ДОДАТОК А

Перелік джерел посилання за науковими напрямами керівника та науковців кафедри програмної інженерії

1. Vyrodov K., Chupryna A., Kotelnykov R. Detecting of Anti-Ukrainian Trolling Tweets //Proceedings http://ceur-ws. org ISSN. – 2023. – Т. 1613. – С. 0073.

ДОДАТОК Б

Звіт результатів перевірки на унікальність тексту в базі ХНУРЕ

UNICHECK
by Turnitin

Ім'я користувача:
Кардаш Євген Вікторович каф.ПІ

ID перевірки:
1016341545

Дата перевірки:
10.06.2024 10:53:18 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
10.06.2024 11:02:44 EEST

ID користувача:
100013622

Назва документа: 2024_М_ПІ_ІПЗм-22-4_Дубок_В_Ю_скорочений

Кількість сторінок: 22 **Кількість слів:** 6809 **Кількість символів:** 52409 **Розмір файлу:** 4.26 MB **ID файлу:** 1016142797

6.51%
Схожість

Найбільша схожість: 4.96% з Інтернет-джерелом (https://uk.wikipedia.org/wiki?curid=81909)

6.24% Джерела з Інтернету | 98 ..... Сторінка 24

2.69% Джерела з Бібліотеки | 38 ..... Сторінка 24

0% Цитат

Вилучення цитат вимкнене

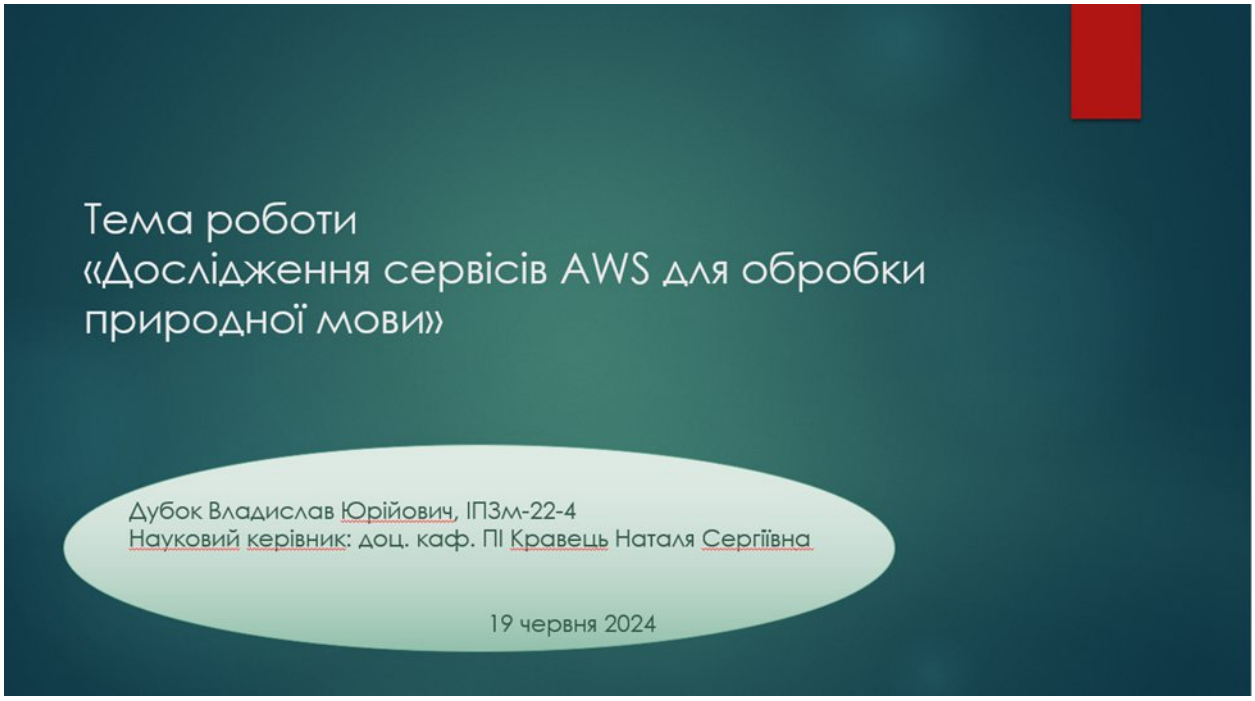
Вилучення списку бібліографічних посилань вимкнене

