

Метод автоматизованого оцінювання відповідності тональності відгуків на товари в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання

Шифр: тональність відгуків

Анотація

Актуальність. Дослідженнями встановлено, що більшість споживачів довіряють відгукам так само, як рекомендаціям від друзів чи родичів, відповідно, відгуки є одним із способів оцінки якості товару перед покупкою. Отже, виявлення відповідності тональності відгуків на товари до їх користувацької оцінки є актуальним напрямом досліджень в галузі обробки природної мови. Робота присвячена розробці методу автоматизованого оцінювання відповідності тональності відгуків на товари повсякденною українською мовою в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання. Для визначення тональності використано модифікацію архітектури нейронної мережі BERT – RoBERTa, яка показала кращі результати в задачах класифікації коротких текстових повідомлень.

Мета – розробка та реалізація методу автоматизованого оцінювання відповідності тональності відгуків на товари в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання.

Завдання – автоматизоване визначення рівня емоційного забарвлення тексту та його подальший автоматизований аналіз шляхом співставлення з користувацькою оцінкою задля визначення рівня їх відповідності.

Використана методика дослідження. Налаштування та навчання нейромережевої мовної моделі RoBERTa була ключовим аспектом у цій роботі, адже для української мови такого роду дослідження не проводились.

Наукова робота містить:

Пояснювальна записка			
Сторінок	Рисунків	Таблиць	Джерел інформації
30	19	4	28

Ключові слова: обробка природної мови, семантичний аналіз, оцінка відповідності, емоційне забарвлення.

Зміст

Вступ.....	2
Розділ 1 Аналіз існуючих моделей, методів та реалізацій.....	3
1.1 Аналіз предметної області	3
1.2 Огляд теоретичних підходів до розв’язку подібних задач	5
1.3 Аналіз існуючих програмних рішень.....	6
1.4 Мета та задачі дослідження	8
Розділ 2 Метод автоматизованого оцінювання відповідності тональності відгуків до їх користувацької оцінки засобами глибокого навчання	9
2.1 Складові методу автоматизованого оцінювання відповідності тональності відгуків.....	9
2.2 Проектування інформаційної системи.....	14
2.3 Вибір засобів розробки інформаційної системи.....	16
2.4 Формування датасету з текстовими відгуками щодо якості товарів.....	18
Розділ 3 Програмна реалізація інформаційної системи	20
3.1 Архітектура інформаційної системи.....	20
3.2 Особливості реалізації програмних складових системи.....	22
3.3 Результати досліджень	24
Висновки	30
Перелік посилань.....	31

Вступ

Останніми роками сфера аналізу емоційної тональності тексту, що є частиною обробки природної мови, привертає увагу науковців. Цей напрям дає змогу визначити інформаційну цінність тексту та виявити важливі настрої користувачів. Пов'язана така тенденція із масовим зростанням сфер потенційного використання. Аналіз емоційного забарвлення тексту – це сукупність методів вилучення та розпізнавання оцінок користувачів продуктів і моделей, який має різні підходи з використанням алгоритмів машинного навчання для класифікації емоцій, що стоять за текстом.

Один з прикладів використання – дослідження твітів для розуміння сприйняття людьми певних новин, оцінка взаємодії людини з роботом, формування рекомендаційної системи для вибору товарів тощо.

Відгуки користувачів, отримані на безлічі інтернет-магазинів слугують важливим джерелом інформації не лише для покупців, що обирають товар чи послугу, а й для бізнес-аналітиків.

Втім серед відгуків можна знайти масу невідповідних: іноді користувачі залишають такі коментарі, що їх зміст діаметрально протилежний оцінці, що виставляється поруч. Такі відгуки псують статистику, можуть формувати хибне судження про товар чи послугу.

Аналіз емоційного забарвлення відгуків дозволяє виявити тренди та патерни серед споживачів, як-от виявлення частоти позитивних або негативних відгуків, спільні позитивні або негативні оцінки, що допомагає зрозуміти загальну реакцію споживачів на продукт або послугу.

Отже, визначення емоційного контексту відгуків та аналіз співставлення оцінок товарів та послуг допомагає краще розуміти сприйняття клієнтів, виявляти тенденції, впливати на репутацію бренду та постійно вдосконалювати продукт або послугу. Це важлива інформація для розробників, виробників та підприємств, яка сприяє покращенню їх діяльності та відповіді на потреби клієнтів.

Розділ 1 Аналіз існуючих моделей, методів та реалізацій

1.1 Аналіз предметної області

Відгуки в мережі є невід’ємною частиною сучасного цифрового простору. Сучасний споживач має доступ до безмежного обсягу інформації, серед неї є і відгуки, які інші користувачі залишають про продукти, послуги, ресторани, готелі та інше. Відгуки в Інтернеті можуть бути знайдені на різних платформах, таких як вебсайти, соціальні медіа, форуми та спеціалізовані сервіси для відгуків. Дослідження [1] демонструють, що відгуки в Інтернеті мають значний вплив на рішення споживачів. Коли потенційні покупці розглядають товари або послуги, вони часто шукають відгуки в мережі. Позитивні відгуки можуть переконати споживача в покупці, в той час як негативні відгуки можуть спричинити відмову від придбання товару.

Дослідження також показали, що більшість споживачів довіряють відгукам так само, як рекомендаціям від друзів чи родичів. З іншого боку компанії розглядають відгуки в інтернеті як важливий інструмент для вдосконалення якості своїх продуктів та послуг. Вони можуть відстежувати відгуки користувачів та використовувати їх для здійснення необхідних змін та поліпшень. Впровадження змін, враховуючи погляди та скарги споживачів, допомагає компаніям зберігати і покращувати свою репутацію.

Захистити себе від можливих обманів у процесі покупки товарів – це важливий крок для будь-якого покупця. Одним із способів оцінки якості товару перед покупкою є пошук та аналіз відгуків інших користувачів, які вже придбали цей товар в інтернет-магазині. Відгуки можуть містити інформацію про якість товару, його функціональність, доставку та обслуговування. Інші покупці залишають свої враження та рекомендації, які можуть стати важливими при прийнятті рішення щодо покупки. Зазвичай, на сторінці товару в інтернет-магазині можна знайти відгуки користувачів, де вони окрім тексту відгуку встановлюють оцінки за певною шкалою, наприклад, за допомогою зірок, як

показано на рисунку 1.1. Крім того, покупці можуть навіть прикріплювати фото чи відео до своїх відгуків, щоб краще проілюструвати свої враження.

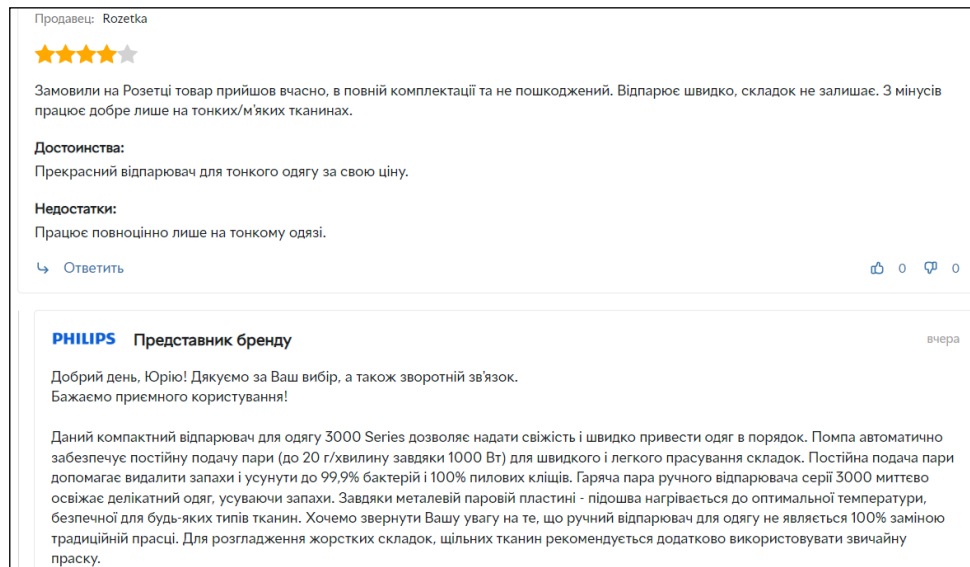


Рисунок 1.1 – Відгук про товар [2]

Оцінка товару на основі відгуків користувачів відображає ступінь задоволення покупців. Проте важливо зберігати обережність і аналізувати кілька різних відгуків, оскільки думки можуть різнитися. Іноді відгуки можуть бути залишені без оцінки клієнтами, і це може статися внаслідок недбалості або невірної натискання під час встановлення оцінки, як показано на рисунку 1.2. Такі ситуації можуть бути непередбаченими та призвести до плутанини серед покупців, які можуть вибирати товари, спираючись на неправильні відгуки, і робити невірні висновки щодо якості товару.



Рисунок 1.2 – Невідповідність відгуку з оцінкою

Не варто забувати, що оцінки та відгуки можуть не завжди точно відображати якість товару через їх суб'єктивність. Деякі інтернет-магазини дозволяють коригувати оцінки та відгуки у випадку помилково залишених відгуків, шляхом звернення до служби підтримки.

Зважаючи на те, що відгуки можуть бути суб'єктивними і не завжди точно відображати реальну якість товару, існує потреба в автоматизації процесу оцінки відповідності тону відгука інтернет-магазину користувацькій оцінці. Це допомагає користувачам отримати більш об'єктивну картину і спрощує процес вибору товарів.

1.2 Огляд теоретичних підходів до розв'язку подібних задач

Штучний інтелект (ШІ) представляє собою комп'ютерні системи та алгоритми, які здатні навчатися та приймати рішення на основі аналізу об'ємних даних. Вивчення ШІ спрямоване на створення комп'ютерних систем, які можуть виконувати інтелектуальні завдання.

Однією з важливих задач в області обробки природної мови є розпізнавання тональності тексту. Цей підхід є дуже популярним в численних країнах. Розпізнавання тональності включає в себе визначення емоційного відтінку тексту, який може бути позитивним, негативним або нейтральним.

