий сценарій оснований на створенні власної моделі розпізнавання мовлення з використанням медичних мовленнєвих елементів та включав підготовку даних для створення моделі, що включала створення бібліотеки аудіо файлів, які містять мовленнєві елементи медичного напрямку, створення моделі, що включала визначення архітектури та конфігурації мережі для навчання моделі та безпосередньо навчання моделі з використанням підготовлених даних, що складались з записаних аудіо файлів, що містять мовленнєві елементи медичного напрямку, запис аудіо файлів для тестування моделі, підготовку та обробку записаних аудіо файлів, тестування моделі з використанням підготовлених аудіо файлів.

Практична реалізація методів і алгоритмів розпізнавання природнього мовлення за двома сценаріями включала наступні етапи: проведення експерименту за першим сценарієм, проведення експерименту за другим сценарієм, оцінку та порівняння результатів експериментів, проведених за першим та другим сценаріями, порівняння результатів експериментів, та винесення висновків по проведеній реалізації методів і алгоритмів розпізнавання природнього мовлення.

Оцінка результатів експериментів з розпізнавання природнього мовлення, проведених за першим та другим сценаріями проводилась з використанням узагальненого коефіцієнту помилок в словах. За результатами тестування коефіцієнт помилок в словах для стандартної моделі становив 29%, для створеної моделі для розпізнавання природнього мовлення специфічного медичного напрямлення коефіцієнт помилок в словах становив 69.2%. Таким чином коефіцієнт помилок слова створеної моделі вищий за коефіцієнт помилок слова стандартної моделі на 40.2 %.

Зроблено висновок про залежність якості розпізнавання природнього мовлення від вихідних даних моделей: кількості файлів для навчання, якості запису аудіо файлів, акценту. Майбутні дослідження будуть направлені на заходи по поліпшенню цих факторів.

Таким чином, вирішено поставлене завдання - підвищення якості розпізнавання природнього мовлення за рахунок визначення вагомих факторів, що впливають на якість розпізнавання природнього мовлення з вживанням вузькоспеціалізованої специфічної термінології.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

В даному розділі проаналізовані потенційно небезпечні й шкідливі виробничі фактори. На підставі цього аналізу розроблені технічні, санітарно-гігієнічні і організаційні заходи, спрямовані на усунення причин виробничого травматизму, професійної захворюваності, підвищення продуктивності праці, на зниження дії на довкілля. Оскільки виконання завдання дипломного проекту передбачено на ПК, то аналіз потенційно небезпечних і шкідливих виробничих чинників виконується для обчислювальної техніки. При роботі з обчислювальною технікою змінюються фізичні і хімічні фактори навколишнього середовища: виникає статична електрика, електромагнітне випромінювання, змінюється температура і вологість, рівень вмісту кисню й озону в повітрі. Також недотримання вимог безпеки призводить до того, що при роботі за комп'ютером працівник може відчувати дискомфорт: виникають головні болі й різь в очах, з'являються втома й дратівливість. У деяких людей порушується сон, апетит, погіршується зір, починають хворіти руки, шия, поперек тощо. При ненормованій роботі можливе нервово виснаження. Забезпечення цих умов покладається на власника або уповноважений ним орган (далі роботодавець). Умови праці на робочому місці, безпека технологічних процесів, машин, механізмів, устаткування та інших засобів виробництва, стан засобів колективного та індивідуального захисту, що використовуються працівником, а також санітарно-побутові умови повинні відповідати вимогам нормативних актів про охорону праці, що чітко врегульовані законодавством України.

4.1 Аналіз стану умов праці

Обчислювальна техніка при функціонуванні має наступні експлуатаційні характеристики:

- робоче живлення 220 В;
- частота живильної мережі 50 Гц;
- споживана потужність в межах 300 Вт.

Робочі місця, обладнані ЕОМ з ВДТ і ПП, виконують із забезпеченням виконання НПАОП 0.00-7.15-18 [19] «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями», які встановлюють вимоги безпеки до обладнання робочих місць, до роботи із застосуванням ЕОМ з ВДТ і ПП. При роботі на персональних ПЕОМ користувач наражається на небезпеку ураження електричним струмом. Приміщення для обчислювальної техніки за ступенем

небезпеки ураження людини електричним струмом відноситься до приміщень без підвищеної небезпеки [20]. Тяжкість роботи персоналу, що обслуговує і працює на ПЕОМ, відноситься до категорії 1а - легкі фізичні навантаження. При обслуговуванні обчислювальної техніки мають місце фізичні і психофізіологічні небезпечні та шкідливі виробничі фактори:

- підвищене значення напруги в електричному ланцюзі, замикання якого може відбутися через тіло людини;
- підвищена або знижена температура повітря робочої зони;
- підвищена або знижена рухливість повітря;
- підвищена або знижена вологість;
- підвищений рівень електромагнітних полів у робочій зоні;
- відсутність або нестача природного світла;
- підвищена пульсація світлового потоку;
- розумове перенапруження;
- монотонність праці;
- емоційні навантаження;
- підвищений рівень шуму;
- недостатнє освітлення робочого місця;
- підвищена статична електроенергія.

4.2 Вимоги до приміщення

Для захисту людей від ураження електричним струмом при дотику до металевих неструмоведучих частин, які можуть опинитися під напругою в результаті пошкодження ізоляції, передбачаються наступні заходи:

- захисне заземлення або занулення металевих частин електроустановок, які доступні для дотику людини й не мають інших видів захисту, що забезпечують електробезпеку:
- захисне відключення;
- електричний поділ мереж;
- використання малої напруги;
- ізоляція струмоведучих частин;
- огорожу електроустановок;
- шина заземлення виконується провідником з опором не більше 4-х Ом.

Завдання захисного заземлення - усунення небезпеки ураження струмом у випадку дотику до корпусу та інших струмоведучих металевих частин електроустановок, які опинилися під напругою.

Розрахунок заземлюючого контуру виконується виходячи з умови:

$$R_3 = \frac{R_3 * R_3}{R_{\Pi} * n * \eta_3 + R_3 * \eta_{\Pi}} \leq 4 \text{ Ом}, \quad (4.1)$$

де R_3 - опір заземлювача (стержня, труби, куточка і т.д.), Ом;

R_{Π} - опір лінії, що з'єднує заземлювачі, Ом;

n - кількість заземлювачів;

η_3 і η_{Π} - коефіцієнти екранування відповідно заземлювача і з'єднує смуги ($\eta_3 = 0,2 \div 0,9$; $\eta_{\Pi} = 0,1 \div 0,7$).

Опір заземлювача розраховується за формулою 1.2

$$R_3 = \frac{\rho}{2 * \pi * l} \left(\ln \frac{2 * l}{d} + \frac{1}{2} * \ln \frac{4 * t + 1}{4 * t - 1} \right), \quad (4.2)$$

де ρ - питомий опір ґрунту (взяти з довідкової літератури);

l - довжина заземлювача (для труб 2-3 м, для стрижнів до 10 м), м;

d - діаметр заземлювача (для стрижнів 0,01 - 0,03 м, для труб 0,03 - 0,05 М);

t - відстань від середини забитого в ґрунт заземлювача до рівня землі (необхідно враховувати, що відстань від верхнього кінця заземлювача до поверхні землі має бути не менше 0,5), м.