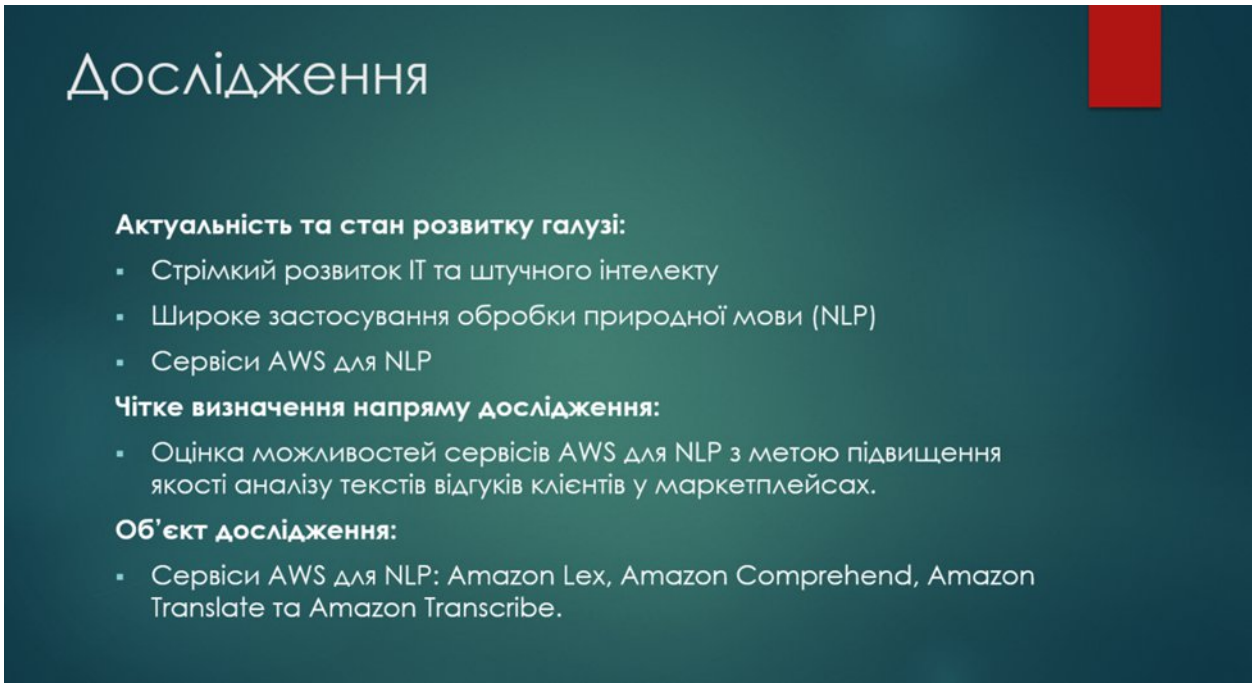
0%
Вилучень

Немає вилучених джерел

ДОДАТОК В

Слайди презентації





Постановка задачі

- **Чітке формулювання проблеми:**
  - Аналіз текстів відгуків клієнтів є трудомістким та складним завданням через великий обсяг даних та їх неоднозначність.
  - Автоматизація аналізу текстів відгуків може значно підвищити якість обслуговування та оптимізацію бізнес-процесів.
- **Опис очікуваних результатів:**
  - Визначення можливостей та обмежень сервісів AWS для NLP.
  - Розробка програмної системи для автоматизації аналізу відгуків користувачів з використанням сервісів AWS.

Методологія

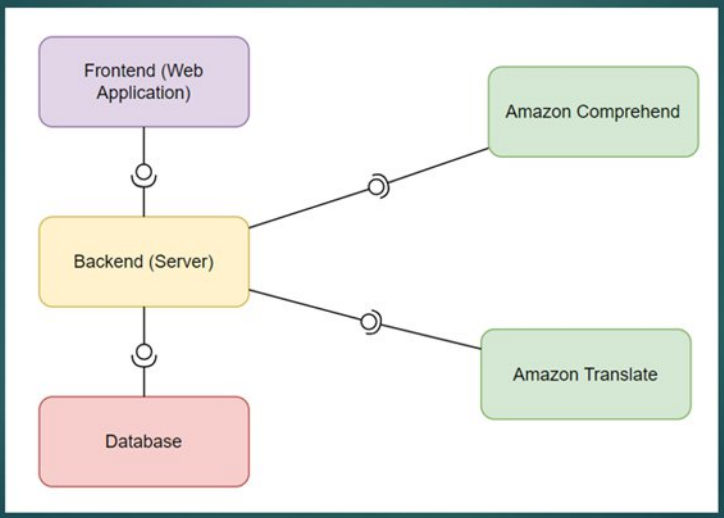
Опис використаних методів дослідження:

- **Емпіричні методи**
- **Порівняльні методи**

Інструментарій та технології, використані в роботі:

- **Amazon Comprehend:** Аналіз тональності, виявлення ключових слів, класифікація відгуків.
- **Amazon Translate:** Машинний переклад текстів для обробки відгуків різними мовами.
- **Amazon Transcribe:** Конвертація мовлення в текст для аналізу аудіо відгуків.
- **Amazon Lex:** Створення голосових та текстових чат-ботів для взаємодії з користувачами, які можуть автоматично відповідати на запити клієнтів та забезпечувати зворотній зв'язок у режимі реального часу.

