

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

Єрмолаєв Олександр Андрійович

УДК 004.8

**Система визначення якості товарів в інтернет-магазинах
на основі штучного інтелекту**

Галузь знань 12 «Інформаційні технології» за спеціальністю
122 «Комп'ютерні науки»
122 - БКР.А - 402.21710210

Автореферат
бакалаврської кваліфікаційної роботи на здобуття освітньої кваліфікації
«бакалавр з комп'ютерних наук»

Миколаїв – 2021

Бакалаврська кваліфікаційна робота є рукопис.

Робота виконана в Чорноморському національному університеті імені Петра Могили Міністерства освіти і науки України на кафедрі інтелектуальних інформаційних систем

Науковий керівник: старший викладач С.Ю. Борольова

Рецензент: старший викладач Ю.О. Нездолій

Захист відбудеться «24» червня 2021 р. о 9⁰⁰ год. на засіданні екзаменаційної комісії (ауд. 2-403) у Чорноморському національному університеті імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

З бакалаврською кваліфікаційною роботою можна ознайомитися в бібліотеці Чорноморського національного університету імені Петра Могили за адресою: 54003, м. Миколаїв, вул. 68-ми Десантників, 10.

Автореферат представлений «___» червня 2021 р.

Секретар
екзаменаційної комісії,
викладач

О.С. Скакодуб

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми – у наш час все більше і більше покупок відбувається в інтернеті. Крім того, нещодавня пандемія COVID-19 посприяла розвитку онлайн торгівлі. З'являється все більше онлайн-магазинів, які в свою чергу продають все більше товарів. Ці інтернет-магазини знаходяться в різних країнах та мають різні вимоги до якості товарів. А так як усе це відбувається в інтернеті, де складно з'ясувати якість товару лише по зображенню, то єдиним критерієм якості являються відгуки (коментарі) реальних людей які вже робили цю покупку.

Метою бакалаврської кваліфікаційної роботи є автоматизація процесу визначення якості товарів в інтернет-магазинах за рахунок розробки системи з використанням штучного інтелекту.

Практичне значення отриманих результатів полягає у наданні користувачам можливості на основі відгуків аналізувати якість товарів в інтернет-магазинах, а також порівнювати якість товарів в різних інтернет-магазинах.

Структура кваліфікаційної роботи. Пояснювальна записка до бакалаврської кваліфікаційної роботи складається із вступу, 5 розділів, висновків, додатків. Загальний обсяг роботи складає 87 сторінки, 34 рисунків, 14 таблиць та 30 посилань на літературні джерела.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі описано актуальність, мету, предмет та об'єкт дослідження, основні завдання та засоби для їх вирішення.

У першому розділі виконується дослідження предметної сфери та розглядаються існуючі аналоги для аналізу тональності тексту, визначається технічне завдання. Розглянуто найпопулярніші рішення та отримано висновок, що на даний момент часу системи, яка б вирішувала всі необхідні задачі та потреби, не існує, а тому розробка такого рішення є досить актуальною. Крім того, у кожному з існуючих рішень потрібно самому збирати відгуки з сайту інтернет-магазину. Також у даному розділі було розроблено технічні та функціональні вимоги до застосунку.

Сфера аналізу тональності тексту є досить нової, а нові технології у ній з'являються кожен рік, тому на даний момент не існує якогось загальноприйнятого рішення. Крім того, багато компаній намагають розроблювати схожі системи, які орієнтовані на вузьку сферу товарів. Популярним шляхом аналізу тональності слів є використання тонального словника, в якому кожне слово співвідноситься оцінка, наприклад “позитивна”, “негативна” або “нейтральна”. Для отримання кінцевого результату потрібно обчислити значення двох оцінок: позитивної складової тексту і негативною. Для того, щоб знайти позитивну складову тексту необхідно знайти суму тональностей всіх позитивних термінів тексту з урахуванням їх ваги.

Назва продукту: Review Analyze (з англійської – аналізатор коментарів).