Розрахуємо опір заземлювача:

$$R_3 = \frac{60}{2 * \pi * 3} \left(\ln \frac{2 * 3}{0.03} + \frac{1}{2} * \ln \frac{4 * 1 + 3}{4 * 1 - 3} \right) = 19.96, \quad (4.3)$$

Опір лінії, що з'єднує заземлювачі розраховується за формулою 4.4

$$R_{\Pi} = \frac{\rho}{2 * \pi * l} * \ln \frac{2 * L^2}{b * t}, \quad (4.4)$$

де L - довжина лінії, що з'єднує заземлювачі (при контурному заземленні вона приблизно дорівнює периметру виробничої будівлі), м;

b - ширина смуги (0,03 - при прокладанні всередині будівлі і 0,05 - при прокладанні поза будівлею), м;

t - глибина заземлення від рівня землі (0,5 м.).

Розрахуємо опір лінії, що з'єднує заземлювачі

$$R_{\Pi} = \frac{60}{2 \cdot \pi \cdot 3} * \ln \frac{2 \cdot 50^2}{0.03 \cdot 5} = 14.37, \quad (4.5)$$

Необхідна кількість заземлювачів, розраховується за формулою 1.6

$$n = \frac{2 \cdot R_3}{4 \cdot \eta_3}, \quad (4.6)$$

де 4 - допустимий загальний опір;

2 - коефіцієнт сезонності.

Розрахуємо необхідну кількість заземлювачів:

$$n = \frac{2 \cdot 19.9}{4 \cdot 0.5} = 19.9 \approx 20, \quad (4.7)$$

Округлимо результат в більшу сторону і отримуємо необхідну кількість заземлювачів - 20.
Маючи всі необхідні дані розрахуємо опір заземлюючого контуру.

$$R_3 = \frac{19.96 \cdot 14.37}{14.37 \cdot 20 \cdot 0.5 + 19.96 \cdot 0.4} = 1.89 \leq 4 \text{ Ом}, \quad (4.8)$$

Опір заземлюючого контуру 1,89 Ом, що відповідає умові $R_3 < 4 \text{ Ом}$.

4.3 Вимоги до організації робочого місця

Робочий стіл на досліджуваному місці містить достатньо простору для ніг. Крісло, що використовується в якості робочого сидіння, є підйомно-поворотним, має підлокітники і можливість регулювання за висотою і кутом нахилу спинки, також воно м'яке і виконане з екологічної шкіри, що дає можливість працювати у комфорті. Екран монітору знаходиться на відстані 0.8 м, клавіатура має можливість регулювання кута нахилу 5-15°. Отже, за всіма параметрами робоче місце відповідає нормативним вимогам.

Приміщення кабінету знаходиться на четвертому поверсі чотирьохповерхової будівлі і має об'єм 87,5 м³, площу — 25 м². У цьому кабінеті обладнано шість робочих місць, з них чотири укомплектовані ПК.

Температура в приміщенні протягом року коливається у межах 18–24°C, відносна вологість — близько 50%. Швидкість руху повітря не перевищує 0,2 м/с. Шум на робочому місці знаходиться на рівні 50 дБА. Система вентиляції приміщення — природна неорганізована, а опалення — централізоване.

Розміщення вікон забезпечує природне освітлення з коефіцієнтом природного освітлення не менше 1,5%, а загальне штучне освітлення, яке здійснюється за допомогою восьми люмінесцентних ламп, забезпечує рівень освітленості не менше 200 Лк.

У кабінеті є електрична мережа з напругою 220 В, яка створює небезпеку ураження електричним струмом. ПК та периферійні пристрої можуть бути джерелами електромагнітних випромінювань, аерозолів та шкідливих речовин (часток тонеру, оксидів нітрогену та озону).

За ступенем пожежної безпеки приміщення належить до категорії В. Кабінет має бути оснащений переносним вуглекислотним вогнегасником ВВК-5.

Наявна аптечка для надання долікарської допомоги, а також у кабінеті роблять вологе прибирання та щоденно провітрюють приміщення.

4.4 Рекомендації із пожежної профілактики

Пожежі в робочому приміщенні становлять небезпеку, тому що пов'язані як з матеріальними втратами, так і з відмовою засобів обчислювальної техніки.

Пожежа може виникати при внесенні джерела запалювання в горючу середу. Горючими матеріалами в приміщенні, де розташовані обчислювальні засоби є будівельні матеріали, віконні рами, двері, підлоги, меблі, ізоляція силових і сигнальних кабелів, радіотехнічні деталі,

конструктивні елементи з пластичних матеріалів, рідини для очищення елементів і вузлів ПЕОМ від забруднень.

- поліамід - матеріал корпусу мікросхеми, горюча речовина, температура самозаймання 420°C;

- полівінілхлорид - ізоляційний матеріал, горюча речовина, температура займання 335°C, температура самозаймання 530°C, теплота згоряння 18000-20700 кДж/кг;

- склотекстоліт ДЦ - матеріал друкованих плат, важко-горючий матеріал, показник горючості 1.74, не схильний до температурного самозаймання; пластик кабельний № 489 - матеріал ізоляції кабелю, горючий матеріал, показник горючості більше 2.1;

- деревина - будівельний і оздоблювальний матеріал, матеріал з якого виготовлені меблі, горючий матеріал, показник горючості більше 2.1, теплота згоряння 18731-20853 кДж/кг, температура займання 399°C, схильна до самозаймання [21].

Приміщення належить до категорії "В", як пожежонебезпечна. Приміщення відноситься до класу П-Па. Пожежа може виникнути в результаті утворення джерела запалювання (іскри й дуги короткого замикання, порушення ізоляції, що приводить до короткого замикання, перегріву радіодеталей внаслідок тривалого перевантаження) та внесення його до горючої середовища. При повному згорянні органічних сполук утворюється CO_2 , SO_2 , H_2O , N_2 , а при згорянні неорганічних сполук - оксиди. Залежно від температури плавлення продукції реакції диму можуть або знаходитися у воді розплаву (Al_2O_3 , TiO_2), або підійматися в повітря у вигляді диму (P_2O_5 , Na_2O , MgO). Розплавлені тверді частинки створюють світимість полум'я. Склад продуктів неповного згоряння горючих речовин складний і різноманітний. Це можуть бути горючі речовини - H_2 , CO , CH_4 та ін.; атомарний водень і кисень; різні радикали - OH , CN і інші. Продуктами неповного згоряння можуть бути також оксиди азоту, спирти альдегіди, кетони й високотоксичні з'єднання, наприклад, синильна кислота.

Для захисту персоналу від можливого отруєння проектом передбачається застосування протигаза з коробкою марки «В» (жовта) [22].

Пожежна безпека при застосуванні ЕОМ забезпечується:

- системою запобігання пожеж;
- системою протипожежного захисту;
- організаційно-технічними заходами.

До системи запобігання пожежі відносяться:

- запобігання утворенню горючого середовища;

- електроживлення ПЕОМ має автоматичне блокування відключення електроенергії на випадок зупинки системи охолодження й кондиціонування;

- система вентиляції обчислювальних центрів обладнуються блокувальними пристроями, що забезпечують її відключення в разі пожежі;

- система обладнується вогнеперепинюваченими клапанами;

- застосування обладнання задовольняє вимогам;

- забезпечення пожежної безпеки обладнання;

- після закінчення роботи, перед закриттям приміщення, все електроустановки відключаються і ПК від мережі електроживлення.

Зменшити горюче навантаження не є можливим, тому проектом передбачається застосування наступних способів і їх комбінацій для запобігання утворення (внесення) джерел запалювання:

- застосоване обладнання задовольняє вимогам електростатичної безпеки;

- застосування в конструкції швидкодіючих засобів захисного відключення можливих джерел запалювання;

- виключення можливості появи іскрового заряду статичної електрики в займистому середовищі з енергією, що дорівнює і вище мінімальної енергії запалювання;

- підтримання температури нагріву поверхні машин механізмів устаткування пристроїв речовин і матеріалів, які можуть увійти в контакт із горючим середовищем, нижче гранично допустимої, що становить 80% найменшої температури самозаймання пального.