Архітектура програмної системи



Опис програмного забезпечення, що було використано у дослідженні

Опис процесу розробки:

- Розробка веб-додатка для взаємодії з користувачами.
- Розробка серверної частини для обробки відгуків користувачів та інтеграції з AWS.
- Створення бази даних для зберігання результатів аналізу.

Вибрані мови програмування та фреймворки:

- **Backend:** C# (ASP.NET) для серверної частини.
- **Frontend:** JavaScript (Angular) для веб-додатка.
- **Database:** PostgreSQL для зберігання даних.

Зміст проведеного експерименту

- **Методи:**
  - Аналіз текстів з різним розміром та стилем для визначення точності роботи сервісів.
- **Вхідні дані:**
  - Розмовні тексти: відгуки користувачів.
  - Художні тексти: літературні твори, аудіо книги.
  - Публіцистичні тексти: статті, репортажі, конференції.
- **Критерії:**
  - Точність настрою.
  - Точність перекладу.
  - Точність розпізнавання мовлення.
  - Час виконання запиту.
  - Розмір тексту.

Результати порівняльного аналізу сервісів

Назва сервісу	Призначення	Ключові особливості	Випадки використання	Кількість підтримуваних мов
Amazon Lex	Створення інтерфейсів для розмов за допомогою голосу та тексту	Автоматичне розпізнавання мови, розуміння природної мови, підтримка багатоплатформності	Чатботи, віртуальні помічники, додатки з підтримкою голосу	27
Amazon Comprehend	Сервіс обробки природної мови (NLP) для отримання інсайтів з тексту	Розпізнавання сутностей, аналіз настроїв, визначення мови, моделювання тем	Аналіз тексту, аналіз відгуків клієнтів, класифікація документів	9
Amazon Translate	Сервіс перекладу мов	Переклад в реальному часі та пакетний переклад, користувацька термінологія, автоматичне визначення	Переклад документів, вебсайтів, комунікація в реальному часі	76
Amazon Transcribe	Автоматичне розпізнавання мови (ASR) для перетворення мови в текст	Транскрипція в реальному часі та пакетна транскрипція, ідентифікація спікерів, користувацький словник	Транскрипція зустрічей, дзвінків в службу підтримки клієнтів, генерація контенту	31

Результати експерименту

Таблиця 4.1 – Результати проведених експериментів з Amazon Comprehend (таблиця виконана самостійно)

№	Точність настрою (%)	Розмір тексту (Кількість символів)	Час виконання (мс)	Стиль мовлення
1	61,98	395	781	Розмовний
2	81,43	4732	1440	Розмовний
3	64,02	423	832	Публіцистичний
4	88,94	4490	1562	Публіцистичний
5	86,75	359	935	Художній
6	94,12	4376	1522	Художній

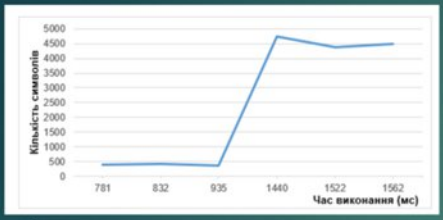
Таблиця 4.2. – Результати проведених експериментів з Amazon Translate (таблиця виконана самостійно)

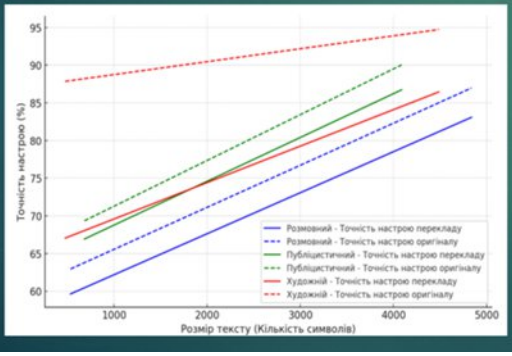
№	Точність настрою перекладу (%)	Точність настрою оригіналу (%)	Точність перекладу (%)	Розмір тексту (Кількість символів)	Стиль мовлення
1	59,64	62,95	80,84	532	Розмовний
2	83,08	86,97	86,06	4837	Розмовний
3	66,93	69,35	82,584	682	Публіцистичний
4	86,73	90,04	83,6	4087	Публіцистичний
5	67,05	87,85	67,68	475	Художній
6	86,45	94,73	70,11	4486	Художній