Розробка повинна складатися з двох частин:

1. створення нейромережі для аналізу коментарів;
2. створення системи яка буде збирати коментарі по запиту користувача та видавати результат використовуючи нейромережу.

Вимоги до функціональних характеристик нейромережі:

- класифікація коментарів по шкалі від 1 до 5 балів;
- середнє квадратичне відхилення оцінки менше одного балу;
- точність більше ніж 70%;
- навчальний розмір датасету більше ніж 500'000 коментарів;
- швидкість передбачення нейромережі менше ніж 100 мс.

Вимоги до функціональних характеристик системи:

- можливість аналізу власного тексту;
- можливість аналізу коментарів за певним посиланням;
- можливість аналізу товару на декількох інтернет-магазинах за його назвою;
- зручне API для нейромережі яке дозволяє аналізувати як один певний коментар, так и декілька одразу;
- зручний інтерфейс з використанням фреймворку Django.

Вимоги до технічного забезпечення:

Навчати рекурентні нейромережі найкраще всього на відеокартах (GPU). Для того щоб процес навчання був швидшим, потрібно мати кращу відеокарту та більше оперативної пам'яті (RAM). Мінімальна апаратна конфігурація комп'ютера, на якому може бути запущене навчання нейромережі:

- відеокарта Tesla K80 або краща;
- процесор: Xeon 2.3 Ghz або кращий;
- від 16 ГБ оперативної пам'яті;
- від 32 ГБ вільного місця на SSD;
- стабільне швидкісне інтернет з'єднання від 100 МБ/с.

У другому розділі розглядаються типи нейромереж, описуються можливі проблеми при навчанні та обирається який тип нейромереж краще підходить.

За результатами даного розділу було досліджено та детально розглянуто моделі та алгоритми, які допоможуть розв'язати задачі, що поставлені перед системою, яка має бути результатом виконання БКР. Для аналізу тональності тексту були використані рекурентні нейронні мережі, а саме структуру нейрону під назвою LSTM яка має довгу короткочасну. Цей вибір було аргументовано здатністю цих нейронних мереж пам'ятати останні проаналізовані слова, а отже запам'ятовує контекст. Було детально описано можливу проблему перенавчання та описано алгоритм по запобіганню цього явища.

Нейронна мережа – математична модель, а також її програмне або апаратне втілення, побудована за принципом організації та функціонування біологічних нейронних мереж – нервових клітин живого організму. Це поняття з'явилося при вивченні процесів, що протікають в мозку, і при спробі змодельювати ці процеси. Після розробки алгоритмів навчання одержувані моделі стали використовувати в практичних цілях: в задачах прогнозування, для розпізнавання образів, в задачах управління.

Нейронні мережі не програмуються в звичному сенсі цього слова, вони навчаються. Можливість навчання – одна з головних переваг нейронних мереж перед традиційними алгоритмами. Технічно навчання полягає в знаходженні коефіцієнтів зв'язків між нейронами. В процесі навчання нейронна мережа здатна

виявляти складні залежності між вхідними даними і вихідними, а також виконувати узагальнення. Це означає, що в разі успішного навчання мережа зможе повернути вірний результат на підставі даних, які були відсутні в навчальній вибірці, а також неповних та “зашумлених”, частково спотворених даних.

Одні і ті ж види моделей машинного навчання можуть вимагати різні розміри, ваги та швидкості навчання для різних видів даних. Ці параметри називаються гіперпараметрами і їх слід налаштовувати так, щоб модель могла оптимально вирішити конкретне завдання. Для цього знаходиться кортеж гіперпараметрів, який дає оптимальну модель, що оптимізує задану функцію (або набір функцій) втрат на заданих незалежних даних. Цільова функція бере кортеж гіперпараметрів і повертає пов'язані з ними втрати. Часто використовується перехресна перевірка для оцінки цієї узагальнюючої здатності.