Для зниження пожежної небезпеки проектом передбачається використовувати систему автоматичної пожежної сигналізації з одним димовим датчиком-оповіщувачем типу ВДМ-1М, який розрахований для контролю площі до 100 м² при висоті стелі до чотирьох метрів, а також первинні засоби пожежогасіння. В якості первинних засобів пожежогасіння передбачається використовувати:

- ручний вуглекислотний вогнегасник ОУ-5 - 1 шт .;

- повітряно-пінний вогнегасник ОВП-5 - 1 шт .;

- азбестове полотно 1.5 × 2 м.

В якості організаційно-технічних заходів рекомендується проводити навчання робочого персоналу пожежної безпеки.

4.5 Мікроклімат

У виробничому приміщенні на організм людини та його працездатність впливають мікрокліматичні фактори. Мікроклімат виробничих приміщень визначається поєднанням температури, вологості й швидкості руху повітря, а також температури навколишніх поверхонь. На даний час основним нормативним документом, щодо нормалізації мікроклімату є [23].

Для робіт категорії 1а для робочої зони виробничих приміщень забезпечуються наступні метеорологічні умови:

- в холодний і перехідний період року температура повітря $22 \div 24^{\circ}\text{C}$;
- відносна вологість повітря $40 \div 60\%$, швидкість руху повітря не більше 0.1 м/с ;
- у теплий період року температура повітря $23 \div 25^{\circ}\text{C}$, відносна вологість повітря $40-60\%$, швидкість руху повітря не більше 0.1 м/с .

Рівень шуму не перевищує санітарних норм. Тому застосування захисту від шуму в роботі не передбачається.

Ще одна проблема полягає в тому, що спектр випромінювання комп'ютерного монітора включає в себе рентгенівську, ультрафіолетову та інфрачервону області, а також широкий діапазон хвиль інших частот. Небезпека рентгенівських променів мізерно мала, оскільки цей вид випромінювання поглинається речовиною екрану. Однак велику увагу слід приділяти біологічним ефектам низькочастотних електромагнітних полів.

Для зниження стомлюваності персоналу в приміщеннях, де розташовані обчислювальні засоби, передбачається використовувати кольорові поєднання й покриття, що не дають відблисків.

У проекті, що розробляється передбачається використовувати сполучне освітлення. У світлий час доби приміщення буде висвітлюватися через віконні прорізи, в решту часу буде використовуватися штучне освітлення. Штучне освітлення створюється лампами розжарювання або газорозрядними лампами. Штучне освітлення в робочому приміщенні передбачається здійснювати з використанням люмінесцентних джерел світла в світильниках загального освітлення, оскільки люмінесцентні лампи мають високу світловою віддачею (до 75 лм/Вт і більше), тривалим терміном служби (до 10000 годин), спектральним складом випромінюваного світла, близьким до сонячного. При експлуатації комп'ютерів виробляється зорова робота IV в розряді точності (середня точність). При цьому нормована освітленість на робочому місці дорівнює 200 лк. Джерелом природного світла (освітлення) є сонячне світло. У приміщенні, де розташовані комп'ютери, передбачається природне бічне освітлення, рівень якого відповідає ДБН В.2.5-28-2018 " Природне і штучне освітлення" [24].

Регулярно проводиться контроль освітленості, який підтверджує, що рівень освітленості задовольняє ДБН і для даного приміщення в світлий час доби достатньо природного освітлення.

Розрахунок штучного освітлення проводиться за коефіцієнтами використання світлового потоку, яким визначається потік, необхідний для створення заданої освітленості при загальному рівномірному освітленні.

Розрахунок кількості світильників здійснюється за формулою:

$$N = E \cdot S \cdot Z \cdot K / (F \cdot U \cdot M), \quad (4.9)$$

де N - число світильників;

E - нормоване освітлення;

S - площа підлоги, м², S=25 м²;

Z - поправний коефіцієнт світильника (Z = 1,15 для ламп розжарювання та ДРЛ; Z = 1,1 для люмінесцентних ламп) приймаємо рівним 1,1;

K - коефіцієнт запасу, що враховує зниження освітленості в процесі експлуатації – 1,5;

U - коефіцієнт використання, що залежить від типу світильника, показника індексу приміщення і т. п. - 0,575;

M - число люмінесцентних ламп у світильнику - 2;

F - світловий потік – 5400лм (для ЛБ - 80).

Згідно вимог ДБН В.2.5-28-2018 [6], освітлення робочого місця оператора обчислювальної техніки повинно бути не менше 200 лк.

$$N = 200 \cdot 25 \cdot 1.1 \cdot 1.5 / (5400 \cdot 0.575 \cdot 2) = 1.33 \approx 1 \quad (4.10)$$

Обираємо кількість світильників, що дорівнює 1.

4.6 Охорона навколишнього природного середовища

Діяльність за темою магістерської роботи, а саме робота за комп'ютером в процесі її виконання впливає на навколишнє природне середовище і регламентується нормами діючого законодавства.

Основним екологічним аспектом в процесі діяльності за даними спеціальностями є процеси впливу на атмосферне повітря та процеси поводження з відходами, які утворюються, збираються, розміщуються, передаються на видалення (знешкодження), утилізацію, тощо в ІТ галузі.

Вплив на атмосферне повітря при нормальних умовах праці не оказує, бо не має в приміщенні сканерів, принтерів та інших джерел викиду забруднюючих речовин в повітря робочої зони.

В процесі діяльності комп'ютера виникають процеси поводження з відходами ІТ галузі. Нижче надано перелік відходів, що утворюються в процесі роботи:

Відпрацьовані люмінесцентні лампи - I клас небезпеки

Батарейки та акумулятори (малі) -III клас небезпеки

Змінні носії інформації - IV клас небезпеки

Відходи друкуючих пристроїв - IV клас небезпеки

Відпрацьований ізолюючий матеріал, дроти та кабелі - IV клас небезпеки

Макулатура - IV клас небезпеки

Побутові відходи - IV клас небезпеки

Відходи в міру їх накопичення збирають у тару, відповідну класу небезпеки, з дотриманням правил безпеки, після чого доставляють до місця тимчасового зберігання відходів відповідно до затвердженої схеми їх розміщення. Зазначені для зберігання відходів місця чи об'єкти повинні використовуватися лише для заявлених відходів.

Не допускається зберігання відходів у невстановлених схемою місцях, а також перевищення норм тимчасового зберігання відходів.

Способи тимчасового зберігання відходів визначаються видом, агрегатним станом і класом небезпеки відходів:

- відходи I класу небезпеки зберігаються в герметичній тарі (сталеві бочки, контейнери). У міру наповнення тару з відходами закривають герметично сталевий кришкою;

- відходи III класу небезпеки зберігаються в тарі, яка забезпечує локалізацію зберігання, дозволяє виконувати вантажно-розвантажувальні і транспортні роботи і виключає поширення в ОС шкідливих речовин;

- відходи IV класу небезпеки можуть зберігатися відкрито на промисловому майданчику у вигляді конусоподібної купи, звідки їх автовантажувачем перевантажують у самоскид і доставляють на місце утилізації або захоронення.

4.7 Висновки до розділу 4

В результаті проведеної роботи було зроблено аналіз умов праці, шкідливих та небезпечних чинників, з якими стикається робітник. Було визначено параметри і певні характеристики приміщення для роботи над запропонованим проектом написаному в дипломній роботі, описано, які заходи потрібно зробити для того, щоб дане приміщення відповідало необхідним нормам і було комфортним і безпечним для робітника. Приведені рекомендації щодо організації робочого місця, а також важливу інформацію щодо пожежної та електробезпеки. Були наведені розміри приміщення та значення температури, вологості й рухливості повітря, необхідна кількість і потужність ламп та інші параметри, значення яких впливає на умови праці робітника, а також – наведені інструкції з охорони праці, техніки безпеки при роботі на комп'ютері.