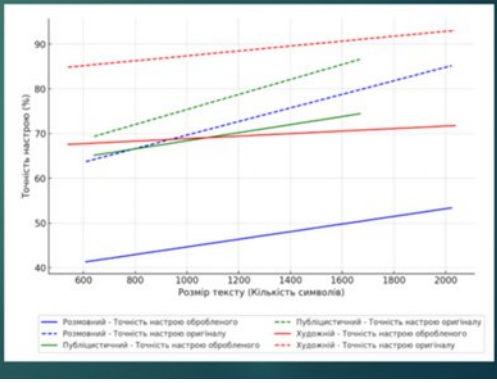
Таблиця 4.3. – Результати проведених експериментів з Amazon Transcribe (таблиця виконана самостійно)

№	Точність настрою обробленого (%)	Точність настрою оригіналу (%)	Точність speech-to-text (%)	Розмір тексту (Кількість символів)	Стиль мовлення
1	41,32	63,74	59.86	609	Розмовний
2	53,39	85,19	56.16	2023	Розмовний
3	65,15	69,37	71.13	641	Публіцистичний
4	74,43	86,61	67.53	1669	Публіцистичний
5	67,58	84,85	65.76	540	Художній
6	71,78	93,07	58.63	2036	Художній

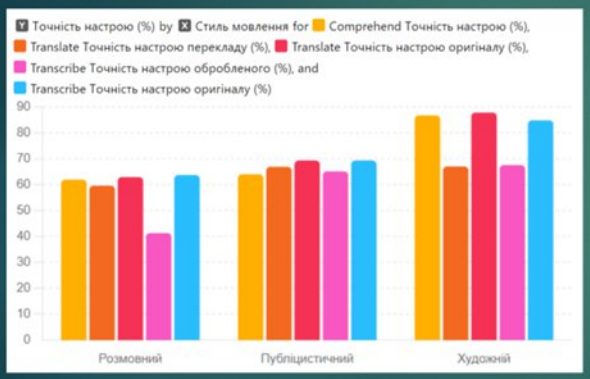
Результати експерименту

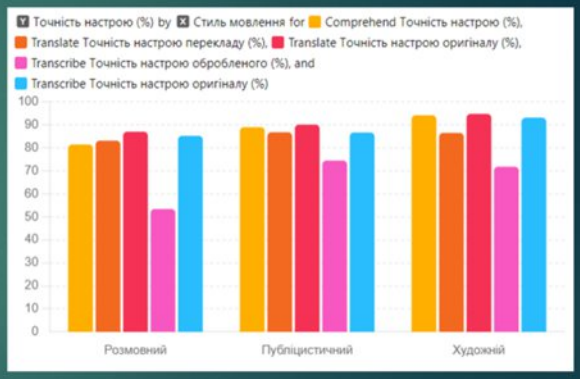






Результати експерименту





Аналіз отриманих результатів

- **Співставлення з цілями дослідження:**
  - Визначення можливостей AWS для NLP.
- **Висновки з отриманих даних:**
  - Вплив стилю мовлення та розміру тексту на точність аналізу.
- **Інтерпретація результатів:**
  - Рекомендації щодо використання сервісів AWS для аналізу відгуків.

Підсумки

- **Реалістичність та корисність отриманих результатів:**
  - Покращення якості аналізу відгуків, автоматизація процесу.
- **Можливий розвиток досліджень:**
  - Подальше вдосконалення системи, інтеграція нових сервісів та методів.

Публікація результатів

УДК 004.75

ДОСЛІДЖЕННЯ СЕРВІСІВ AWS ДЛЯ ОБРОБКИ ПРИРОДНОЇ МОВИ

Дубок В.Ю.
Науковий керівник – к.т.н., доц. Кравець Н.С.
Харківський національний університет радіоелектроніки, каф. ІП,
м. Харків, Україна
e-mail: vladyslav.dubok@nure.ua

This work explores the current state and prospects of natural language processing (NLP). The main objectives of NLP, such as understanding, generation, and processing of text, have been described. The text scrutinizes the challenges inherent in grappling with the inherent ambiguity of human language, as well as the methodologies adopted to mitigate such complexities. The purpose of this work is to compare AWS services and identify the best one for processing user feedback. AWS has been discussed as a cloud platform offering a wide range of NLP services, such as Amazon Lex, Amazon Polly, Amazon Transcribe, Amazon Translate, and Amazon Comprehend.