Аналіз емоційного відтінку текстової інформації – це активний напрямок досліджень, який досліджується різними підходами, як це відображено у роботах [3, 4]. У науковій роботі [4], яка зосереджена на аналізі мовних настроїв, був представлений фреймворк під назвою "двонапрямковий емоційний рекурентний блок". Цей підхід використовує загальний блок нейронних тензорів, за яким слідує двоканальний класифікатор, призначений для контекстуальної обробки та класифікації настроїв, пов'язаних з текстом.

Автори систематизують ряд наукових публікацій і надають ілюстрації найновіших тенденцій у дослідженнях аналізу настроїв та суміжних галузях [5].

У дослідженні, проведеному авторами [6], було встановлено, що поєднання методів машинного навчання та аналізу настроїв на основі словників дозволяє досягти більш високої точності, у порівнянні з будь-яким іншим методом аналізу настроїв. Дослідники використовували різні підходи до аналізу настроїв, включаючи машинне навчання та аналіз настроїв на основі словників, для проведення тестів та порівняння ефективності в контексті користувацької поведінки. У роботі [7], науковці запропонували метод аналізу настроїв в Twitter, використовуючи словники, і отримали релевантні результати відносно настроїв, пов'язаних з вакцинами проти COVID-19, такими як AstraZeneca/Oxford, Moderna та Pfizer/BioNTech, протягом 4 місяців.

У роботі [8] запропоновано використовувати TextBlob з векторизацією TF-IDF та класифікаційною моделлю LinearSVC для аналізу настроїв. За допомогою цього підходу була досягнута вражаюча точність на рівні 0,96752 для англomовних твітів.

У дослідженні [9] продемонстровано, що сучасні маркетингові дослідження переважно використовують словникові інструменти для вилучення настроїв з текстових даних, і ці інструменти мають певні переваги з точки зору інтерпретації. Однак вони відстають за точністю. Автори також ретельно оцінили існуючі методи аналізу настроїв і продемонстрували, що методи на основі машинного навчання мають вищу точність у класифікації, але можуть бути менш інтерпретованими.

З метою виявлення розбіжностей між тональністю відгуків на товари в інтернет-магазинах і користувацькими оцінками цих товарів, використовуються методи глибокого навчання на основі нейромереж штучного інтелекту.

1.3 Аналіз існуючих програмних рішень

На сьогоднішній день програми для розпізнавання тональності тексту стали досить популярними та застосовуються в різних сферах. Програми для розпізнавання тональності тексту повинні надійно визначати емоційний настрій

тексту, відповідаючи його справжнім відчуттям. Приклади таких програм повинні бути здатні працювати з різними мовами та враховувати їхні особливості. Приклади повинні бути здатні ефективно обробляти великі обсяги тексту. Програми повинні враховувати та розпізнавати різні емоційні відтінки тексту.

Як приклад, розглянуто програму під назвою "AFINN" [10], яка призначена для аналізу та визначення тональності тексту. Ця програма була опублікована вперше в жовтні 2011 року і розроблена Олом Нільсеном. Роботу програми демонструється нижче (рисунок 1.3).

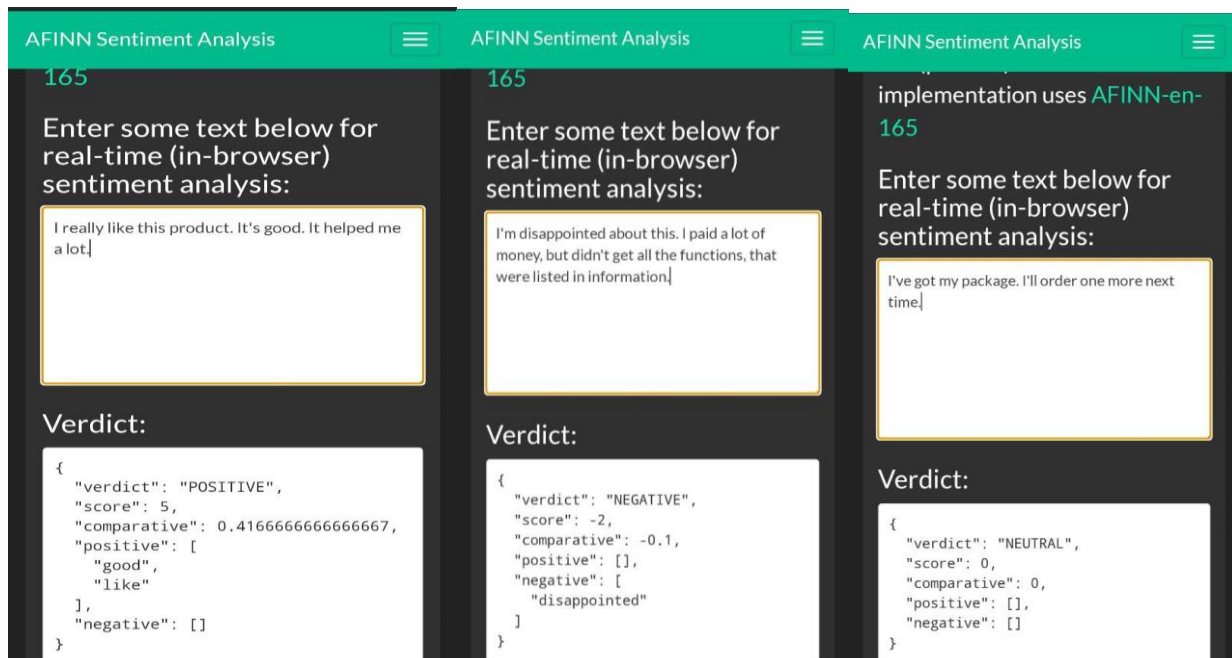


Рисунок 1.3 – Процес роботи програми [10]

Програма "AFINN" є відкритою для вільного використання та базується на словнику, який містить числові оцінки емоційного забарвлення окремих слів у англійській мові. Основна концепція "AFINN" полягає в тому, що кожне слово отримує свою числову оцінку, яка вказує на його емоційне забарвлення. Позитивні слова мають позитивні оцінки, негативні слова - негативні оцінки, а нейтральні слова мають оцінки, близькі до нуля. Програма надає можливість проводити аналіз тональності тексту шляхом сумування оцінок слів, які

зустрічаються в тексті. Чим вище сумарна оцінка, тим більше позитивний настрій тексту, і навпаки. Серед переваг можна зазначити доступність готових словників для англійської мови, а також можливість розширення словника власними словами та їх оцінками, простоту використання, загальну доступність для користувачів.

Підсумовуючи вищесказане, програма "AFINN" представляє собою корисний інструмент для аналізу тональності тексту, зокрема в англійській мові. Вона базується на числових оцінках слів і дозволяє визначати емоційне забарвлення тексту, відповідно до суми цих оцінок. Проте важливо враховувати, що програма має обмежену мовну підтримку та призначена переважно для англійської мови. Також, відсутність регулярних оновлень словника може обмежувати актуальність програми в довгостроковому плані.

Таким чином, відгуки є одним із способів оцінки клієнтами якості товару перед покупкою. Тому виявлення відповідності тональності відгуків на товари до їх користувацької оцінки є актуальним напрямом досліджень в галузі обробки природної мови.

1.4 Мета та задачі дослідження

Метою роботи є розробка та реалізація методу автоматизованого оцінювання відповідності тональності відгуків на товари в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання.

Відповідно, завдання роботи – автоматизоване визначення рівня емоційного забарвлення тексту та його подальший автоматизований аналіз шляхом співставлення з користувацькою оцінкою задля визначення рівня їх відповідності.

Розділ 2 Метод автоматизованого оцінювання відповідності тональності відгуків до їх користувацької оцінки засобами глибокого навчання

2.1 Складові методу автоматизованого оцінювання відповідності тональності відгуків

Враховуючи мету дослідження, оцінка тональності буде проводитися стосовно відгуків про електронну комерцію. Самі відгуки про електронну комерцію мають наступні особливості:

- обмежений обсяг контенту (до 500 слів);
- невеликий обсяг контенту (1-3 слова);
- використання суржику, слів-покручів та професійних слів, жаргонізмів та інтегрованого багатомовного контенту.

У більшості випадків відгуки мають обмежений обсяг і не перевищують 100 слів. Триваліші відгуки, зазвичай, супроводжуються негативною оцінкою.

Крім цього, характерною особливістю відгуків є їхній короткий обсяг. Серед позитивних відгуків часто можна знайти вирази, подібні до таких: *"Рекомендую"*, *"Мені все сподобалося"*, *"Найкращий магазин"*, а серед негативних, відповідно: *"Не рекомендую"*, *"Жах!"* тощо. Зазначено, що відгуки можуть бути недовгими та одночасно містити в собі різноманітні лінгвістичні особливості, такі як:

- Вживання жаргонізмів і сленгу. Відгуки можуть містити специфічну лексику, що є характерною для електронної комерції або специфічних груп споживачів.

- Використання слів, які не відповідають нормам української літературної мови. Це можуть бути іншомовні слова, спотворені слова або запозичення з інших мов.

- Професіоналізми. Відгуки можуть містити терміни та вирази, розуміння яких вимагає фахових знань у галузі електронної комерції.

– Назви товарів. Відгуки можуть містити згадки про конкретні товари або бренди.

Урахування цих мовних особливостей також дозволяє покращити якість аналізу та зробити його більш відповідним конкретній галузі, яку представляють відгуки про електронну комерцію. Це може виявитися корисним для підприємств, які бажають краще розуміти відгуки своїх клієнтів і покращити якість обслуговування. Приклад частини відгуку: *"Мені потрібно було вивести USB 3.0 на передню частину корпусу, тому що в мене є флешки USB 3.0, а підходить до задньої частини комп'ютера і вставляти їх незручно, тому що спереду є тільки USB 2.0. Тому я замовив на Rozetka перехідник Chieftec USB 3.0..."*.