А також визначені основні екологічні аспекти впливу на навколишнє природне середовище та зазначені заходи щодо поводження з ними.

ВИСНОВКИ

Метою дипломної роботи визначено підвищення якості розпізнавання природнього мовлення за рахунок визначення вагомих факторів, що впливають на якість розпізнавання природнього мовлення з вживанням вузькоспеціалізованої специфічної термінології.

В ході дослідницької частини роботи були отримані наступні результати:

1. Проведено аналіз основних проблем розпізнавання мовлення.
2. Виконано аналіз галузі застосування, методів обробки й алгоритмів розпізнавання природнього мовлення.
3. Проведено аналіз використання глибоких нейронних мереж для розпізнаванні природнього мовлення.

4. Розглянуто математичні основи роботи глибоких нейронних мереж.

В ході практичної частини роботи були отримані наступні результати:

1. Проведення експерименту за першим сценарієм із використанням нейронних мереж.
2. Проведення експерименту за другим сценарієм із використанням нейронних мереж.
3. Порівняння результатів експериментів.
4. Оцінка отриманих результатів із використанням узагальненого коефіцієнту помилок у словах.
5. Зроблено висновки по проведеній реалізації методів і алгоритмів розпізнавання природнього мовлення.
6. Доведено залежність якості розпізнавання природнього мовлення від вихідних даних моделей: кількості файлів для навчання, якості запису аудіо файлів, акценту.

Таким чином, вирішено поставлене завдання - досліджено методи й алгоритми розпізнавання природнього мовлення для підвищення якості розпізнавання за рахунок визначення вагомих факторів, що впливають на якість розпізнавання природнього мовлення з вживанням вузькоспеціалізованої специфічної термінології.

ПЕРЕЛІК ПОСИЛАНЬ

1. Сфери застосування NLP [Електронний ресурс]//Режим доступу www. URL <https://neurohive.io/ru/osnovy-data-science/5-metodov-v-nlp-kotorye-izmenjat-obshhenie-v-budushhem/>
2. Завдання, які вирішує NLP [Електронний ресурс]//Режим доступу www. URL https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp.html
3. Що являє собою NLP [Електронний ресурс]//Режим доступу www. URL https://en.wikipedia.org/wiki/Natural_language_processing
4. Додатки, в яких використовується NLP [Електронний ресурс]//Режим доступу www. URL <https://towardsdatascience.com/introduction-to-natural-language-processing-nlp-323cc007df3d>
5. Методи NLP [Електронний ресурс]//Режим доступу www. URL <https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1>
6. NLP в медицині [Електронний ресурс]//Режим доступу www. URL <https://winterlightlabs.com/>
7. Основні проблеми розвитку NLP [Електронний ресурс]//Режим доступу www. URL <https://ruder.io/4-biggest-open-problems-in-nlp/>
8. Методи процесу обробки природнього мовлення [Електронний ресурс]//Режим доступу www. URL <https://witanworld.com/article/2018/10/28/naturallanguageprocessing-nlp/>
9. Приховані марковські моделі https://en.wikipedia.org/wiki/Hidden_Markov_model
10. Yoav Goldberg Neural Network Methods for Natural Language Processing, pp. 282, 2017.
11. Christiane Fellbaum. WordNet: An Electronic Lexical Database. Bradford Books, 2017.
12. Christopher Manning, Prabhakar Raghavan, and Hinrich Schütze. Introduction to Information Retrieval. Cambridge University Press, 2017
13. P. J. Werbos. Backpropagation through time: What it does and how to do it. Proc. of the IEEE, 1990
14. Wang Ling, Chris Dyer, Alan W. Black, and Isabel Trancoso. Two/too simple adaptations of Word2Vec for syntax problems. In Proc. of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 1299–1304, Denver, Colorado, 2015b.
15. Wenduan Xu, Michael Auli, and Stephen Clark. CCG supertagging with a recurrent neural network. In Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th

International Joint Conference on Natural Language Processing—(Volume 2: Short Papers), pages 250–255, Beijing, China, July 2015.

16. Rafal Jozefowicz, Oriol Vinyals, Mike Schuster, Noam Shazeer, and Yonghui Wu. Exploring the limits of language modeling. arXiv:1602.02410 [cs], February 2016.

17. Xuezhong Ma and Eduard Hovy. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF. In Proc. of the 54th Annual Meeting of the Association for Computational Linguistics-Volume 1: Long Papers), pages 1064–1074, Berlin, Germany, August 2016.

18. Опис DeepSpeech <https://hacks.mozilla.org/2019/12/deepspeech-0-6-mozillas-speech-to-text-engine/>

19. НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Зареєстровано в Міністерстві юстиції України 25 квітня 2018 року за № 508/31960. Режим доступу: www. URL: <https://zakon.rada.gov.ua/laws/show/z0508-18>

20. ДСТУ Б А.3.2-13:2011 «Будівництво. Електробезпека. Загальні вимоги». Наказ від 29 грудня 2011 року № 405. Режим доступу: www. URL: <https://zakon.rada.gov.ua/rada/show/z0072-13/paran21>

21. Закон України «Про забезпечення санітарного та епідемічного благополуччя населення». Вводиться в дію Постановою ВР № 4005-ХІІ від 24 лютого 1994 року, ВВР, 1994, № 27, ст. 219. Режим доступу: www. URL: <https://zakon.rada.gov.ua/laws/show/4004-12>

22. НПАОП 0.00-1.04-07 «Правила вибору та застосування засобів індивідуального захисту органів дихання». Зареєстровано в Міністерстві юстиції України 4 квітня 2008 року за № 285/14976. Режим доступу: www. URL: <https://zakon.rada.gov.ua/laws/show/z0285-08>

23. ГОСТ 12.1.004-91 ССБТ «Пожежна безпека. Загальні вимоги». Постанова від 14 червня 1991 року №875. Режим доступу: www. URL: <https://zakon.rada.gov.ua/laws/show/z0252-15>

24. ДБН В.2.5-28-2018 «Природне і штучне освітлення». Наказ від 3 жовтня 2018 року № 264. Режим доступу: www. URL: https://dbn.co.ua/load/normativy/dbn/dbn_v_2_5_28/1-1-0-1188