Сучасне суспільство стрімко розвивається у галузі інформаційних технологій та штучного інтелекту. Завдяки швидкому зростанню обчислювальної потужності, доступності великих обсягів даних і новаторським методам в галузі алгоритмів, технології розвиваються надзвичайно швидко, впливаючи на всі сфери життя. Однією з ключових сфер, яка відіграє важливу роль у цьому процесі, є обробка природної мови.

Обробка природної мови або NLP (Natural Language Processing) це напрям лінгвістики інформаційних технологій та штучного інтелекту, спрямований на вивчення проблем комп'ютерного аналізу та синтезу природної мови. Основними цілями NLP є: розуміння — створення алгоритмів для розпізнавання речень, фраз, слів та граматичної структури, а також зрозуміти значення тексту, генерація — можливість створення текстів які будуть граматично правильними та зрозумілими для людей, обробка — виконання різних задач з текстом таких як: переклад, пошук.

Розуміння природної мови іноді вважають AI-повною задачею, тому що розпізнавання живої мови потребує величезних знань системи про навколишнє середовище та можливості взаємодіяти з ним. Проблема розуміння людського мовлення полягає саме у його неоднозначності. Одні і ті ж самі фрази та слова можуть мати різні значення в залежності від контексту речення. Під час генерації тексту важливо передати не тільки сенс, а й врахувати особливості природної мови. Вісім вище перелічених вище є особливо відчутними при створенні систем машинного перекладу.

Для вирішення задач NLP використовується широкий спектр методів, таких як:

Статистичні методи: грунтуються на аналізі статистичних закономірностей мови.

Методи машинного навчання: використовують алгоритми машинного навчання для вивчення мовних закономірностей.

Лінгвістичні методи: грунтуються на знаннях про лінгвістичну структуру мови.

Наразі обробка природної мови одна з найпоширеніших тем яку зараз досліджують та розвивають протягом останніх років. Можливість аналізувати, обчислювати та розуміти просту людську мову за допомогою комп'ютера і машинного навчання була впроваджена в різні програми та платформи. Таким чином зараз стало можливо автоматизувати деякі процеси які є більш ефективнішими ніж використання людей. Одним з таких процесів є обробка відгуків користувачів.

В останні десятиліття зростає значення відгуків користувачів як важливого джерела інформації для прийняття рішень. Аналіз цих відгуків став необхідним завданням для компаній, які прагнуть зберегти конкурентоспроможність і задовольнити потреби своїх клієнтів. Однак, разом зі зростанням обсягу відгуків, ускладнюється і складності. Проблема полягає в тому, щоб ефективно аналізувати велику кількість даних, що надходять у різних форматах та від різних джерел. При цьому необхідно враховувати не лише текстовий зміст відгуків, але і їх контекст, тон, емоційне забарвлення та інші аспекти, що можуть впливати на їх значущість. В даній роботі розглядається вирішення цих проблем за допомогою сервісів AWS.

Метою роботи є порівняння сервісів AWS з визначенням найкращого з них для обробки відгуків користувачів.

AWS пропонує широкий спектр сервісів для NLP, які дозволяють розробникам швидко та легко створювати інтелектуальні додатки.

Amazon Lex — платформа для розробки чат-ботів та віртуальних помічників, які розуміють та генерують людську мову. Вони можуть вести діалог з користувачами, відповідати на питання, надавати інформацію та виконувати завдання. Завдяки таким широким можливостям їх можна застосувати у створенні системи особистого віртуального помічника.

Amazon Polly — сервіс для перетворення тексту в мову з реалістичним звучанням з можливістю налаштування тему, гучності, інтонації і підтримкою десятків мов та голосів.

Amazon Transcribe — сервіс для перетворення мови в текст. Забезпечує високу точність розпізнавання мови навіть при наявності фонових шумів або присутності акцентів.

Amazon Comprehend — сервіс для аналізу тексту, який допомагає отримувати цінні знання текстових даних. Він використовує машинне навчання для автоматичного аналізу тексту та виявлення ключових інформаційних блоків. У результаті аналізу тексту, отримуємо список сутностей, які поділені та розпізнані по категоріям (імена людей, місця, організації, дати, тощо). Також є можливість отримати настрій

проаналізованого тексту, і сутностей у ньому. Наразі не всі мови підтримуються для аналізу, але є можливість виявити якою мовою написаний текст.

