

ДОДАТОК А

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

8. Intelligent information system of heterogeneous medical data analysis / A. Yerokhin et al. 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, 5–8 September 2017. 2017. URL: <https://doi.org/10.1109/stc-csit.2017.8098798> (дата звернення: 30.06.2024).

10. A new intelligence-based approach for rhinomanometric data processing / A. Yerokhin et al. 2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 19–21 April 2016. 2016. URL: <https://doi.org/10.1109/elnano.2016.7493047> (дата звернення: 01.05.2024).

11. Usage of phase space diagram to finding significant features of rhinomanometric signals / A. Yerokhin et al. 2016 XIth International Scientific and Technical Conference “Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 6–10 September 2016. 2016. URL: <https://doi.org/10.1109/stc-csit.2016.7589871> (дата звернення: 03.05.2024).

ДОДАТОК Б

Слайди презентації

Дослідження методів визначення актуальності та релевантності запитань до комерційних пропозицій

Годованець Д.С., ПЗМ-22-1

Науковий керівник: доц. Бабій А.С.

Дослідження

- Наразі, дослідження в галузі визначення актуальності та релевантності запитань до комерційних пропозицій стають все більш актуальними через потребу в точній оцінці та аналізі запитів, що підвищує якість та ефективність комерційних взаємодій.
- Напрямок дослідження полягає у вивченні моделей та методів для автоматизованої оцінки актуальності та релевантності запитань до комерційних пропозицій. Це дослідження має на меті розробку, аналіз та впровадження передових методів автоматизації оцінки запитань з використанням сучасних технологій обробки даних та машинного навчання.
- Об'єктом дослідження є моделі та методи для автоматизованої оцінки актуальності та релевантності запитань до комерційних пропозицій.

Постановка задачі

Робота спрямована на виявлення найкращого методу або алгоритму, який забезпечить точне і швидке визначення актуальності та релевантності запитань до комерційних пропозицій, що дозволить компаніям оптимізувати свої бізнес-процеси та підвищити конкурентоспроможність.

Реалізація наукового дослідження складається з наступних етапів:

- аналіз предметної області комерційних пропозицій та проблем, пов'язаних з їх аналізом;
- розгляд основних характеристик та функціональності алгоритмів ML.NET та TensorFlow;
- проведення порівняльного аналізу підходів на основі ML.NET та TensorFlow.
- розробка алгоритму для автоматизації процесу аналізу комерційних пропозицій.
- реалізація розробленого алгоритму у бізнес-середовищі та провести його тестування.
- оцінка результатів впровадження з точки зору підвищення ефективності бізнес-процесів та задоволеності клієнтів.



Методологія

Опис використаних методів дослідження: Дослідження проводилось за допомогою аналізу предметної галузі, що включало вивчення існуючих методів визначення актуальності та релевантності запитань до комерційних пропозицій. Ми аналізували наукові статті, патерни та результати досліджень у цій галузі. Використовувались експериментальні методи для проектування та тестування алгоритмів, а також порівняльний аналіз для оцінки ефективності та адаптованості різних моделей машинного навчання. Наші дослідження включали експериментування з алгоритмами ранжування інформації та нейронними мережами для досягнення найкращих результатів.

Інструментарій та технології, використані в роботі:

- Платформи: ML.NET, TensorFlow
- Мови програмування: C#
- Фреймворки для машинного навчання: ML.NET, TensorFlow
- Методи обробки даних: Алгоритми машинного та глибинного навчання, нейронні мережі



Дослідження проблематики

Актуальність та важливість дослідження:

- В сучасних умовах конкуренції, ефективність комерційних пропозицій є критичним фактором для успішності бізнесу.
- Збільшення обсягів інформації та різноманітність запитань ускладнюють процес оцінки актуальності та релевантності комерційних запитів.

Основні проблеми:

- Відсутність автоматизованих систем для точної оцінки релевантності та актуальності комерційних запитів.
- Труднощі в адаптації існуючих методів машинного навчання до специфічних вимог бізнесу.
- Необхідність інтеграції нових технологій в існуючі системи управління для покращення ефективності процесу прийняття рішень.



5

Огляд існуючих методів

Машинне навчання (МН) – це підгалузь штучного інтелекту, яка фокусується на розробці алгоритмів та моделей, здатних навчатися на основі даних і приймати рішення або робити прогнози без прямого програмування. Системи на основі МН автоматично вдосконалюються з досвідом.