Додаток А

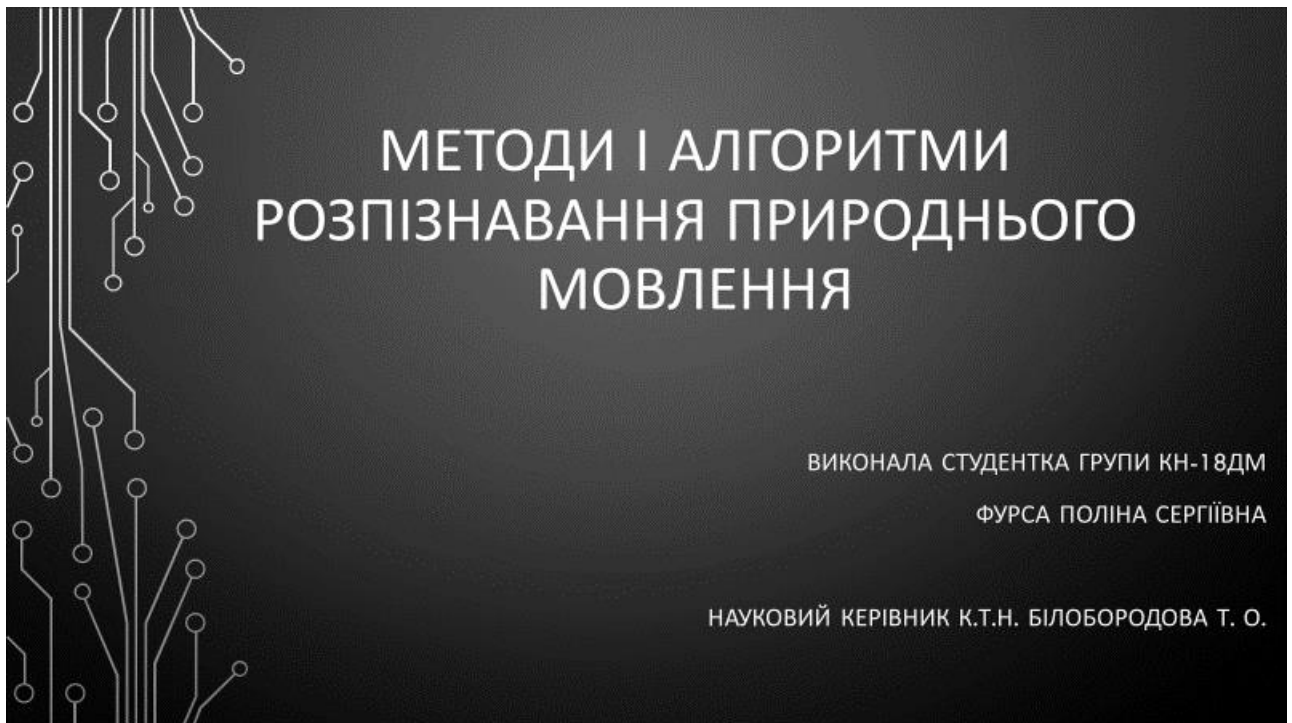


Рисунок А.1 – Слайд 1



Рисунок А.2 – Слайд 2

АНАЛІЗ ПРОБЛЕМ ЗАСТОСУВАННЯ ОБРОБКИ ПРИРОДНОГО МОВЛЕННЯ

- неоднозначність речень;
- позначка частин мови та створення графіків залежностей;
- створення відповідного словникового запасу;
- зв'язування різних компонентів словникового запасу;
- установка контексту;
- виділення смислових значень.

Рисунок А.3 – Слайд 3

АНАЛІЗ МЕТОДІВ ПРОЦЕСУ ОБРОБКИ ПРИРОДНОГО МОВЛЕННЯ

На першому етапі виконується лексичний аналіз, у ході якого з тексту програми послідовно зчитуються лексеми, або токени - мінімальні неподільні одиниці мови. До таких неподільних одиниць відносяться слова, роздільники, знаки алгебраїчних операцій і т. д. Отриманий ланцюг токенів записується компілятором у спеціальну таблицю, в якій для кожного токена вказуються його код, тип (число, рядок, оператор, ідентифікатор, роздільник і т.д.), деякі інші характеристики. На другому етапі виконується синтаксичний аналіз, у ході якого проводиться розбір отриманої послідовності токенів. В результаті розбору формується абстрактне синтаксичне дерево, що є проміжною і, як впливає з назви, абстрактною формою представлення програми. Абстрактне синтаксичне дерево не є буквального відображенням коду, яке воно представляє. Воно містить тільки необхідні дані для запису операцій. На третьому етапі виконується генерація машинного коду. Спочатку абстрактне синтаксичне дерево перетворюється в лінеаризоване абстрактне представлення, наприклад, у трьохадресний код, кожна команда якого складається не більше ніж із трьох операндів. Потім лінеаризоване представлення перекладається в машинний код.