Дані сервіси будуть порівняні за наступними критеріями: мова — яку кількість мов підтримує сервіс для обробки тексту, відповідність функціональності — чи достатньо функцій сервісу для вирішення питання обробки відгуків користувачів, моделі — чи підтримує сервіс можливість додавати власні моделі для обробки запитів і які є обмеження.

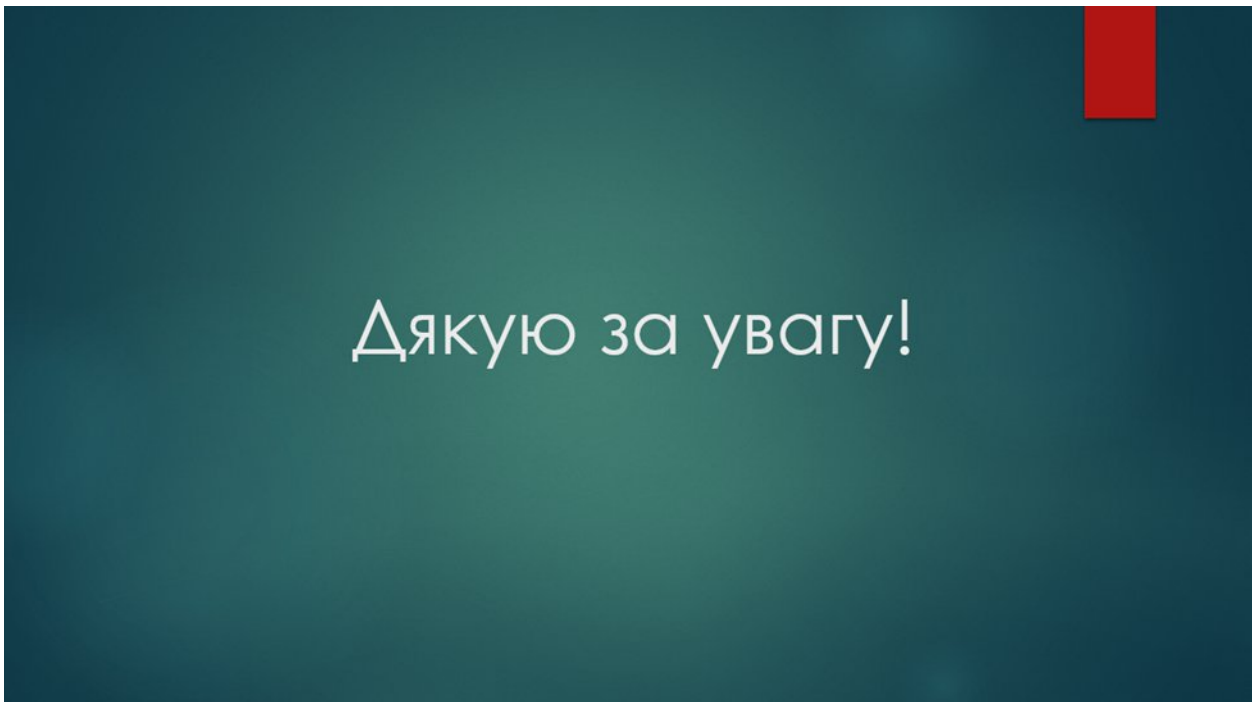
Для тестування даних сервісів буде створено веб додаток, за допомогою якого можна буде перевірити чи придатний сервіс AWS для обробки відгуків користувачів. Така схема допоможе дізнатися обмеження кожного сервісу та можливості комбінувати функціонал між собою, вирішуючи недоліки один одного.

За результатами порівняння сервісів за критеріями очікуються отримання комплексної оцінки за допомогою якої можна буде вирішити які сервіси підходять або не підходять для вирішення проблеми обробки відгуків користувачів.

Отже, обробка природної мови – це область яка динамічно розвивається і має широкий спектр застосувань. Сервіси AWS для NLP дають можливість розробникам створювати інтелектуальні додатки, які розуміють, генерують та обробляють текст, подібно до того, як це роблять люди. В рамках даної роботи було розглянуто можливість використання сервісів AWS для обробки тексту відгуків користувачів. За результатами аналізу текстів з допомогою Amazon Comprehend отримаємо необхідні дані для: формування загальної оцінку продукту, автоматичної класифікації відгуків на позитивні та негативні та виявлення відношення користувачів до сутностей продукту для типізації відгуків, що підвищить розуміння сильних та слабких сторін продукту.