Типи навчання:

- Навчання з учителем: Модель навчається на наборі даних, що містить вхідні приклади та відповідні мітки, з метою робити правильні прогнози для нових даних (класифікація та регресія).
- Навчання без учителя: Модель виявляє приховані структури або закономірності у даних без міток (кластеризація та асоціація).
- Навчання з підкріпленням: Модель навчається на основі винагород і покарань, намагаючись максимізувати загальну винагороду.

Застосування: Використовується для класифікації текстів, прогнозування, аналізу даних, розпізнавання образів та багатьох інших завдань.



6

Огляд існуючих методів

Штучні нейронні мережі (ШНМ) – це моделі обчислювальної техніки, натхненні біологічними нейронними мережами людського мозку. Вони складаються з вузлів (нейронів), об'єднаних у шари, які працюють разом для обробки даних.

Типи ШНМ:

- Перцептрон: Найпростіший тип ШНМ, який використовується для задач класифікації.
- Багатошаровий перцептрон (MLP): Містить один або більше прихованих шарів, здатних моделювати складні функції.
- Рекурентні нейронні мережі (RNN): Використовуються для обробки послідовних даних, таких як текст або часові ряди.
- Конволюційні нейронні мережі (CNN): Застосовуються в основному для обробки зображень та відео.

Застосування: Використовується для класифікації текстів, прогнозування, аналізу даних, розпізнавання образів та багатьох інших завдань.



Вибір методу аналізу

Машинне навчання було обрано для аналізу комерційних пропозицій з кількох вагомих причин, які виділяють його серед інших методів. Нижче наведено основні причини, чому машинне навчання є оптимальним вибором для цього завдання:

Гнучкість: Машинне навчання дозволяє створювати моделі, які можуть адаптуватися до різних типів даних і задач. Це особливо важливо в контексті комерційних пропозицій, де дані можуть бути різномірними та змінюватися з часом.

Автоматизація: Машинне навчання дозволяє автоматизувати процес аналізу комерційних пропозицій, зменшуючи потребу в ручній обробці даних. Це значно підвищує ефективність та зменшує витрати часу і ресурсів.

Висока точність: Автоматизовані системи машинного навчання можуть швидко і точно аналізувати великі обсяги даних, виявляти релевантні пропозиції та надавати рекомендації, що забезпечує конкурентні переваги для бізнесу.



Порівняння бібліотек машинного навчання

ML.NET:

Переваги:

- Підтримує широкий спектр завдань машинного навчання, таких як класифікація, регресія, кластеризація, рекомендаційні системи тощо.
- Універсальний інструмент для різних видів аналізу комерційних пропозицій.
- Можливість інтеграції з ONNX для використання передтренованих моделей BERT.

Недоліки:

- Не має рідної підтримки для трансформерів, таких як BERT.

TensorFlow:

Переваги:

- Одна з найпотужніших платформ для глибокого навчання.
- Підтримує складні архітектури нейронних мереж, такі як трансформери.
- Рідна підтримка для трансформерів, включаючи BERT, що дозволяє легко імплементувати передтреновані моделі.

Недоліки:

- Високі вимоги до обчислювальних ресурсів та часу на навчання.
- Складність реалізації, що вимагає глибоких знань і досвіду.



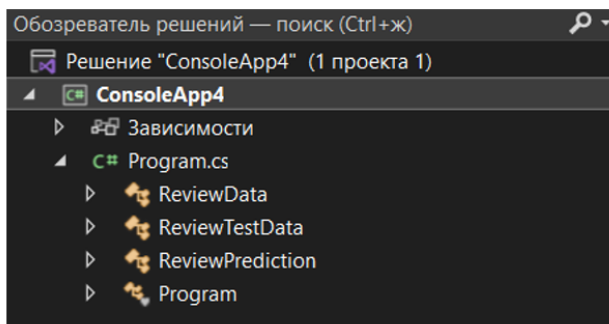
9

Експериментальне дослідження

Для порівняння ми обрали ML.NET та TensorFlow з імплементацією BERT у кожен підхід з кількох вагомих причин.

Ці платформи мають різні сильні сторони та особливості, що робить їх відповідними для різних сценаріїв застосування в аналізі комерційних пропозицій.

Рішення має просту тестову файлова структуру:



- ConsoleApp (Кореневий проект);
- Dependencies (Залежності);
- Program.cs (Головний файл);
- ReviewData (Клас даних відгуків);
- ReviewTestData (Клас тестових даних відгуків);
- ReviewPrediction (Клас передбачення);
- Program (Головний клас з методами для тренування та тестування моделі).