Багатомовний контент також досить поширений у відгуках. Наприклад, ось відгук, який містить помилки та росіянізми: "Я замовив акумулятор у інтернет-магазині, оскільки перевірів, що у них є хороші відгуки." У цьому реченні виявляються орфографічні помилки, зокрема вони виникають через використання слів, що походять із російської мови. Українська версія цього речення має наступний вигляд: "Я замовив акумулятор у магазині в інтернеті, оскільки переконався, що вони мають високі рейтинги та хороші відгуки."

Ураховуючи ці обмеження, виникає необхідність знайти експериментальні дані, які відповідають вищезазначеним критеріям. Схема виявлення відмінностей між відгуками та користувацькою оцінкою представлена на рисунку 2.1.

Початковим етапом методу є визначення бінарної оцінки відгуку, яке включає розділення повідомлень на негативні та позитивні та визначення оцінки відгуку.

Наступним кроком є обрахунок числової оцінки за тональністю за допомогою нейромережі. Для позитивних відгуків числова оцінка обчислюється за допомогою формули 2.1.

$$Level = 5 * NeuralNetworkSentimentMark \quad (2.1)$$

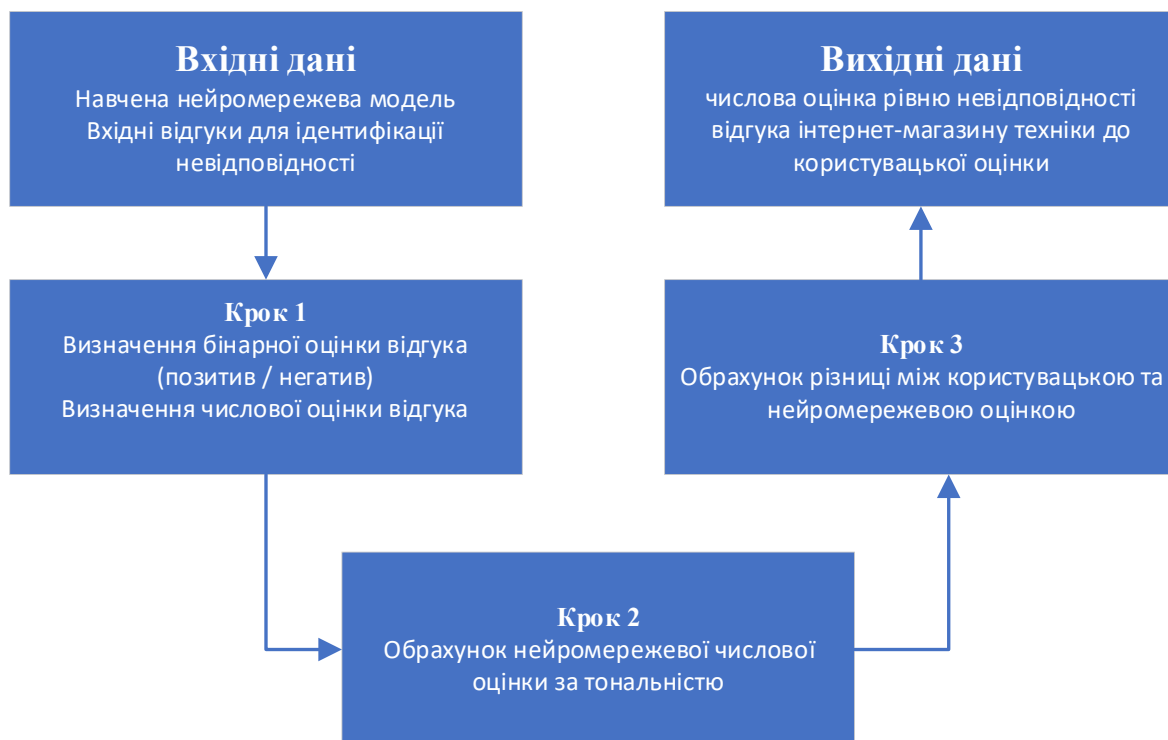


Рисунок 2.1 – Схема виявлення невідповідності відгука до користувацької оцінки

Для негативних відгуків неймережева оцінка визначається наступним чином (формула 2.2):

$$Level = 5 * (1 - NeuralNetworkSentimentMark) \quad (2.2)$$

На третьому етапі проводиться розрахунок різниці між користувацькою та неймережевою оцінкою, що виконується за допомогою формули 2.3.

$$\Delta = |NeuralNetworkSentimentMark - UserMark| \quad (2.3)$$

Якщо метод повертає значення різниці, яке перевищує 1.5 за модулем, то цей відгук вважається несумісним з користувацькою оцінкою. Результати методу включають числову оцінку рівня несумісності між відгуком користувача щодо продукту у вашому інтернет-магазині техніки та його користувацькою оцінкою. Спочатку відбувається вивантаження відгуку користувача щодо придбаного товару із бази даних.

Перший крок методу полягає в установленні бінарної оцінки відгука (позитивного / негативного) та числової оцінки відгука.

Подальший крок включає розрахунок числової оцінки за тональністю з використанням нейромережі, як це показано у прикладі з рисунку 2.2. На завершальному етапі проводиться обрахунок різниці між користувацькою та нейромережевою оцінкою. Наприклад, якщо користувацька оцінка відгуку становить 4, а нейромережева оцінка - 1.748, то модуль різниці складає 2.2519, що перевищує поріг 1.5, і тому відгук вважається несумісним з користувацькою оцінкою.

З використанням глибокої нейромережі програма може визначити ступінь позитивності або негативності відгуку, що допомагає користувачам отримати інформацію щодо емоційного відтінку відгука. Цей аспект є істотним для споживачів, оскільки він полегшує їм краще зрозуміти враження та думки інших покупців про конкретний товар.

Розроблений метод для виявлення невідповідності відгуку в інтернет-магазині техніки до користувацької оцінки, використовуючи глибоку нейромережу, спрямований на визначення некоректно оцінених товарів шляхом аналізу відгуків користувачів. Проводячи аналіз текстових описів відгуків, програма намагається встановити відповідність між відгуками та оцінками, які надали користувачі. Схематично роботу методу проілюстровано на рисунку 2.2.

Ключовим аспектом роботи системи є саме нейромережева мовна модель, що буде визначати оцінку емоційного забарвлення відгуку. Для реалізації системи було обрано модель RoBERTa.

RoBERTa, яке є скороченням від "Robustly optimized BERT approach," є варіацією моделі BERT (Bidirectional Encoder Representations from Transformers), розробленої командою дослідників з Facebook AI [11]. Так само, як BERT, RoBERTa – це мовна модель, яка базується на технології трансформаторів, використовуючи контекст для обробки вхідних текстів та створення контекстуалізованих репрезентацій слів у тексті.

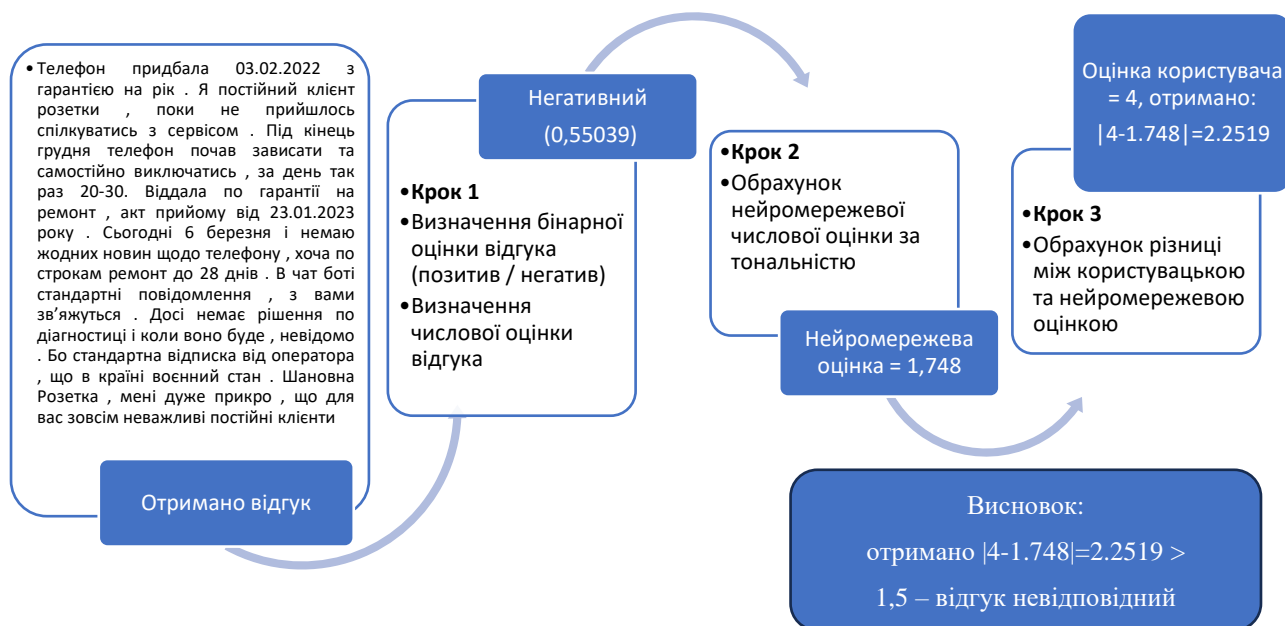


Рисунок 2.2 – Ілюстрація роботи методу із отриманими результатами

Однією з головних відмінностей RoBERTa від BERT є те, що RoBERTa було навчено на великому обсязі даних і використовувалася більш ефективна процедура навчання. Під час навчання RoBERTa використовує техніку динамічного маскування, яка сприяє покращенню надійних та узагальнених репрезентацій слів.

Оскільки використання нейромережевих методів для семантичного аналізу є актуальним напрямком досліджень, існують розробки, спрямовані на українську мову. Один з прикладів - це навчена модель для багатомовної обробки текстів, яка підтримує українську та понад 50 інших мов [12], а також вбудовання [13] Ukjae Jeong, що входить до бібліотеки Tensorflow_hub для мови Python. З використанням цих моделей запропоновано розробити модель, яка буде навчатися на зазначеній вище вибірці експериментальних даних. Вибір багатомовних моделей обумовлений тим, що, як вже було зазначено, відгуки можуть містити текст не лише у літературній українській мові.