Список використаних джерел

1. What is Natural Language Processing NLP? : вебсайт URL: https://aws.amazon.com/ru/what-is/nlp/?nc1=h_ls (дата звернення: 03.03.2024).
2. Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc.
3. Vytrodov K., Chupryna A., Kotelnykov R. Detecting of Anti-Ukrainian Trolling Tweets //Proceedings http://ceur-ws. org ISSN. – 2023. – T. 1613. – C. 0073.


Дякую за увагу!

ДОДАТОК Г

Апробація результатів роботи

УДК 004.75

ДОСЛІДЖЕННЯ СЕРВІСІВ AWS ДЛЯ ОБРОБКИ ПРИРОДНОЇ МОВИ

Дубок В.Ю.
Науковий керівник – к.т.н., доц. Кравець Н.С.
Харківський національний університет радіоелектроніки, каф. ПІ,
м. Харків, Україна
e-mail: vladyslav.dubok@nure.ua

This work explores the current state and prospects of natural language processing (NLP). The main objectives of NLP, such as understanding, generation, and processing of text, have been described. The text scrutinizes the challenges inherent in grappling with the inherent ambiguity of human language, as well as the methodologies adopted to mitigate such complexities. The purpose of this work is to compare AWS services and identify the best one for processing user feedback. AWS has been discussed as a cloud platform offering a wide range of NLP services, such as Amazon Lex, Amazon Polly, Amazon Transcribe, Amazon Translate, and Amazon Comprehend.

Сучасне суспільство стрімко розвивається у галузі інформаційних технологій та штучного інтелекту. Завдяки швидкому зростанню обчислювальної потужності, доступності великих обсягів даних і новаторським методам в галузі алгоритмів, технології розвиваються надзвичайно швидко, впливаючи на всі сфери життя. Однією з ключових сфер, яка відіграє важливу роль у цьому процесі, є обробка природної мови.

Обробка природної мови або NLP (Natural Language Processing) це напрям лінгвістики інформаційних технологій та штучного інтелекту, спрямований на вивчення проблем комп'ютерного аналізу та синтезу природної мови. Основними цілями NLP є: розуміння — створення алгоритмів для розпізнавання речень, фраз, слів та граматичні структури, а також зрозуміти значення тексту, генерація — можливість створення текстів які будуть граматично правильними та зрозумілими для людей, обробка — виконання різних задач з текстом таких як: переклад, пошук.

Розуміння природної мови іноді вважають AI-повною задачею, тому що розпізнавання живої мови потребує величезних знань системи про навколишнє середовище та можливості взаємодіяти з ним. Проблема розуміння людського мовлення полягає саме у його неоднозначності. Одні і ті ж самі фрази та слова можуть мати різні значення в залежності від контексту речення. Під час генерації тексту важливо передати не тільки сенс, а й врахувати особливості природної мови. Вплив вище перелічених явищ є особливо відчутним при створенні систем машинного перекладу.

Для вирішення задач NLP використовується широкий спектр методів, таких як:

Статистичні методи: ґрунтуються на аналізі статистичних закономірностей мови.

Рисунок – Перша сторінка публікації

Методи машинного навчання: використовують алгоритми машинного навчання для вивчення мовних закономірностей.

Лінгвістичні методи: ґрунтуються на знаннях про лінгвістичну структуру мови.

Наразі обробка природної мови одна з найпоширеніших тем яку зараз досліджують та розвивають протягом останніх років. Можливість аналізувати, обчислювати та розуміти просту людську мову за допомогою комп'ютерів і машинного навчання була впроваджена в різні програми та платформи. Таким чином зараз стало можливо автоматизувати деякі процеси які є більш ефективнішими ніж використання людей. Одним з таких процесів є обробка відгуків користувачів.

В останні десятиліття зростає значення відгуків користувачів як важливого джерела інформації для прийняття рішень. Аналіз цих відгуків став необхідним завданням для компаній, які прагнуть зберегти конкурентоспроможність і задовольнити потреби своїх клієнтів. Однак, разом зі зростанням обсягу відгуків, з'являються і складності. Проблема полягає в тому, щоб ефективно аналізувати велику кількість даних, що надходять у різних форматах та від різних джерел. При цьому необхідно враховувати не лише текстовий зміст відгуків, але і їх контекст, тон, емоційне забарвлення та інші аспекти, що можуть впливати на їх значущість. В даній роботі розглядається вирішення цих проблем за допомогою сервісів AWS.

Метою роботи є порівняння сервісів AWS з визначенням найкращого з них для обробки відгуків користувачів.