Найпоширеніші типи нейромереж:

- Перцептрон
- Багат шаровий перцептрон
- Рекурентна нейронна мережа
- Згорочна нейронна мережа
- Імпульсна нейронна мережа

Найвідоміший варіант алгоритму навчання нейронної мережі – алгоритм зворотного поширення *back propagation*. В ньому обчислюється вектор градієнта поверхні помилок. Цей вектор вказує напрямком найкоротшого спуску по поверхні з даної точки, тому якщо “трохи” просунутись по ньому, помилка зменшиться. Послідовність таких кроків, розмір яких буде зменшуватись з часом, врешті-решт призведе до мінімуму того чи іншого типу. Певні труднощі тут представляє питання про те, яку потрібно брати довжину кроків.

При великій довжині кроку збіжність буде швидшою, але є небезпека перестрибнути через рішення або (якщо поверхня помилок має особливо химерну форму) піти в неправильному напрямку. Класичним прикладом такого явища при навчанні нейронної мережі є ситуація, коли алгоритм дуже повільно просувається по вузькому яру з крутими схилами, стрибаючи з одного його боку на іншу.

Навпаки, при маленькому кроці, ймовірно, буде схоплено вірний напрям, однак при цьому потрібно дуже багато ітерацій. На практиці величина кроку береться пропорційної крутизни схилу (так що алгоритм уповільнює хід поблизу мінімуму) з деякою константою, яка називається швидкістю навчання. Правильний вибір швидкості навчання (значення lr (learning rate) у даній реалізації) залежить від конкретного завдання і зазвичай здійснюється досвідченим шляхом. Ця константа може також залежати від часу, зменшуючись у міру просування алгоритму.

Основною здатністю нейромереж є навчання – це одне з головних переваг нейронних мереж перед традиційними алгоритмами. Навчання нейромережі полягає в знаходженні коефіцієнтів зв'язків між нейронами. У процесі навчання нейронна мережа здатна виявляти складні залежності між вхідними даними й вихідними, а також виконувати узагальнення. Це значить, що, у випадку успішного навчання, мережа зможе повернути вірний результат на підставі даних, які були відсутні в навчальній вибірці, а також неповних або частково змінених даних.

У третьому розділі описується процес створення та навчання нейромережі. Також описуються рішення для підвищення точності передбачень та метрики, за допомогою яких оцінюється точність нейромережі. За результатами даного розділу було створено, навчено та оцінено нейромережу для аналізу відгуків користувачів в інтернет-магазинах. Було описано вимоги до датасету та обрано відповідний набір даних. Навчальна та валідаційні частини датасету містили у сумі 650'000 коментарів, тестова частина – 50'000 коментарів. Після завантаження цей датасет було нормалізовано та перетворено у вектор.

Для створення та навчання нейромережі було використано фреймворк Torch та середовище Jupyter Notebook. Навчання нейромережі відбувалося на ресурсах платформи Google Colab. Для того щоб навчання було швидшим було обрано віртуально машину із GPU. За результатами тестів, найкращий результат мала нейромережа після 20 епох навчання із використанням dropout з ймовірністю 0.2.

Для оцінки нейромережі було обрано три метрики – логістична функція втрат, точність та середньоквадратична помилка. Отримані значення похибок на тестовому

датасеті близькі до значень на навчальному, отже можна зробити висновок що нейромережа дійсно навчилася аналізувати тональність коментарів

Отримана структура нейромережі та її параметри були збережені у файл для подальшого використання у системі оцінки відгуків користувачів в інтернет-магазинах.

Як було писано у другому розділі, у роботі було використано рекурентну нейромережу, а саме структуру LSTM. Створення та навчання нейромережі відбувалось відразу після завантаження та обробки датасету, у тому ж файлі в Jupyter. Для зручної роботи з нейромережами було обрано бібліотеку Torch. Вона має широкий функціонал та зрозумілу документацію. Цей фреймворк машинного навчання для мови Python має відкритий програмний код. У даний момент він активно розробляється групою штучного інтелекту в компанії Facebook. Також навколо цього фреймворка вибудована екосистема, що складається з різних бібліотек, що розробляються сторонніми командами. У своїй роботі Torch використовує тензорні обчислення, як і більшість бібліотек машинного навчання на Python.