Рисунок А.4 – Слайд 4

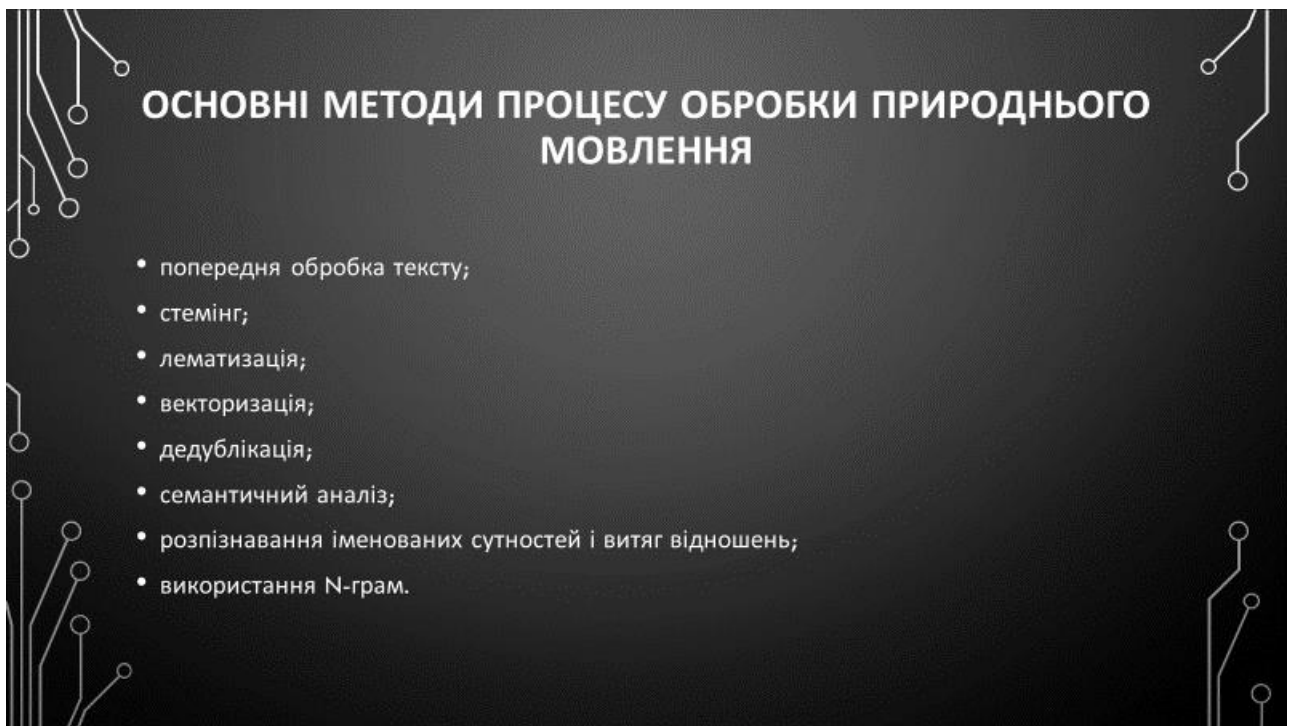


Рисунок А.5 – Слайд 5

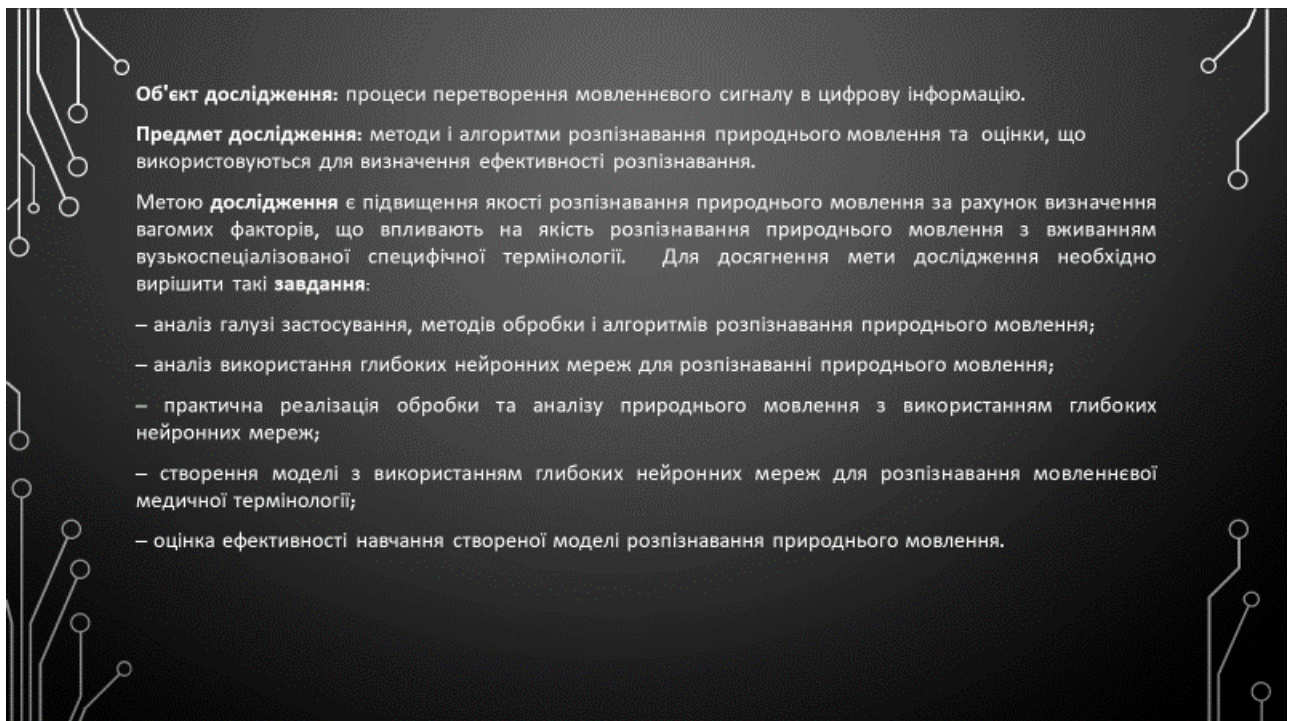


Рисунок А.6 – Слайд 6

РЕКУРЕНТНА НЕЙРОННА МЕРЕЖА

У рекурентної штучної нейронної мережі прихований шар зберігає всю попередню історію, таким чином, розмір контексту необмежений

Мережа має вхідний шар x , прихований шар z (також званий контекстним шаром або станом) і вихідний шар y . Вхідний шар складається з вектора $x(t)$, який є об'єднанням вектора $w(t)$, що представляє собою поточне слово, і вектора $z(t-1)$, який представляє собою вихідні значення прихованого шару, отримані на попередньому кроці. Розмір вектора $w(t)$ дорівнює розміру словника. Вихідний шар $y(t)$ має таку ж розмірність, як і $w(t)$, і після навчання нейронної мережі є ймовірнісний розподіл наступного слова при даному попередньому слові і стані прихованого шару в попередній часовий крок. Розмір прихованого шару зазвичай вибирається емпірично.

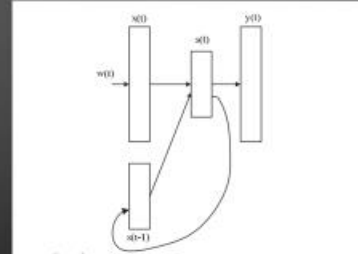


Рисунок А.7 – Слайд 7

АРХІТЕКТУРА МЕРЕЖІ

На рисунку представлена архітектура, що складається зі змінних компонентів. Подана система - це рекурентна нейронна мережа (RNN) з одним або декількома згортковими вхідними шарами, за ним слідує кілька повторюваних (односпрямованих або двоспрямованих) шарів і один повністю зв'язаний шар перед softmax шаром. Мережа навчається наскрізним способом із використанням функції втрати нейромережевої часової класифікації (НЧК, Connectionist temporal classification), яка дозволяє безпосередньо прогнозувати послідовності символів із вхідного аудіо.

Входи в мережу представляють собою послідовність лог-спектрограм нормалізованих за потужністю звукових аудіокліпів на 20 мс вікна. Виходи - алфавіт кожної мови. На кожному виході тимчасовий крок t . Рекурентна нейронна мережа (RNN) робить прогноз, $p(t_i|x)$, де t_i це або символ в алфавіті, або порожній символ. Англійською у нас $t_i \in \{a, b, c, \dots, z, \text{пробіл}, \text{апостроф}, \text{пропуск}\}$, де додали символ пробілу для позначення границь слова. Під час виводу моделі, НЧК з'єднуються з мовою.



Рисунок А.8 – Слайд 8

ПАКЕТНА НОРМАЛІЗАЦІЯ ДЛЯ ГЛИБОКИХ РНМ

Для ефективного поглинання даних при масштабуванні навчального набору збільшується глибина мереж, шляхом додавання більшої кількості повторюваних шарів. Проте, стає все складніше навчати мережі, що використовують градієнтний спуск у міру збільшення розміру й глибини. При застосуванні глибоких мереж РНМ, на великих наборах даних, варіант використання пакетної нормалізації, істотно покращує остаточну помилку узагальнення на додаток до прискорення навчання. При застосуванні до глибоких мереж РНМ, на великих наборах даних, варіант використання пакетної нормалізації істотно покращує остаточну помилку узагальнення на додаток до прискорення навчання. Рекурентний шар реалізований як:

$$h_t^l = f(W^l h_{t-1}^{l-1} + U^l h_{t-1}^l + b)$$

Є два способи застосування пакетної нормалізації для періодичної операції. Це вставити перетворення пакетної нормалізації, $\mathcal{B}()$, перед кожною нелінійністю наступним чином:

$$h_t^l = f(\mathcal{B}(W^l h_{t-1}^{l-1} + U^l h_{t-1}^l))$$

В цьому випадку статистика середнього й дисперсії накопичується за один часовий інтервал міні-пакета. Це не є ефективним. Альтернативою є нормалізувати пакетно тільки вертикальні зв'язки. Періодичне обчислення задається

$$h_t^l = f(\mathcal{B}(W^l h_{t-1}^{l-1}) + U^l h_{t-1}^l)$$

Рисунок А.9 – Слайд 9

КРИТЕРІЇ ОЦІНЮВАННЯ ЯКОСТІ РОЗПІЗНАВАННЯ ПРИРОДНОГО МОВЛЕННЯ

Розпізнавання мовлення оцінюється за двома критеріями: коефіцієнт помилок в словах - КПС (англ. Word Error Rate -WER) та затримкою (англ. Latency).

1) **Коефіцієнт помилок в словах (КПС)** вимірює невідповідність рівня слів: він порівнює слова, видані розпізнавачем, з тими, які користувач насправді вимовив. Кожна помилка (заміна, вставка або видалення) враховується в розпізнавальнику. КПС можна обчислити так:

$$\text{КПС} = \frac{\text{кількість замін} + \text{кількість вставок} + \text{кількість видалень}}{\text{загальна кількість слів}} * 100\%$$

2) Затримка

Затримка визначається як загальний час (у секундах), необхідний для розпізнавання мовлення. Точніше, затримка є часом з моменту закінчення запису до моменту появи розпізнавання на екрані.