2.2 Проектування інформаційної системи

Конфігурація нейронної мережі на основі обраного набору даних та типу нейронної мережі має структуру, показану на рисунку 2.3.

Вхідний шар перетворює вхідну текстову інформацію в форму тензору в бібліотеці Keras, іншими словами, у символний тензороподібний об'єкт, який обгортатиметься атрибутами, що потрібні для побудови моделі Keras на основі вхідних та вихідних даних. Після цього тензор подається на вхід до шару попередньої обробки, включаючи обгортку викликаного об'єкта, для використання як Keras-сумісний шар, побудований на основі заздалегідь навченої моделі обробки тексту [13].

Ця модель використовує SentencepieceTokenizer [14,15], який токенизує тензор строків у кодуванні UTF-8 та діє як інструмент для розділення тексту на токени та їх подальшого об'єднання.

Наступним в шаровій структурі є кодер RoBERTa. Цей шар базується на попередньо навченій моделі "xlm_roberta_multi_cased_L-12_H-768_A-12" [16], яка виникла в результаті масштабного безконтрольного міжмовного навчання (XLM-RoBERTa) [16]. Модель була навчена на 2,5 терабайтів відфільтрованих даних з CommonCrawl, що охоплюють 100 мов [17].

Наступним шаром є шар відсіву, який рандомно обирає певні вхідні одиниці та встановлює їх значення в 0 на кожному навчальному кроці, що допомагає уникнути перенавчання [18]. Вхідні дані, які не дорівнюють 0, масштабуються так, щоб збережено було загальну суму всіх вхідних даних.

Для класифікації настроїв україномовних відгуків про електронну комерцію розглядалися різні підходи, включаючи нейромережеві та інші методи. Аналіз наукових публікацій вказує на те, що дослідження, які головним чином використовують словникові інструменти для аналізу настроїв у текстових даних та надають перевагу інтерпретації, мають обмежену точність. Серед нейромережевих методів, які були вивчені, BERT-подібні мережі вважаються найбільш ефективними.

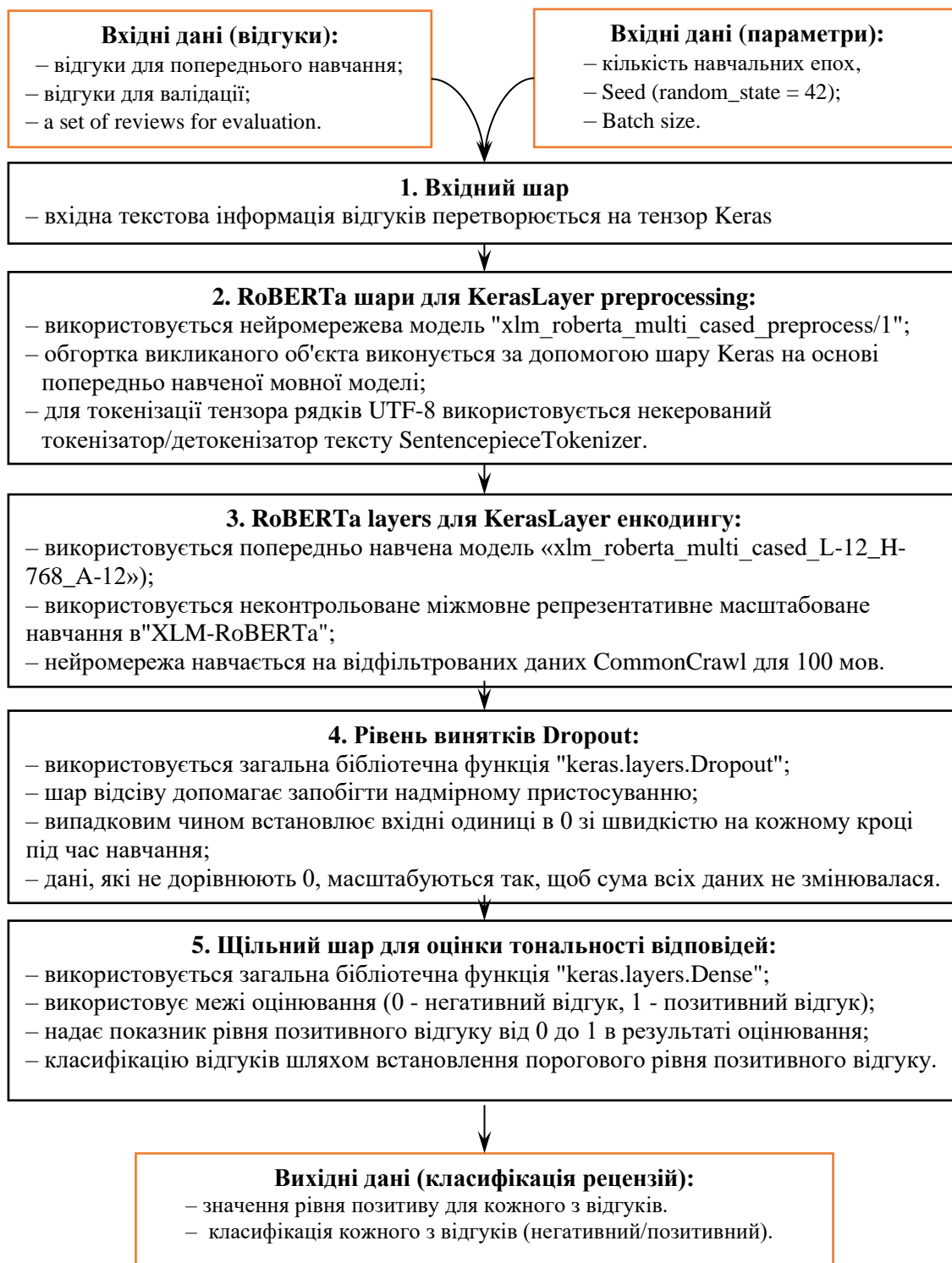


Рисунок 2.3 – Схема класифікатора на основі RoBERTa для класифікації тональності відгуків електронної комерції

BERT був розроблений для того, щоб допомагати комп'ютерам розуміти сенс неоднозначних виразів у тексті, використовуючи контекст для з'ясування

того, в якому контексті був написаний текст [19]. Але дослідники виявили, що моделі, такі як ukr-RoBERTa, ukr-ELECTRA та XLM-R large, показують найкращі результати. Зокрема, XLM-R large та ukr-ELECTRA відзначаються в кращому розпізнаванні довших текстів, тоді як ukr-RoBERTa перевершує інші моделі на коротких послідовностях.

Оскільки в даному дослідженні аналізуються відгуки з інтернет-платформи "HotLine" [20], які зазвичай є короткими текстовими повідомленнями, то було вирішено використовувати нейронну мережу RoBERTa.

Заключним кроком у моделі є процедура класифікації, яка виконується за допомогою функції Dense та генерує результат у діапазоні від 0 до 1, що вказує на ступінь позитивності у відгуках про електронну комерцію українською мовою. Наприклад, 0 вказує на негативний відгук, а 1 - на позитивний. Далі запропоновану модель навчають на наведеній вище вибірці. Навчання проводиться з різними комбінаціями кількісних параметрів, таких як кількість епох навчання, Seed та Batch size.

Кількість епох навчання визначає, скільки разів модель пройде навчальні дані. Параметр Seed встановлюється рівним 42, враховуючи [21], що використання random_state=42 гарантує однаковий набір тестових даних при кожному запуску програми. Batch size визначає кількість навчальних прикладів, які використовуються за одну ітерацію. Оскільки оптимальний розмір партії складно визначити заздалегідь, цей параметр буде вибрано експериментально.

2.3 Вибір засобів розробки інформаційної системи

Для реалізації програмного застосунку необхідно скомпонувати засоби розробки для досягнення коректного результату.

Мова C# володіє значною потужністю та ефективністю, що дозволяє розробляти програми з великим функціоналом. Вона пропонує легкий у засвоєнні синтаксис, має сильну типізацію та включає багато вбудованих

можливостей, що полегшують написання коду та забезпечують високу продуктивність.

Microsoft Visual Studio є високо оціненим середовищем розробки (IDE), відомою своєю інтеграцією з Microsoft ML (Machine Learning). Ця інтеграція є ключовою складовою платформи .NET, розробленою компанією Microsoft. Вона надає розробникам потужні інструменти для вирішення різних завдань у галузі машинного навчання та глибокого навчання [22, 23].

Microsoft ML пропонує розширений набір функцій, призначених для розробки та впровадження моделей машинного навчання. Цей набір включає алгоритми для класифікації, регресії, кластеризації, виявлення аномалій, побудови рекомендаційних систем, обробки природної мови та багато інших можливостей. Навчання нейромережі можна проводити у спеціалізованому розділі машинного навчання [24] від MS Visual Studio, однак для вирішення проблеми цієї роботи було обрано Python, адже він пропонує набагато ширші можливості. Однак побудова інтерфейсу була проведена на базі Windows Forms, адже додатки, що створені таким чином набагато зручніші та простіші в реалізації.

Python [25] надає значні переваги для навчання нейромереж з наукової точки зору. Завдяки тому, що Python має розширену колекцію бібліотек, призначених для роботи з нейромережами, існує широкий спектр інструментів для розробки та навчання нейромереж. Також він має різноманітні бібліотеки для роботи з даними, такі як NumPy, Pandas та Matplotlib. Це полегшує підготовку та аналіз даних перед навчанням нейромереж.

Для вирішення поставленої задачі, процес навчання та дослідження проводився засобами Python. Далі модель була збережена та додана до проєкту у MS Visual Studio, до Windows Forms застосунку. Таким чином, поєднання Python та C# дало змогу реалізувати програмний продукт на базі методу автоматизованого оцінювання відповідності тональності відгуків на товари в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання.