Значною перевагою бібліотеки Torch є підтримка відеокарт (GPU) для навчання нейромереж. Саме це прискорює навчання у десятки разів, так як задача навчання нейромереж ідеально підходить для розпаралелювання на GPU. На даний момент підтримується лише програмно-апаратна архітектура CUDA, яка використовується на відеокартах від фірми Nvidia. В основі інтерфейсу програмування додатків CUDA лежить мова C з деякими розширеннями. Для успішної трансляції коду на цій мові до складу CUDA SDK входить власний компілятор для мови C. CUDA широко використовується в різних областях, особливу у машинному навчанні. Для перевірки того, чи підтримує бібліотека Torch відеокарту на певному пристрої існує спеціальна функція. Цю функцію потрібно викликати перед початком використання бібліотеки для навчання.

Для того щоб оцінити якість та похибку нейромережі використовуються різні функції оцінки, які називаються метриками. Їх існує багато, кожна метрика оцінює нейромережу з певної сторони, враховуючи певні критерії. У даній роботі для

оцінки нейромережі було використано три метрики: логістична функція втрат, точність та середньоквадратична помилка.

На вхід нейромережі буде приходити вектор у якому закодован текст. Також вона має п'ять виходів – по одному для кожної оцінки коментаря, від 0 до 4. Після декількох тестів було додано метод dropout, так як він підвищує точність нейромережі. Суть методу dropout полягає в тому, що в процесі навчання із загальної мережі випадковим чином виділяється підмережа, для якої здійснюється навчання. Після навчання обраної підмережі випадковим чином обирається нова підмережа і навчання продовжується. Вибір нейронів для підмережі відбувається випадковим чином, тобто кожен нейрон може бути виключеним з підмережі з певною ймовірністю. Ця ймовірність розраховується методом підбору. В даній задачі найкращий результат був при ймовірності виключення 0.2.

Для розв'язку задачі аналізу коментарів інтернет-магазинів було обрано досить відомий датасет. Він підходить під усі описані вище необхідні ознаки. Сам датасет складається з 700'000 відгуків, які були взяті з популярного сайту [yelp.com](https://www.yelp.com), а також оцінку кожного коментаря від 1 до 5.

Попередню обробку та саме навчання було проведено у Jupyter – інтерактивній оболонці для мови програмування Python. Jupyter розроблений на основі движка IPython. Він представляє додатковий командний синтаксис, має підсвічування коду і автоматичне доповнення. IPython дозволяє здійснювати неблокуючі взаємодії з графічним інтерфейсом у браузері. Також він може інтерактивно керувати паралельними обчисленнями використовуючи асинхронні зворотні виклики (callback). IPython може використовуватися як заміна стандартної командної оболонки операційної системи так як дозволяє отримати більшу гнучкість.

Для перевірки нейромережі було використано тестовий датасет. У попередньому пункті було описано процес його завантаження, та те що він містить 50'000 коментарів з оцінками. Ця частина датасету ще до цього жодного разу не подавалася на вхід до нейромережі, тому ідеально підходить для оцінки нейромережі.

Спочатку тестовий набір даних було токенізовано та перетворено на вектор. Після цього нормалізований вектор було подано до нейромережі на оцінку, а її відповідь була зіставлена із даними з датасету. У результаті було отримано точність у 71.5% на тестовому датасеті. Значення метрики логістична функція втрат та середньоквадратична помилка склали 0.885 та 0.811.

У четвертому розділі описується процес створення API та фронтенду для системи, зв'язок цих частин з нейромережою. Крім того, описується рішення для контейнеризації системи для її запуску на сервері. У цьому розділі були описані програмні засоби за допомогою яких була створена програмна реалізація. Описана діаграма прецедентів яка містить чотири прецеденти та описана структура БД яка містить чотири таблиці та описано зв'язок між ними. Було обрано СУБД – PostgreSQL.

Було описано процес скрапінгу коментарів із DOM-дерева за допомогою мови XPath та процес їх подальшої обробки нейромережою. Для використання нейромережі було створено API з використанням, який автоматично згенерував Swagger документацію до API. Було описано процес створення фронтенду за допомогою фреймворку Django на Python.