Рисунок А.10 – Слайд 10

ПРАКТИЧНА РЕАЛІЗАЦІЯ РОЗПІЗНАВАННЯ ПРИРОДНЬОГО МОВЛЕННЯ З ВИКОРИСТАННЯМ ГЛИБОКИХ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ

Апаратне забезпечення, що використане для проведення експерименту з практичної реалізації методів і алгоритмів розпізнавання природнього мовлення, має наступні характеристики:

- 1) Операційна система – 64-розрядна, Ubuntu 18.04.
- 2) Оперативна пам'ять – 24 Гб.
- 3) Процесор – AMD Ryzen 5 1600(3,8 ГГц).
- 4) Відеокарта Nvideo GeForce GTX 1060 3GB

Практична реалізація проведена з використанням **Mozilla DeepSpeech** за двома сценаріями.

Рисунок А.11 – Слайд 11

ПРАКТИЧНА РЕАЛІЗАЦІЯ РОЗПІЗНАВАННЯ ПРИРОДНЬОГО МОВЛЕННЯ З ВИКОРИСТАННЯМ ГЛИБОКИХ РЕКУРЕНТНИХ НЕЙРОННИХ МЕРЕЖ

Перший сценарій включає наступні умови та дії:

- використання навченої моделі, що пропонується Mozilla Deep Speech;
- запис аудіо файлів для тестування моделі за участі 15 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

Другий сценарій включає наступні умови та дії:

- підготовка даних для створення моделі, що включає створення бібліотеки аудіо файлів, які містять мовленнєві елементи медичного напрямку за участі 45 добровольців;
- створення моделі, що включає визначення архітектури та конфігурації мережі для навчання моделі та безпосередньо навчання моделі з використанням підготовлених даних, що складаються з записаних аудіо файлів, які містять мовленнєві елементи медичного напрямку;
- запис аудіо файлів для тестування моделі за участі 5 добровольців;
- підготовка та обробка записаних аудіо файлів;
- тестування моделі з використанням підготовлених аудіо файлів.

Рисунок А.12 – Слайд 12

ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ З РОЗПІЗНАВАННЯ ПРИРОДНОГО МОВЛЕННЯ ЗА ПЕРШИМ СЦЕНАРІЄМ

У першому сценарії використовувалася стандартна модель для створення якої було залучено 39577 добровольців. З них 67 % осіб чоловічої статі та 33 % осіб жіночої статі.

Дані, які використовувались для навчання стандартної моделі:

- сумарний об'єм аудіофайлів становить - 30 ГБ;
- загальна сума годин – 1087;
- акценти: 23% - США, 9% - Великобританія, 4% - Індія та Південна Азія (Індія, Пакистан, Шрі-Ланка), 3% - Австрія, 3% - Канада, 1% - Шотландія, 1% - Південна Африка (Південна Африка, Зімбабве, Намібія), 1% - Нова Зеландія.

Рисунок А.13 – Слайд 13

ТЕСТУВАННЯ МОДЕЛІ

№	Текст	Вік (років)	Стать	Акцент	Текст на виході	Коефіцієнт помилки слова (%)
1	down below in the darkness were hundreds of people sleeping in peace	20	чоловік	США	then below and the darkness were hundreds of people sleeping in peace	17
2	this was the strangest of all things that ever came to earth from outer space	30	чоловік	Великобританія	this was the strangest of all things that ever came to earth from outer space	13
3	follow the instructions here	20	чоловік	Шотландія	ful of these structions here	100
4	he didn't need to seek out the old woman for this	20	жінка	США	you didn't need to seek out the old woman for this	9
5	don't point the gun at me	20	жінка	Індія	don't hold the gun at me	0
6	i told you to have the ice box fixed	50	жінка	Австралія	i don't do it to him the eye both ex	78
7	but finally the merchant appeared and asked the boy to shear four sheep	30	чоловік	Індія	but finally the reather appeared and as the boy was sheared for sheep	46
8	most meerkats are more or less rounded	20	жінка	Великобританія	not me i am more less rounded	71
9	from here on you will be alone the accountant said	30	чоловік	Канада	from hereon you will be alone the accountant said	30
10	my tax dollars pay for those public school prons	40	чоловік	Нова Зеландія	my text dollars pay for those public school prons	22
11	an army is coming the boy said	20	жінка	Вікторіні	an army colling the voice that	57
12	did the discuss it with you	30	жінка	Ірландія	did the discuss it with you	0
13	i'm crazy about edison	30	чоловік	Гонконг	i am crazy about engwing	35
14	he decided to spend the night there	підліток	чоловік	Африка	he decided to spend the night there	0
15	he's in the insurance business	30	чоловік	Південна Африка	he's in the insurance business	0
						39

Рисунок А.14 – Слайд 14

ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ З РОЗПІЗНАВАННЯ ПРИРОДНОГО МОВЛЕННЯ ЗА ДРУГИМ СЦЕНАРІЄМ

У другому сценарії модель навчалась на наборі для створення якого було залучено 45 добровольців. З них 76 % осіб чоловічої статі та 24 % осіб жіночої статі.

Дані, які використовувались для навчання стандартної моделі:

- сумарний об'єм аудіофайлів становить - 302 МБ;
- загальна сума годин – 3;
- акценти: 100% - Україна.

Рисунок А.15 – Слайд 15

№	Фраза	Розпізнана фраза	$t_{\text{розп.с}}$	НДС, %	№	Фраза	Розпізнана фраза	$t_{\text{розп.с}}$	НДС, %
1	Left traction up here	Left catching up here	0.675	25	11	Left move anterior above bladder	left must endure above blade	0.745	80
		left direction up he	0.687	50			let me and darrel about la	0.713	100
		a less traced appose	0.695	100			oh lattimore and dare a ball blow	1.011	100
		left decompose	0.603	100			left a very orebo later	0.862	80
2	Right move toward Manipulator left	miss curran up her	0.730	75	12	Right move up	love to it and far about lat	1.093	100
		reduction down here	0.815	100			right movat	0.492	66
		rise rection dan here	0.726	83			right mit	0.614	66
		right dragon down to her	0.819	50			written out	0.585	100
3	Right traction down here	achard clainthe	0.675	100	13	Right move down the bladder	right mobile	0.582	66
		right action down here	0.814	83			ride a	0.662	100
		reduction down here	0.815	50			right don't said winter	0.778	80
		rise rection dan here	0.638	75			right moved down a blade	0.714	60
4	Left close	right dragon down to her	0.759	75	14	Left move toward Manipulator left	writing down the bl	0.628	80
		achard clainthe	0.618	100			all right move down to blot her	0.786	60
		right action down here	0.755	50			at dooble	0.594	100
		let lose	0.552	100			left more to god do inchie	1.146	83
5	Left grasp here	left close	0.618	0	15	Assitant Left traction up seminal vesicle	her eyes moved over in chief	0.815	100
		the left close	0.591	33			left more tread the reef	0.718	83
		left	0.499	50			left move toward the we dashed	0.807	50
		left close	0.581	0			ah to obtain sure	0.657	100
6	Left traction down here	left grasped her	0.611	66	16	Right move down the bladder	left action up sembovere	0.815	60
		left dress here	0.616	33			s left direction up simon with thick mist	0.851	87.5
		the lad grasped her	0.636	100			lean direction up seven over ice	0.808	83
		let earth pear	0.643	100			let the extra up seven recte	0.952	100
7	Right move up here	lost trap here	0.727	66	17	Left grasp prostate	this traction up singing out there beside	1.190	71
		left direction down hole	0.699	50			right moved down a blade	0.699	60
		left direction down here	0.754	25			writing down the bl	0.603	80
		the last train down here	0.704	50			all right move down to blot her	0.802	57
8	Left open	laughed to condone her	0.737	100	18	Left move out the prostate	at dooble	0.609	100
		lived traction down here	0.767	25			right on to the platter	0.850	80
		her eyes moved up here	0.670	50			lived gross rose	1.033	100
		writing of her	0.597	100			left grasped prostate	0.691	33
9	Right move toward da Vinci right	or i too up her	0.694	100	19	Left grasp the needle	clar to grasp prostate	0.735	50
		at both apperture	0.625	100			left it for state	0.873	75
		right more up here	0.670	25			left great thrust at	0.740	75
		left open	0.606	0			left me all the pro	0.806	60
10	Left stay here	left open	0.603	0	20	Right cut here	left mobile sir state	0.713	80
		let open	0.540	50			let move out the prostate	0.782	20
		left open	0.585	0			left more at a petase	0.770	80
		left open	0.584	0			he moved more houseboat	0.814	100
		right moved toward the inter it	0.785	33			lived up in de	0.711	100
		right moved toward the insure it	0.804	33			let dropped the need	0.715	100
		bride i dread the writer it	0.723	100			the left grasps any do	0.627	80
		or i more dread the writer i	0.787	100			left grasps and he do	0.700	50
		or i to move to the red right	0.800	100			we grasped in no	0.687	100
		he left that here	0.626	33			her right got here	0.647	50
		left they hear	0.642	33			right god he	0.674	33
		the left they here	0.603	33			it caught the her	0.558	100
		let they he	0.557	100			at that her	0.736	100
		well i say	0.600	100			it gut her	0.696	100