2.4 Формування датасету з текстовими відгуками щодо якості товарів

Як показано в огляді літератури, враховуючи вищезазначені критерії, доступні корпуси слів не підходять для даного дослідження. По-перше, їх загальна кількість обмежена до 3400, що вважається відносно невеликою кількістю. По-друге, особливість твіту полягає в тому, що він завжди є коротким повідомленням і, як правило, містить лише одну фразу. Тому було вирішено використовувати набір відповідей з платформи "HotLine", приклади яких наведені в Таблиці 1.

Таблиця 1 – Відгуки з платформи "HotLine"

Відгук	Користувацька оцінка
<i>Розетка, у вас є совість? С початком війни відмінили в односторонньому порядку усі замовлення. Гроші обіцяли повернути протягом 7 днів. Через 5 днів вже буде місяць як чекаю. При цьому, оператори не відповідають, боти у месенджерах не працюють. Зв'язок відсутній і ще вони приймають нові замовлення</i>	Не рекомендую
<i>Товар замовив і оплатив ще 11 лютого, з тих пір ні слуху ні духу(((невже так важко подзвонити і уточнити?</i>	Не рекомендую
<i>Замовляла товар зі складу розетки (не від партнерів), відправили швидко через днів два, 31.03, очікую вже на оперативну роботу укрпошти.</i>	Рекомендую

Обираючи експериментальні дані, було враховувано потребу у розмовному україномовному контенті, який також повинен бути маркованим. Оцінки ґрунтуються на рейтингах користувачів, що пишуть відгуки, де "Не рекомендую" вказує на негативні відгуки, а "Рекомендую" - на позитивні. Навчальний набір не включає дані з іншими оцінками. Для видобування відгуків було розроблено відповідне програмне забезпечення на основі бібліотеки Crawler [26], яке подальше оброблялося мовою C# та розділено на дві категорії - "позитивні" та "негативні". Такий підхід був використаний також авторами в роботі [27]. Нижче наведено приклад парсингу відгуку:

```
"id": "hotline-464516",
"source": "hotline",
```

"text": "Розетка, у вас є совість? С початком війни відмінили в односторонньому порядку усі замовлення. Гроші обіцяли повернути протягом 7 днів. Через 5 днів вже буде місяць як чекаю. При цьому, оператори не відповідають, боти у месенджерах не працюють. Зв'язок відсутній і ще вони приймають нові замовлення",

"mark": 1

},

Розподіл відгуків у масиві даних проілюстровано на рисунку 2.5.



Рисунок 2.5 – Діаграми розподілів відгуків у датасеті

Загалом набір даних включає 7656 документів, із яких 6655 документів складають навчальний набір, тоді як для валідації використовувалося 1331 документ (що становить 20% від навчального набору). Особливістю цього набору даних є наявність російських слів, ненормативної лексики та частково російськомовних відгуків. Це пов'язано з тим, що навіть при втраті російською мовою домінуючої позиції в соціальних мережах після початку війни, вона все ще є присутньою - 37% постів на українській мові порівняно з 63% на російській. Проте варто зауважити, що ця статистика може варіюватися в залежності від конкретних соціальних мереж [28, 29]. Крім того, відгуки часто містять слова з орфографічними помилками.

Розділ 3 Програмна реалізація інформаційної системи

3.1 Архітектура інформаційної системи

Для розробки програмного застосунку, з використанням нейронної мережі глибокого навчання на платформі .NET для виявлення розбіжностей між відгуками в онлайн-магазині та оцінками користувачів, необхідно визначити структуру та функціональні ролі компонентів системи.

На рисунку 3.1 показано діаграму класів, що представляє різні компоненти програмного застосунку.

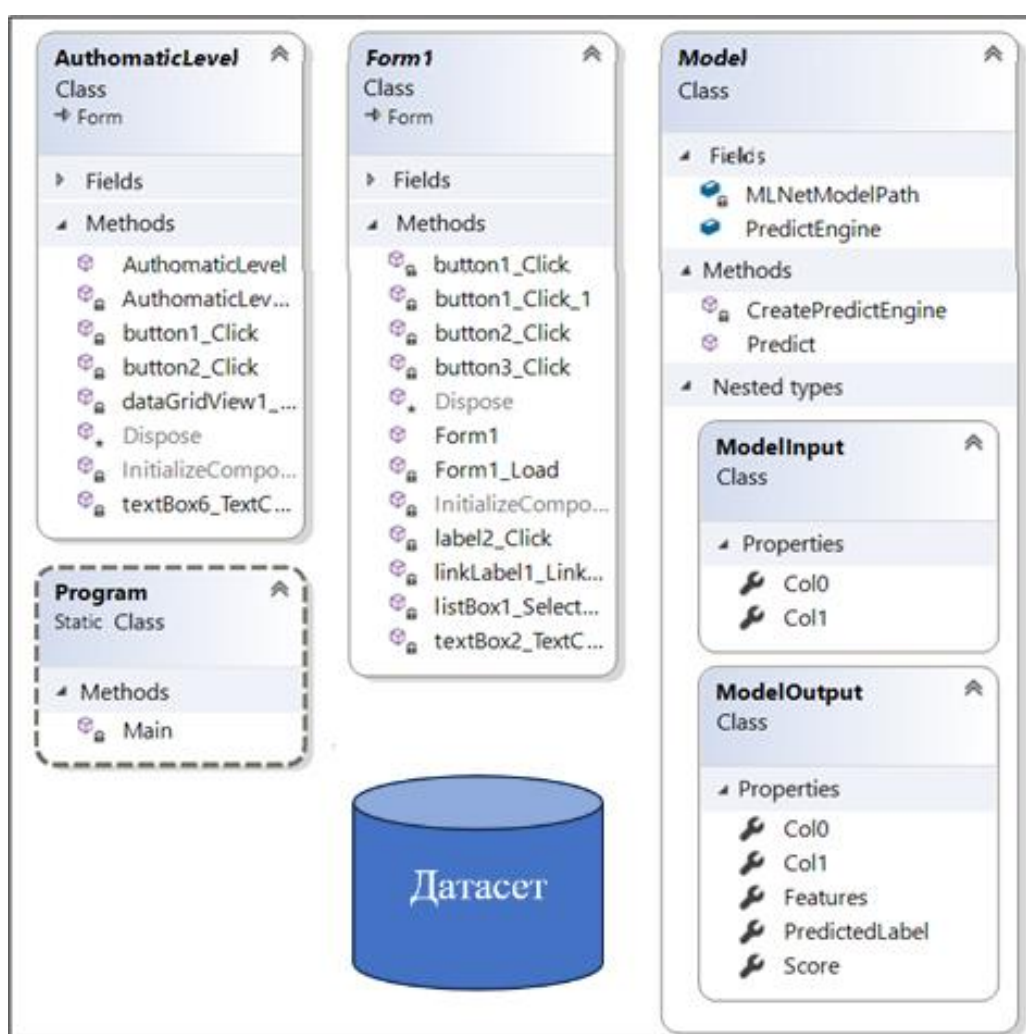


Рисунок 3.1 – Діаграма класів програмного продукту

Програма включатиме в себе форми інтерфейсу користувача, нейромережеву модель і набір даних, з яким будуть взаємодіяти різні модулі в межах програмного забезпечення.

Клас "Form1" призначено для спілкування користувача з підсистемою, відповідальною за визначення рівня позитивного відгука користувача. Ця підсистема виконує наступні функції:

- перегляд користувацьких відгуків.
- додавання нових відгуків до бази даних.
- видалення існуючих відгуків із бази даних.

Основна функція цієї форми - визначення емоційного тону текстового відгука та перехід до наступної форми, "Автоматизатор виявлення невідповідності", де автоматизовано виявляються розходження між відгуками та оцінками.

Клас "Автоматизатор виявлення невідповідності до оцінки" виконує наступні завдання:

- вибір відгуку для подальшого аналізу.
- перегляд змісту відгуку та порівняння його емоційного забарвлення з користувацькою оцінкою.
- розрахунок ступеня відповідності відгуку до користувацької оцінки.

Клас "ModelRoberta" містить інформацію про шлях до навченої нейромережевої моделі і відповідає за обчислення, пов'язані із визначенням емоційного забарвлення користувацьких текстових відгуків.

Крім того, в програмному продукті присутній датасет, який містить розмічені відгуки про товари. Цей датасет взаємодіє з усіма модулями системи.

Отже, було описано структуру та функціональне призначення програмних компонентів системи, і була створена діаграма класів для програмного застосунку, який використовує метод виявлення невідповідності між відгуком про техніку в інтернет-магазині та користувацькою оцінкою за допомогою неймережі глибокого навчання.

3.2 Особливості реалізації програмних складових системи

Для реалізації програмного застосунку було створено відповідний інтерфейс (рисунок 3.2).