В останньому пункті було описано процес деплою з використанням технологій контейнеризації Docker та Docker Compose. Для деплою на сервер було використано Docker. Docker – програмне забезпечення з відкритим вихідним кодом, що застосовується для розробки, тестування, доставки і запуску веб-додатків в середовищах з підтримкою контейнеризації. Він потрібен для більш ефективного використання системи і ресурсів, швидкого розгортання готових програмних продуктів, а також для їх масштабування і перенесення в інші середовища з гарантованим збереженням стабільної роботи.

Для опису діаграм було використано уніфіковану мову моделювання UML. UML (Unified Modeling Language) – це система позначень, яку застосовують для об'єктно-орієнтованого аналізу і проектування систем. Її можна використовувати для візуалізації, специфікації, конструювання та документування програмних систем.

Основними перевагами використання UML є:

- можливість подивитися на завдання з різних точок зору;
- іншим програмістам легше зрозуміти суть завдання і спосіб її реалізації;
- діаграми порівняно прості для читання після досить швидкого ознайомлення з їх синтаксисом.

Під час розробки програмного забезпечення, було розроблено систему для скрапінгу коментарів для їх подальшої передачі у нейромережу. Розробка відбувалася на мові програмування Python. Це дозволило зручно інтегрувати нейромережу з основною системою, так як обидві частини написані на Python.

Для зручної роботи с нейромережою було створено API на Python за допомогою фреймворку FastAPI. FastAPI – це фреймворк для створення лаконічних і швидких HTTP API-серверів з вбудованими можливостями валідації, серіалізації та асинхронністю, із коробки. В його основі лежать два інших фреймворки: за роботу з інтернет-з'єднанням відповідає бібліотека Starlette, а за валідацію даних у API відповідає Pydantic. Цей фреймворк дозволяє швидко створити HTTP сервер який буде приймати як GET, так і POST запити.

Основний принцип роботи Docker – контейнеризація. Цей тип віртуалізації дозволяє упаковувати програмне забезпечення по ізольованим середам – контейнерів. Кожен з цих віртуальних блоків містить всі необхідні елементи для роботи програми. Це дає можливість одночасного запуску великої кількості контейнерів на одному хості.

Так як у система має два Dockerfile та залежить від БД PostgreSQL, то для зручного та швидкого деплою було використано технологію Docker Compose. Docker Compose використовується для одночасного управління декількома контейнерами, що входять до складу програми. Цей інструмент пропонує ті ж можливості, що і Docker, але дозволяє працювати з більш складними додатками.

Для деплою за допомогою Docker Compose було створено docker-compose.yml файл. Цей файл було скопійовано на сервер за допомогою консольної команди SCP та запущено за допомогою команди `docker-compose up --build -d`.

У розділі з охорони праці було описано безпечні і здорові умови праці на робочих місцях, в робочих зонах, у виробничих приміщеннях. Досягнуто це було за допомогою опрацювання питань умов праці, виробничої санітарії, техніки безпеки, пожежної безпеки. Виконано аналіз умов праці в серверному приміщенні закладу ресторанного типу. Було зроблено наступне:

1. Перевірено забезпечення вимог охорони праці.
2. Виявлено, що оцінка умов праці на робочому місці відноситься до IV категорії, коли спостерігається робота у несприятливих умовах праці.
3. Визначено категорії видів діяльності, та характеристика робочих місць, від якої також залежать допустимі величини параметрів мікроклімату робочого місця в виробничому приміщенні.
4. Після замірів фактичних величин параметрів мікроклімату виявлено, що в даному виробничому приміщенні температура повітря для працівників є нижче норми, тобто її необхідно збільшити, а швидкість руху повітря навпаки необхідно зменшити. З метою їх покращення розраховано припливно-витяжну вентиляцію повітропродуктивністю $V \geq 2500$ м³/год.
5. Підібрано вентилятор з необхідною витратно-напірною характеристикою.