Тестова архітектура додатку

10

Набір даних та тренування

```
CommercialDataSet – Блокнот
Файл Правка Формат Вид Справка
"The delivery was super fast and the package was in perfect condition.", "Delivery"
"I'm disappointed with the 'product quality' for the 'price' paid.", "Quality"
"The setup instructions for the 'laptop' were clear and easy to follow.", "Product Information"
"Customer service was unhelpful and rude when I called.", "Customer Service"
"I love how easy it is to navigate the 'website' and make purchases.", "Website"
"The 'laptop battery life' is shorter than advertised.", "Product Complaints"
"Very pleased with the professional 'installation service' provided.", "Service"
"The 'food' arrived cold and two hours late.", "Delivery"
"Great value for the 'price', will definitely shop here again!", "Price"
"The 'shoes' fell apart after a few wears, not worth the money.", "Quality"
"Amazing 'customer support', they solved my issue within minutes.", "Customer Service"
"The 'software subscription' is too expensive for the features offered.", "Price"
"Extremely comfortable 'chair', I can sit for hours without discomfort.", "Product Praise"
"The return process was complicated and time-consuming.", "Customer Service"
"Highly recommend this 'coffee maker', it makes amazing coffee every time.", "Product Praise"
"The new 'gaming mouse' glides smoothly and has customizable buttons.", "Product Praise"
"The 'flower delivery' service was prompt, but the bouquet was smaller than expected.", "Delivery"
"The 'winter jacket' keeps me warm but the zipper breaks easily.", "Product Complaints"
"Enjoyed the 'documentary' on streaming, very insightful and well-produced.", "Content Quality"
"The 'smartwatch' syncs well with my phone but battery life is below expectations.", "Product Quality"
"The 'bookstore' has a wonderful selection but the staff can be unattentive.", "Customer Service"
"Customer support for the 'software' was knowledgeable and solved my issue quickly.", "Customer Service"
"The 'chocolate cake' from the local bakery was divine and rich in flavor.", "Product Praise"
"The 'scented candle' has a soothing aroma but burns out too fast.", "Product Quality"
"The 'yoga mat' is durable and non-slip, perfect for my daily sessions.", "Product Praise"
"The 'online course' for digital marketing was informative but overly expensive.", "Service"
"Airline service was mediocre, with delayed flights and unapologetic staff.", "Service"
"The 'portrait lens' produces sharp images, though it's quite heavy.", "Product Quality"
"The 'historical novel' was captivating, filled with intrigue and accurate details.", "Content Quality"
"Home insurance process was complicated and the terms were not clear.", "Service"
"The 'gym membership' offers many classes but it's always too crowded.", "Service"
"The 'Italian restaurant' near downtown serves authentic dishes with fresh ingredients.", "Product Praise"
"The 'vacuum cleaner' is powerful but its loud noise is a major downside.", "Product Complaints"
```

Тренувальний набір даних

На рисунках можна побачити датасети використані для навчання та тестування, щоб забезпечити належну якість і релевантність даних для нашої моделі машинного навчання

```
TestDataSet – Блокнот
Файл Правка Формат Вид Справка
"The delivery was super fast and the package was in perfect condition."
"I'm disappointed with the product quality for the price paid."
"The setup instructions for the laptop were clear and easy to follow."
"Customer service was unhelpful and rude when I called."
"I love how easy it is to navigate the website and make purchases."
"The laptop battery life is shorter than advertised."
"Very pleased with the professional installation service provided."
"The food arrived cold and two hours late."
"Great value for the price, will definitely shop here again!"
```

Тестовий набір даних

11

Аналіз результатів

Результати роботи двох моделей, представлені на зображеннях, показують порівняння продуктивності моделей ML.NET та TensorFlow у завданні класифікації текстових відгуків.

```
Консоль отладки Microsoft Visual Studio
2: Test Model
3: Exit
Select an option: 2
Testing model, please wait...
Text: The delivery was super fast and the package was in perfect condition.
Predicted Category: Delivery
Commercial Object: package

Text: I'm disappointed with the product quality for the price paid.
Predicted category: Quality
Commercial Object: object

Text: The setup instructions for the laptop were clear and easy to follow.
Predicted Category: Product Information
Commercial Object: laptop

Text: Customer service was unhelpful and rude when I called.
Predicted Category: Customer Service
Commercial Object: object

Text: I love how easy it is to navigate the website and make purchases.
Predicted Category: Website
Commercial Object: website

Model accuracy: 86%

E:\Users\lbogac\Desktop\ConsoleApp4\bin\Debug\net6.0\ConsoleApp4.exe (процесс 12128)
Чтобы автоматически закрывать консоль при остановке отладки, включите параметр "Ce
Чтобы автоматически закрывать консоль при остановке отладки".
```