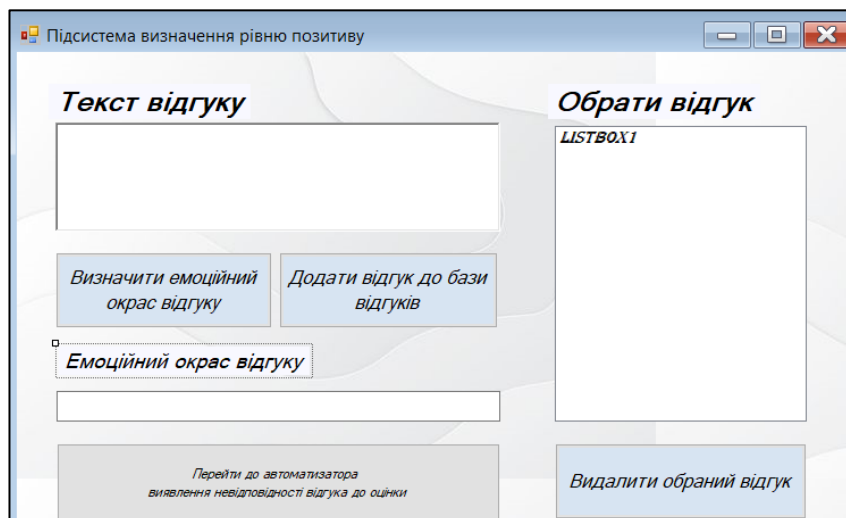


Рисунок 3.2 – Інтерфейс головного вікна застосунку

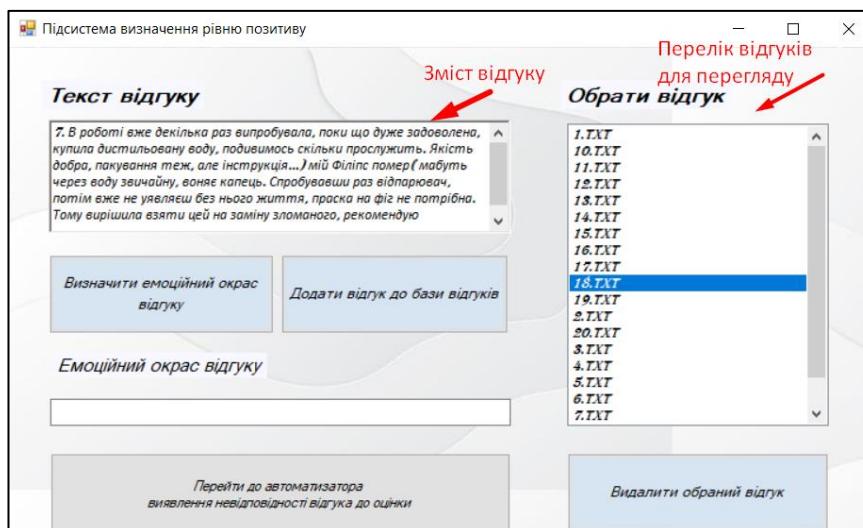


Рисунок 3.3 – Приклад роботи програми

В полі "Текст відгуку" користувач може переглянути зміст кожного текстового повідомлення. В розділі "Обрати відгук" користувач має можливість вибрати конкретний відгук для подальшого дослідження (див. рисунок 3.3). У

спеціально відведеному полі "Рівень позитиву у відгуку" буде відображатися числова оцінка ступеня емоційного забарвлення відгуку.

Після натискання кнопки "Перейти до автоматизатора виявлення невідповідності відгука до оцінки" з'являється нова форма, яка називається "Автоматизатор виявлення невідповідності відгука до оцінки". Тут реалізована можливість перегляду відгука, його емоційного забарвлення та користувацької оцінки в формі таблиці. До цієї форми також додано функціональність, яка дозволяє виконати наступні дії:

- переглянути бінарну користувацьку оцінку відгука;
- переглянути числову оцінку відгука;
- переглянути бінарну нейромережеву оцінку відгука;
- переглянути числову нейромережеву оцінку відгука;
- отримати значення різниці між користувацькою та нейромережевою оцінками.

Результат наведено на рисунку 3.4.

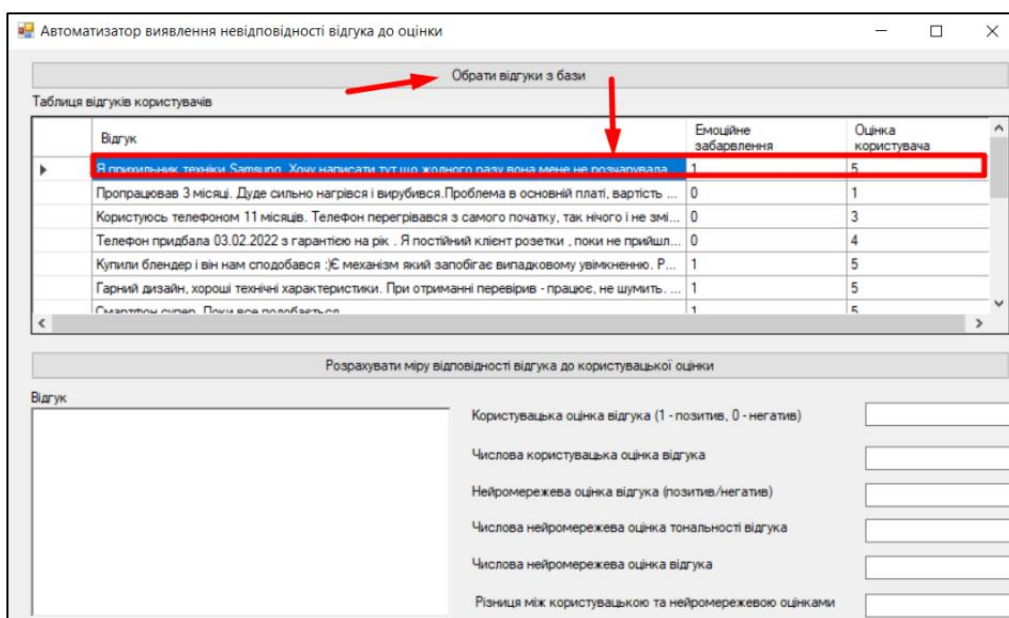


Рисунок 3.4 – Вікно «Автоматизатор виявлення невідповідності відгука до оцінки»

Після вибору відгуку та натискання кнопки "Розрахувати міру відповідності відгука до користувацької оцінки" виконується функція, яка

обчислює різницю між нейромережевою оцінкою відгука та оцінкою користувача. Якщо ця різниця перевищує значення 1.5, то відгук вважається невідповідним користувачькій оцінці. Результат роботи функції наведено на рисунку 3.5.

Автоматизатор виявлення невідповідності відгука до оцінки

Обрати відгуки з бази

Таблиця відгуків користувачів

Відгук	Емоційне забарвлення	Оцінка користувача
Я прихильник техніки Samsung. Хочу написати тут що жодного разу вона мене не розчарувала...	1	5
Пропрацював 3 місяці. Дуже сильно нагрівся і вирубився. Проблема в основній платі, вартість ...	0	1
Користуюсь телефоном 11 місяців. Телефон перегрівався з самого початку, так нічого і не змі...	0	3
Телефон придбала 03.02.2022 з гарантією на рік. Я постійний клієнт розетки, поки не прийшл...	0	4
Купили блендер і він нам сподобався ;С механізм який запобігає випадковому увімкненню, Р...	1	5
Гарний дизайн, хороші технічні характеристики. При отриманні перевірів - працює, не шумить, ...	1	5
Смартфон супер. Поки все роблять...	1	5

Розрахувати міру відповідності відгука до користувачької оцінки

Відгук

Телефон придбала 03.02.2022 з гарантією на рік. Я постійний клієнт розетки, поки не прийшлося спілкуватися з сервісом. Під кінець грудня телефон почав зависати та самостійно вимкнутися, за день так раз 20-30. Віддала по гарантії на ремонт, акт прийому від 23.01.2023 року. Сьогодні 6 березня і немаю жодних новин щодо телефону, хоча по строкам ремонт до 28 днів. В чат боті стандартні повідомлення, з вами зв'язується. Досі немає рішення по діагностиці і коли воно буде, невідомо. Бо стандартна відписка від оператора, що в країні воєнний стан. Шановна Розетка, мені дуже прикро, що для вас зовсім неважливі постійні клієнти

Зміст відгука

Відгук не відповідає числовій оцінці користувача

Користувачька оцінка відгука (1 - позитив, 0 - негатив) 0

Числова користувачька оцінка відгука 4

Нейромережева оцінка відгука (позитив/негатив) Відгук негативний

Числова нейромережева оцінка тональності відгука 0,5503924

Числова нейромережева оцінка відгука 1,748038

Різниця між користувачькою та нейромережевою оцінками 2,251962

Рисунок 3.5 – Результат роботи функцій розділу «Автоматизатор виявлення невідповідності відгука до оцінки»

Отже, було розроблено програмний продукт на базі методу автоматизованого оцінювання відповідності тональності відгуків на товари в інтернет-магазинах до їх користувачької оцінки з використанням нейромереж глибокого навчання. Конкретно, були реалізовані модулі для оцінки рівня позитивності коментарів та для автоматичного виявлення невідповідностей між відгуком та оцінкою користувача.

3.3 Результати досліджень

За обраними параметрами були визначені метрики для оцінки продуктивності моделі, включаючи час навчання в секундах, точність і втрати.

Для функції втрат використовувалася бінарна крос-ентропійна функція, виражена наступною формулою [33]:

$$Loss = -\frac{1}{N} \left[\sum_{j=1}^N [t_j \log(p_j + (1 - t_j) \log(1 - p_j))] \right],$$

де N – кількість вибірок даних, t_j – істинне значення, яке приймає значення 0 або 1, p_j – ймовірність Softmax для i -ої точки даних.

Точність дослідження визначається як відношення кількості правильних відповідей до загальної кількості відповідей. Отримані результати оцінки продуктивності (час навчання, точність, втрати) різних параметрів налаштування моделі (кількість епох навчання, seed, розмір партії) нейронної мережі наведені в таблиці 3.1. Експеримент проводився на комп'ютері з процесором Intel Core i7 восьмого покоління, 16 ГБ оперативної пам'яті та відеокартою NVIDIA GeForce MX150.

Таблиця 3.1 – Параметри перепідготовки класифікатора

Параметри	V1	V2	V3	V4	V5	V6
Кількість епох навчання	3	3	4	5	3	10
Seed	42	42	42	42	42	42
Batch size	64	32	32	32	16	64
Час навчання (sec)	10028	9224	12158	15894	10248	33952
Точність	0.92	0.91	0.91	0.91	0.91	0.92
Loss	0.29	0.31	0.30	0.32	0.31	0.30

Як можна побачити з таблиці 3.1, модель V1 досягає найвищої точності, яка становить 0,92, та найнижчих втрат, які дорівнюють 0,29. У той час як модель V6 також досягає точності на рівні 0,92, втрати у неї становлять 0,30, при цьому час навчання значно більший.

Навіть при невеликих відхиленнях в точності, майже всі версії навчених моделей дали результати на реальних прикладах, які були подібні до експертних оцінок. Деякі з цих оцінок представлені в таблиці 3.2 для порівняння різних версій навчених моделей (V1-V6 з таблиці 3.1).