AWS пропонує широкий спектр сервісів для NLP, які дозволяють розробникам швидко та легко створювати інтелектуальні додатки.

Amazon Lex — платформа для розробки чат-ботів та віртуальних помічників, які розуміють та генерують людську мову. Вони можуть вести діалог з користувачами, відповідати на питання, надавати інформацію та виконувати завдання. Завдяки таким широким можливостям їх можна застосувати у створенні системи особистого віртуального помічника.

Amazon Polly — сервіс для перетворення тексту в мову з реалістичним звучанням з можливістю налаштування темпу, гучності, інтонації і підтримкою десятків мов та голосів.

Amazon Transcribe — сервіс для перетворення мови в текст. Забезпечує високу точність розпізнавання мови навіть при наявності фонових шумів або присутності акценту.

Amazon Comprehend — сервіс для аналізу тексту, який допомагає отримувати цінні знання текстових даних. Він використовує машинне навчання для автоматичного аналізу тексту та виявлення ключових інформаційних блоків. У результаті аналізу тексту, отримуємо список сутностей, які поділені та розпізнані по категоріям (імена людей, місця, організації, дати, тощо). Також є можливість отримати настрій

Рисунок – Друга сторінка публікації

проаналізованого тексту, і сутностей у ньому. Наразі не всі мови підтримуються для аналізу, але є можливість виявити якою мовою написаний текст.

Дані сервіси будуть порівнянні за наступними критеріями: мова — яку кількість мов підтримує сервіс для обробки тексту, відповідність функціональності — чи достатньо функцій сервісу для вирішення питання обробки відгуків користувачів, моделі — чи підтримує сервіс можливість додавати власні моделі для обробки запитів і які є обмеження.

Для тестування даних сервісів буде створено веб додаток, за допомогою якого можна буде перевірити чи придатний сервіс AWS для обробки відгуків користувачів. Така схема допоможе дізнатися обмеження кожного сервісу та можливості комбінувати функціонал між собою, вирішуючи недоліки один одного.

За результатами порівняння сервісів за критеріями очікується отримання комплексної оцінки за допомогою якої можна буде вирішити які сервіси підходять або не підходять для вирішення проблеми обробки відгуків користувачів.

Отже, обробка природної мови – це область яка динамічно розвивається і має широкий спектр застосувань. Сервіси AWS для NLP дають можливість розробникам створювати інтелектуальні додатки, які розуміють, генерують та обробляють текст, подібно до того, як це роблять люди. В рамках даної роботи було розглянуто можливість використання сервісів AWS для обробки тексту відгуків користувачів. За результатами аналізу текстів з допомогою Amazon Comprehend отримаємо необхідні дані для: формування загальної оцінку продукту, автоматичної класифікації відгуків на позитивні та негативні та виявлення відношення користувачів до сутностей продукту для типізації відгуків, що підвищить розуміння сильних та слабких сторін продукту.

Список використаних джерел

1. What is Natural Language Processing NLP? : вебсайт URL: https://aws.amazon.com/what-is/nlp/?nc1=h_ls (дата звернення: 03.03.2024).

2. Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc.

3. Vyrodov K., Chupryna A., Kotelnykov R. Detecting of Anti-Ukrainian Trolling Tweets //Proceedings http://ceur-ws. org ISSN. – 2023. – T. 1613. – C. 0073.

Рисунок – Третя сторінка публікації

ДОДАТОК Д

Експертний висновок результатів перевірки кваліфікаційної роботи на відповідність оформлення вимогам ДСТУ 3008: 2015

Експертний висновок результатів перевірки кваліфікаційної роботи

студент (посада) | програмної інженерії (кафедра) | ІПЗм-21-1 (група)

Дубок Владислав Юрійович

(прізвище, ім'я, по батькові)

Зауваження

Пункт ДСТУ 3008-2015	Зміст пункту	Сторінка кваліфікаційної роботи
1	2	3
	7.1 Загальні положення	
	7.3 Нумерація сторінок звіту	
	7.4 Нумерація розділів, підрозділів, пунктів, підпунктів	
	7.5 Рисунки	
	7.6 Таблиці	
	7.7 Переліки	
	7.8 Примітки	
	7.9 Виноски	
	7.10 Формули та рівняння	
	7.11 Посилання	
	7.13 Список авторів	
	7.14 Скорочення та умовні познаки	
	7.15 Додатки	

зауважень немає

Експерт _____ (підпис) Олена ОЛІЙНИК (прізвище, ініціали)

11.06.2024