Рисунок А.16 – Слайд 16

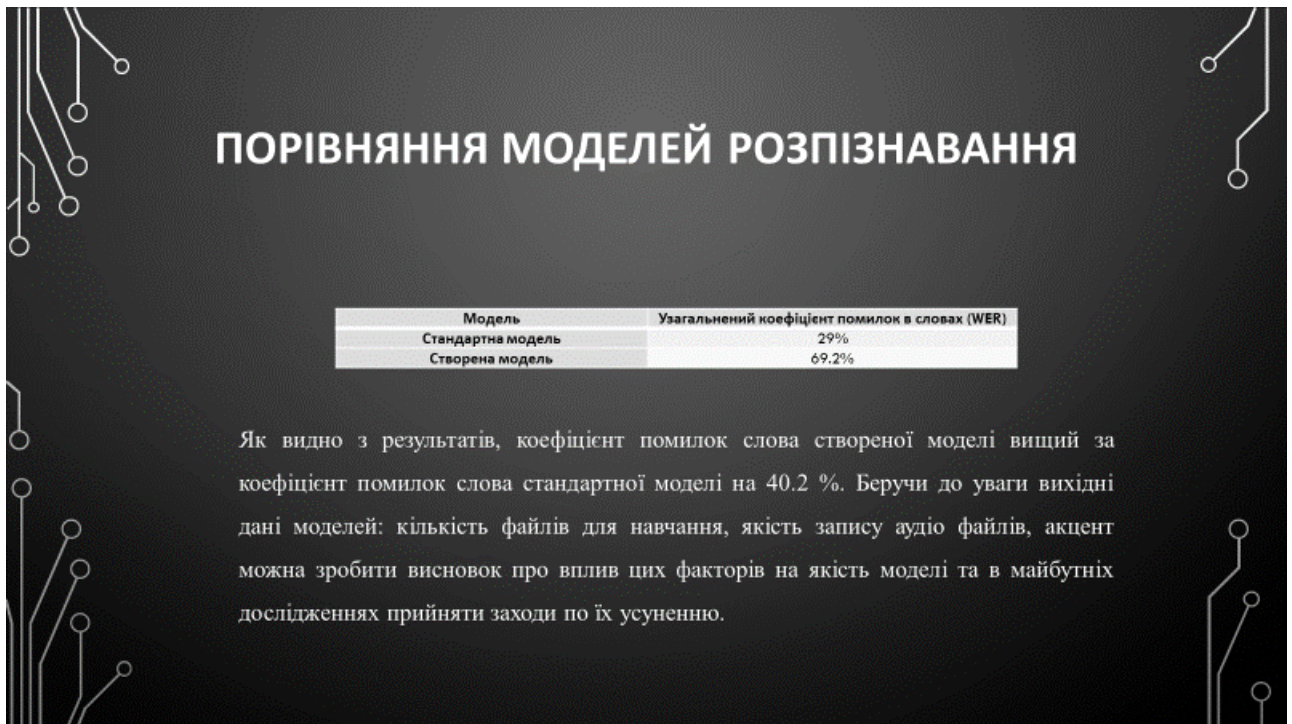


Рисунок А.17 – Слайд 17

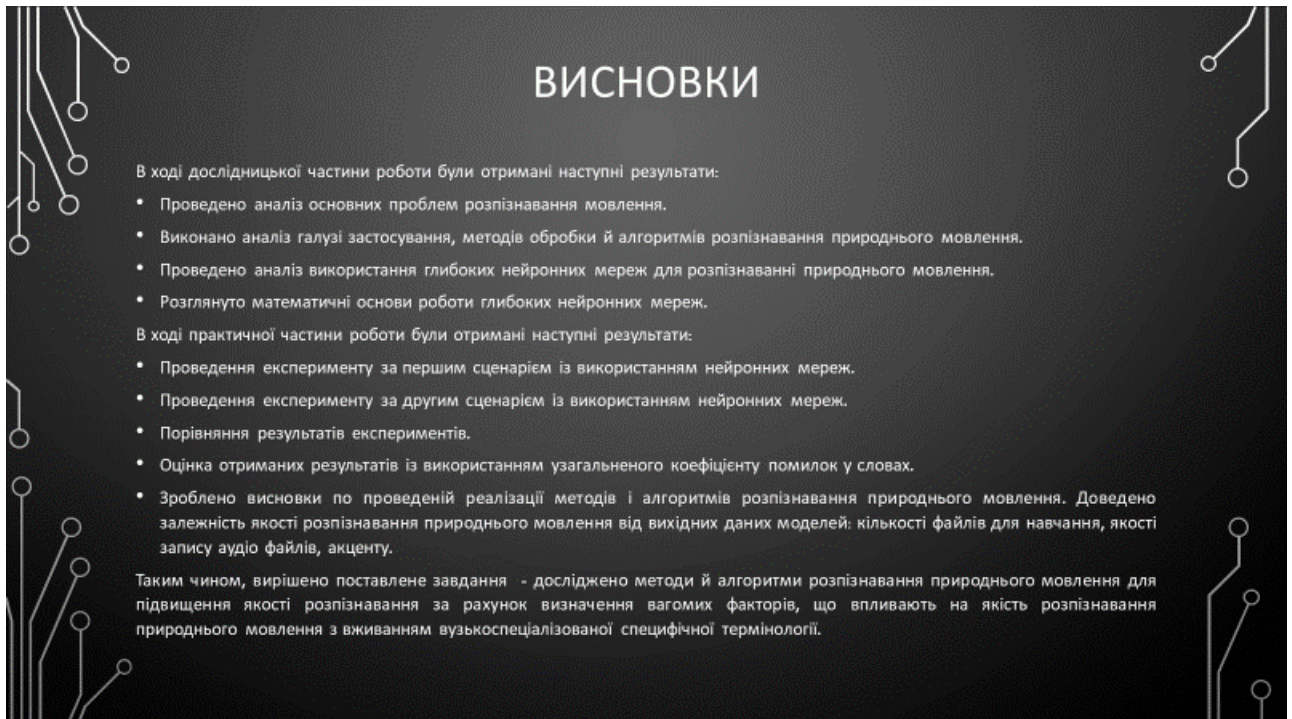


Рисунок А.18 – Слайд 18