Маючи на увазі, що тестова версія RoBERTa є багатомовною моделлю-трансформером, яка була навчена на двомовних даних, важливо зауважити, що

нейронна мережа успішно впоралася з ідентифікацією настрою, як це показано в таблиці 3.2.

Таблиця 3.2 – Класифікація тональності відгуків

Відгуки	Evaluation (V4)	Evaluation (V2)	Evaluation (V6)
<i>(ua) Ваш товар повне лайно, гіршого і годі шукати</i>	0.005181	0.014710	0.000641
<i>(ua) Як добре, що є такі хороші продавці, як ви.</i>	0.988962	0.991397	0.995478
<i>(ua) Наша сім'я вкотре купує тут товари і завжди сервіс на висоті, рекомендуємо</i>	0.948719	0.990182	0.871778
<i>(ua) Ніколи не рекомендую користуватись цим сервісом! Це просто жах!</i>	0.002086	0.011778	0.000665
<i>(ua) Драйверів на комп'ютері не було взагалі. 13/01/2023 з самого ранку поніс комп'ютер в магазин на повернення або обмін на іншу модель, як виявилось обміняти вони його не змогли, не дивлячись на те, що я вибрав дорожчу модель а лише оформили повернення. Мусили сидіти в магазині 2години та чекати поки продавець не скине до заводських налаштувань Yero, лише тоді сказали зможуть оформити повернення (жах просто, ми навіть ним не користувались і це видно було)</i>	0.004561	0.016255	0.001831
<i>(ua) Як на мене, Розетка - найкращий магазин. Великим плюсом є безкоштовна доставка в усі їх відділення. По гарантії теж питань немає, тож рекомендую цей магазин</i>	0.995170	0.988024	0.957222

Аналіз відповідей, які не входили в навчальну та тестову вибірки, підтвердив ефективність запропонованої архітектури. Треба відзначити, що навчальна вибірка не була очищена вручну, тому можливі деякі неправильно класифіковані відгуки. Проте це не суттєво вплинуло на загальну точність бінарної класифікації емоційної тональності відгуків, які містять не лише чисту українську мову, але й суржик та двомовні дані.

На рисунку 3.5 видно, як точність змінюється залежно від кількості пройдених епох, і на рисунку 3.6 представлена зміна функції втрат для комбінації параметрів навчання V1 із таблиці 2 (3 епохи, 64 розміри партії).

З аналізу графіка на рисунку 3.6 видно, що обмежена кількість епох навчання не вистачає для досягнення стабільного результату, оскільки точність показує тенденцію до зростання, а функція втрат - до зменшення, не досягаючи стабільності.

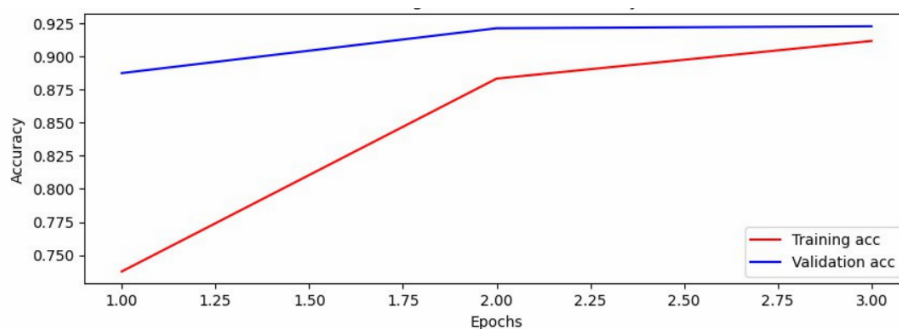


Рисунок 3.5 – Ілюстрація процесу навчання за епохами з точки зору точності (V1)

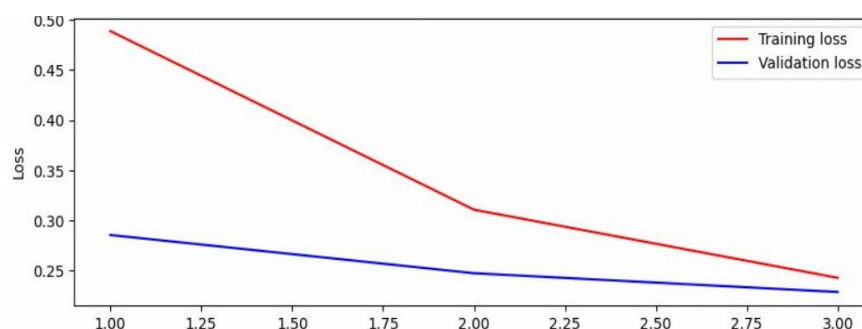


Рисунок 3.6 – Ілюстрація процесу навчання за епохами за функцією втрат (V1)

Продовживши експеримент та збільшивши кількість епох навчання до 10, що відповідає параметру V6 із таблиці 2, було отримано результати, які ілюструються на рисунках 3.7 і 3.8.

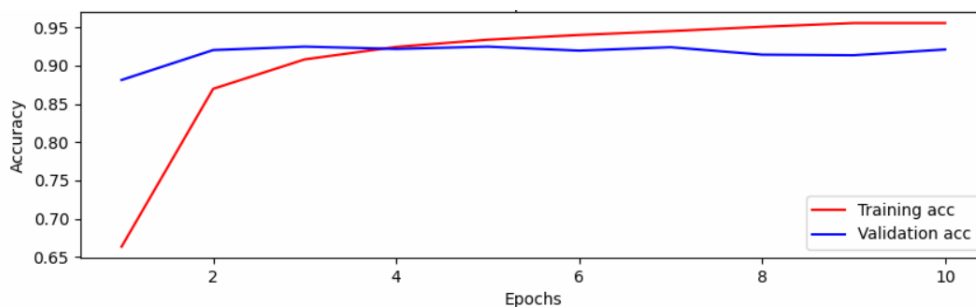


Рисунок 3.7 – Ілюстрація процесу навчання за епохами з точки зору точності (V6)

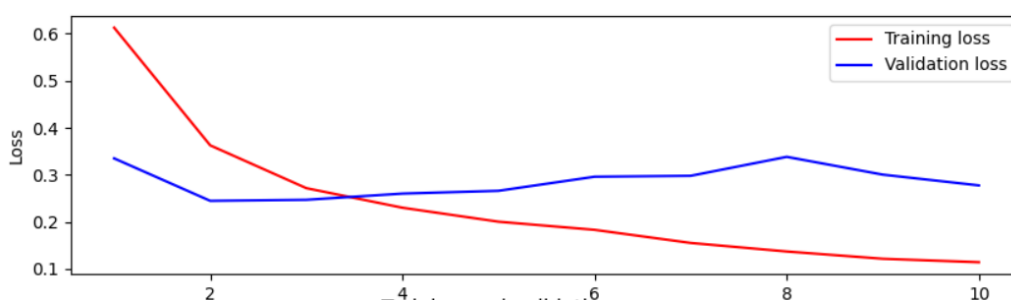


Рисунок 3.8 – Ілюстрація процесу навчання за епохами за функцією втрат (V6)

Результати вказують на те, що використання валідаційної вибірки не призвело до підвищення точності класифікації. Загалом, функція втрат мала тенденцію незначного зростання після третьої ітерації для валідаційної вибірки. Однак ці результати можуть вказувати на те, що вибірки, ймовірно, не були достатньо обгрунтовано відфільтровані. Проведене тестування нейромережі на відгуках, які відсутні в базі даних, показало майже безпомилкові результати для 40 відгуків, що містять емоції. До позитивної вибірки увійшли такі відгуки, як "Мікрохвильовка", "Купив комп'ютер", "Купив навушники", "Купив пілосос" і так далі. Однак схожі відгуки можуть зустрічатися і в негативній вибірці.

Графік, який ілюструє завершення процесу перекваліфікації по епохах для параметра V4 із таблиці 2, представлено на рисунках 3.9 та 3.10.

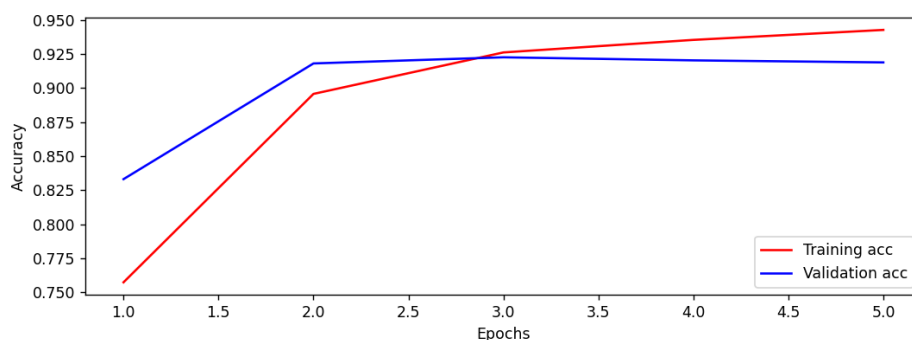


Рисунок 3.9 – Ілюстрація процесу навчання за епохами з точки зору точності (V4)

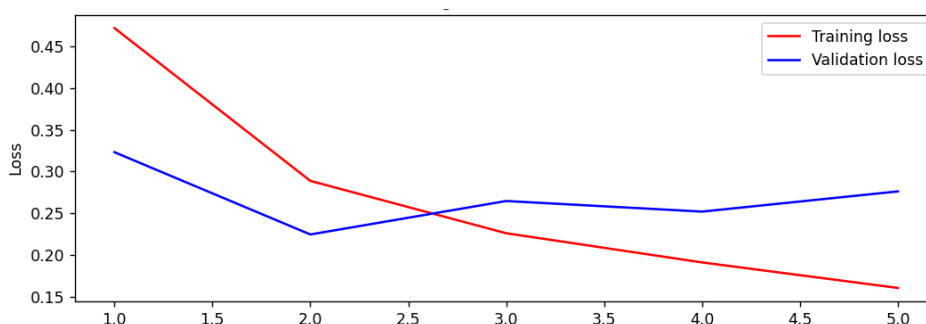


Рисунок 3.10 – Ілюстрація процесу навчання за епохами за функцією втрат (V4)

Результати цього експерименту підтверджують, що набір даних не був вручну очищений. За збільшення кількості епох нейромережа починає "запам'ятовувати", які відгуки належать до конкретної епохи. Це видно з червоної лінії на графіках 8-9 та 10-11. Хоча функція втрат для навчальної

вибірки набагато менше, точність вища. Проте отримані значення функції втрат і точності пояснюються тим, що вибірка не була відфільтрована вручну і включала в себе відгуки, які містили неемоційні коментарі, часто складаючись з одного слова або фрази, такі як "мікрохвильовка", "купив комп'ютер", "купив навушники", "купив пілосос" і так далі.