Результати ML.NET

```
Консоль отладки Microsoft Visual Studio
2: Test Model
3: Exit
Select an option: 2
Testing model, please wait...
Text: The delivery was super fast and the package was in perfect condition.
Predicted Category: Quality
Commercial Object: object

Text: I'm disappointed with the product quality for the price paid.
Predicted Category: Quality
Commercial Object: object

Text: The setup instructions for the laptop were clear and easy to follow.
Predicted Category: Product Information
Commercial Object: laptop

Text: Customer service was unhelpful and rude when I called.
Predicted Category: Customer Service
Commercial Object: object

Text: I love how easy it is to navigate the website and make purchases.
Predicted Category: Website
Commercial Object: website

Model accuracy: 78%

E:\Users\lbogac\Desktop\ConsoleApp4\bin\Debug\net6.0\ConsoleApp4.exe (процесс 12128)
Чтобы автоматически закрывать консоль при остановке отладки, включите параметр "Ce
Чтобы автоматически закрывать консоль при остановке отладки".
```

Результати TensorFlow

На першому зображенні ми бачимо результати моделі ML.NET, яка передбачає категорії відгуків і визначає комерційні об'єкти. Модель ML.NET показала точність у 86%, що свідчить про її здатність правильно класифікувати текстові відгуки у більшості випадків. Комерційні об'єкти, такі як "package", "laptop", і "website", також були ідентифіковані для деяких відгуків.

На другому зображенні представлені результати моделі TensorFlow, яка виконувала аналогічне завдання класифікації. Точність цієї моделі склала 70%, що є нижчим показником порівняно з моделлю ML.NET. Хоча і TensorFlow вдалося правильно передбачити деякі категорії відгуків, кількість правильних передбачень була меншою. Комерційні об'єкти, такі як "laptop" і "website", також були розпізнані, але частіше модель визначала об'єкти як "object", що свідчить про складність розпізнавання конкретних об'єктів у тексті для цієї моделі.

12

Висновки

- Кваліфікаційна робота включала дослідження аналізу комерційних пропозицій і впровадження моделей машинного навчання для класифікації текстових відгуків.
- Було створено та протестовано моделі на основі ML.NET і TensorFlow. Модель ML.NET показала вищу точність (86%) у порівнянні з моделлю TensorFlow (70%).
- Робота продемонструвала, що ML.NET є більш адаптованою для класифікації текстів, хоча TensorFlow також може бути використаний з додатковою оптимізацією.
- Дослідження підкреслює важливість правильного вибору інструментів та методів у системах машинного навчання.

ДОДАТОК В

Результат проходження на академічний плагіат



Ім'я користувача:
Кардаш Євген Вікторович каф.ПІ

ID перевірки:
1016386144

Дата перевірки:
24.06.2024 18:58:02 EEST

Тип перевірки:
Doc vs Internet + Library

Дата звіту:
24.06.2024 19:00:10 EEST

ID користувача:
100013622

Назва документа: 2024_М_ПІ_ІПЗм_22_1_Годованец_Д_С_скорочений

Кількість сторінок: 40 Кількість слів: 6991 Кількість символів: 55040 Розмір файлу: 649.68 KB ID файлу: 1016197315

6.51% Схожість

Найбільша схожість: 3.98% з джерелом з Бібліотеки (ID файлу: 1016183606)

1.63% Джерела з Інтернету 90 Сторінка 42

5.69% Джерела з Бібліотеки 30 Сторінка 42

0% Цитат

Вилучення цитат вимкнено

Вилучення списку бібліографічних посилань вимкнено

0% Вилучень

Немає вилучених джерел

ДОДАТОК Г

Апробація результатів роботи



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-22-1
(група)

Годованець Денис Сергійович

(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 позначки	
	7.15 Додатки	

зауважень немає

Експерт

(підпис)

Олена ОЛІЙНИК

(прізвище, ініціали)

24.06.2024