Крім того, аналіз оцінки тональності показав, що нейромережа успішно впоралася із завданням на 40 фразах, яких не було в навчальній або тестовій вибірках. Ці фрази були попередньо оцінені експертом і містили як стилістичні, так і орфографічні помилки, а також були представлені багатомовними даними.

В таблиці 3.3 наведено приклад використання застосунку, зокрема виділені відгук, оцінка користувача, нейромережева оцінка, визначено різницю між нейромережевою та користувацькою оцінкою та висновок (відповідає коментар оцінці чи ні).

Таблиця 3.3 – Порівняння користувацької та нейромережевої оцінки

Відгук	Оцінка користувача	Нейромережева оцінка	Різниця	Висновок
<i>(ua) Я дуже не задоволена доставкою, кур'єр хамло</i>	5 зірок	0.014710	0.98529	Не відповідає
<i>(ua) Ми дуже задоволені покупкою, повернемося ще</i>	5 зірок	0.990176	0.007575	Відповідає
<i>(ua) Якщо чесно, товар так собі. Якісь зазначена висока, а прийшла тряпочка просто</i>	2 зірки	0.3998613	0.1998613	Відповідає
<i>(ua) Наша сім'я вкотре купує тут товари і завжди сервіс на висоті, рекомендуємо</i>	5 зірок	0.990182	0.009818	Відповідає
<i>(ua) Ніколи не рекомендую користуватись цим сервісом! Це просто жах!</i>	1 зірка	0.011778	0.188222	Відповідає

Дослідження вказують на те, що нейромережа дійсно може "розуміти" контекст. В деяких випадках в нейромережі можливі незначні коливання, особливо при нейтральних відгуках, наприклад: "Ціна правильна, наявність теж." Цей відгук був оцінений як "Рекомендую," але нейромережа визначила його як позитивний з майже маргінальною оцінкою 0,505790. Подібні обробляються і подібні відгуки, наприклад: "Замовив стіл Ambrosio Halmar. Дуже задоволений покупкою ???? повна відповідність фото і швидка доставка (менше двох тижнів). Рекомендую ??????" Оцінка нейромережі для цього відгуку - 0,902363, але експерт може неоднозначно сприймати знаки питання.

Висновки

У роботі було розроблено та програмно реалізовано метод автоматизованого оцінювання відповідності тональності відгуків на товари повсякденною українською мовою в інтернет-магазинах до їх користувацької оцінки з використанням нейромереж глибокого навчання. У процесі розробки методу було досліджено ряд ключових аспектів, включаючи створення маркованого набору даних для навчання нейронної мережі, вибір та налаштування нейромережевого класифікатора, а також побудову семантичної моделі мови. Серед нейронних мереж було обрано й використано архітектуру BERT, яка вважалася однією з найточніших, зокрема її модифікацію RoBERTa, яка є більш ефективною для аналізу коротких текстових документів.

Оскільки україномовні відгуки електронної комерції мають свої унікальні особливості, було підготовлено власний набір даних, який включав 7656 відгуків. Цей набір даних був поділений на навчальну та тестову вибірки, і в кожній з них були представлені як негативні, так і позитивні відгуки. Для оцінки продуктивності запропонованої архітектури були використані метрики точності та втрат. Результати показали, що для об'єднаних багатомовних відгуків точність склала 0,92, і функція втрат мала значення 0,29.

Запропонований підхід має свої обмеження. Він ефективний для коротких текстових рецензій, не довших за 500 слів, написаних українською мовою, включаючи суржик та мовні вкраплення. Майбутні дослідження спрямовані на використання цього класифікатора для оцінювання роботи менеджерів інтернет-магазинів у спілкуванні з клієнтами, впровадження моделей зворотного маркетингу та підвищення ефективності багатомовних класифікаторів. Одержані результати доповідались автором даної роботи на 3 науково-практичних конференціях, опубліковано 3 праці (1 у виданні, що індексується в Scopus, та 2 в збірниках тез Всеукраїнських конференцій).

Перелік посилань

1. Anderson, S. (2012). Online Product Reviews: Insights for Improving Satisfaction and Loyalty. *Journal of Interactive Marketing*, 26(2), 89-97. doi:10.1016/j.intmar.2011.12.003
2. Недоліки інтернет-магазинів. URL: <https://www.tviysvit.com.ua/pokupky-v-interneti-perevahy-ta-nedoliky/>
3. Wei Li, Wei Shao, Shaoxiong Ji, Erik Cambria. "BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis". *Neurocomputing* (2022): 73-82. doi: 10.1016/j.neucom.2021.09.057.
4. Kovalchuk O., Slobodzian V., Sobko O., Molchanova M., Mazurets O., Barmak O., Krak I., Savina N. Visual Analytics-Based Method for Sentiment Analysis of COVID-19 Ukrainian Tweets. Book Chapter. *Lecture Notes on Data Engineering and Communications Technologies*. 2023. Vol. 149. pp. 591–607. DOI: 10.1007/978-3-031-16203-9_33
5. Walaa Medhat, Ahmed Hassan, Hoda Korashy. "Sentiment analysis algorithms and applications: A survey". *Ain Shams Engineering Journal*, Vol. 5, Issue 4, (2014): 1093-1113. doi: 10.1016/j.asej.2014.04.011.
6. Hui Li, Qi Chen, Zhaoman Zhong, Rongrong Gong, Guokai Han. "E-word of mouth sentiment analysis for user behavior studies". *Information Processing & Management* (2022). doi: 10.1016/j.ipm.2021.102784.
7. Marcec, R., Likic, R. "Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines". *Postgraduate Medical Journal*, Volume 98, Issue 1161, (2022): 544–550. doi: 10.1136/postgradmedj-2021-140685
8. Qorib, M., Oladunni, T., Denis, M., Ososanya, E., Cotae, P. "Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset". *Expert Systems with Applications* (2023). doi: 10.1016/j.eswa.2022.118715.

9. Jochen Hartmann, Mark Heitmann, Christian Siebert, Christina Schamp. "More than a Feeling: Accuracy and Application of Sentiment Analysis". International Journal of Research in Marketing (2022). doi: 10.1016/j.ijresmar.2022.05.005.
10. AFINN. URL: <https://darenr.github.io/afinn/>
https://issuu.com/arctionltdchart/docs/what_is_windows_forms_and_its_benefits
11. Ai.Facebook.Com., RoBERTa: An optimized method for pretraining self-supervised NLP systems. URL: <https://ai.facebook.com/blog/roberta-an-optimized-method-for-pretraining-self-supervised-nlp-systems>.
12. Tfhub.Dev., Text preprocessing model xlm_roberta_multi_cased_preprocess. URL: https://tfhub.dev/jeongukjae/xlm_roberta_multi_cased_preprocess/1.
13. Tfhub.Dev., Unsupervised Cross-lingual Representation Learning at Scale. xlm_roberta_multi_cased_L-12_H-768_A-12. URL: https://tfhub.dev/jeongukjae/xlm_roberta_multi_cased_L-12_H-768_A-12/1.
14. Tensorflow.Org., Text.SentencepieceTokenizer. URL: https://www.tensorflow.org/text/api_docs/python/text/SentencepieceTokenizer.
15. Tfhub.Dev., Unsupervised Cross-lingual Representation Learning at Scale. xlm_roberta_multi_cased_L-12_H-768_A-12. URL: https://tfhub.dev/jeongukjae/xlm_roberta_multi_cased_L-12_H-768_A-12/1.
16. Huggingface.Co., XLM-RoBERTa (base-sized model). URL: <https://huggingface.co/xlm-roberta-base>
17. ATensorflow.Org., Tf.keras.layers.Dropout. URL:https://www.tensorflow.org/api_docs/python/tf/keras/layers/Dropout.
18. Sentiment Analysis of Movie Reviews with Google's BERT. URL: <https://medium.com/mllearning-ai/sentiment-analysis-of-movie-reviews-with-googles-bert-c2b97f4217f>
19. Hotline.Ua., Reviews of the store Rozetka. URL: <https://hotline.ua/ua/yp/2476/reviews>

20. Thecleverprogrammer.com., Why Random_state=42 in Machine Learning? URL: https://thecleverprogrammer.com/2020/12/17/why-random_state42-in-machine-learning.
21. Microsoft ML. URL: <https://learn.microsoft.com/en-us/dotnet/machine-learning/how-does-ml-dotnet-work>
22. ML for Visual Studio. URL: <https://dotnet.microsoft.com/en-us/apps/machinelearning-ai/ml-dotnet/model-builder>
23. What is Model Builder and how does it work? URL: <https://learn.microsoft.com/en-us/dotnet/machine-learning/automate-training-with-model-builder>
24. Python. URL: <https://www.python.org/>
25. Crawlee.Dev., A web scraping and browser automation library. URL: <https://crawlee.dev>.
26. G. Maccario, M. Naldi. "Alexa, Is My Data Safe? The (Ir)relevance of Privacy in Smart Speakers Reviews" in: International Journal of Human–Computer Interaction (2022), 13. doi: 10.1080/10447318.2022.2058780.
27. Speka.Media., What language do Ukrainian social networks speak URL: <https://speka.media/socialni-merezi/yakoyu-movoyu-govoryat-ukrayinski-socmerezi-v5m019>.
28. BBC. News. Ukrainian has significantly strengthened in all spheres: at home, at work and on the Internet. Poll. URL: <https://www.bbc.com/ukrainian/news-64201995>