ЗАГАЛЬНІ ВИСНОВКИ

За результатом виконання дипломної роботи було зроблено аналіз предметної області систем для аналізу тональності тексту, який показав що не існує універсального рішення, а більшість із існуючих не підходять для аналізу коментарів в інтернет-магазинах. Після цього було сформовано технічне завдання.

Наступним кроком був вибір, проектування та навчання нейромережі для аналізу тональності тексту. Було обрано LSTM нейромережу та датасет для її навчання та перевірки. Було описано три метрики для оцінки якості нейромережі, а датасет було проаналізовано та розбито на навчальну, валідаційну та текстову вибірку. Навчання нейромережі відбувалося на платформу Google Colab з використанням GPU.

У результаті нейромережа змогла оцінювати тональність коментаря по шкалі від 1 до 5, де чим вище оцінка – тим більш позитивний відгук і навпаки. Після навчання нейромережа досягла точності у 71.5% на тестовому датасеті, а середня квадратична помилка становила 0.811, що менше ніж 1 бал.

Було розроблено інформаційну систему, яка дозволяє збирати відгуки до товарів у певних інтернет-магазинах та аналізувати їх. Для цього було спроектовано БД з використанням СУБД PostgreSQL. Для зручного зв'язку з нейромережею для неї був створений API за допомогою фреймворку FastAPI. А для зручного використання усієї системи було створено фронтенд за допомогою фреймворку Django. Для зручного розгортання отриманої системи було створено два Dockerfile та Docker Compose файл для їх об'єднання.

За результатами спеціального розділу з охорони праці були створені безпечні і здорові умови праці на робочих місцях, в робочих зонах, у виробничих приміщеннях. Досягнуто це було за допомогою опрацювання питань умов праці, виробничої санітарії, техніки безпеки, пожежної безпеки.

АНОТАЦІЯ

Єрмолаєв Олександр Андрійович. Система визначення якості товарів в інтернет-магазинах на основі штучного інтелекту. – На правах рукопису.

Бакалаврська кваліфікаційна робота на здобуття освітньої кваліфікації «бакалавр з комп'ютерних наук» в галузі знань 12 «Інформаційні технології» за спеціальністю 122 «Комп'ютерні науки».

Чорноморський національний університет імені Петра Могили, Миколаїв.

Керівник: старший викладач Боровльова Світлана Юріївна.

Об'єкт – процеси організації визначення якості товарів в інтернет-магазинах.

Предмет – підходи та інформаційні технології до організації визначення якості товарів в інтернет-магазинах.

Мета – автоматизація процесу визначення якості товарів в інтернет-магазинах за рахунок розробки системи з використанням штучного інтелекту.

У першому розділі представлено огляд наявних засобів для аналізу тональності тексту, розроблено технічні та функціональні вимоги до застосунку. У другому розділі описано види нейромереж та обґрунтовано вибір саме рекурентної структури. У третьому розділі представлено опис процесу розробки та навчання нейромережі, а також обрано метрики оцінки її точності. В четвертому розділі було описано процес збору коментарів в інтернет-магазинах, описано основні моделі поведінки користувача та увесь процес створення програмного застосунку. В останньому розділі було розглянуто норми та заходи з охорони праці й техніки безпеки.

Сторінок – 87, таблиць – 14 , рисунків – 34, додатків – 3.

Ключові слова: інтернет-магазини, штучний інтелект, нейромережі, аналіз тексту, Python.

ABSTRACT

Iermolaiev Oleksandr. System for determining the quality of goods in online stores based on artificial intelligence. – On the rights of the manuscript.

Bachelor's qualification work for the educational qualification "Bachelor of Computer Science" in the field of knowledge 12 "Information Technology" in the specialty 122 "Computer Science".

Petro Mohyla Black Sea National University, Mykolaiv.

Supervisor: senior teacher Svetlana Borovleva.

The object is the processes of organizing the determination of product quality in online stores.

The subject is the approaches and information technologies for the organization of determining the quality of goods in online stores.

The goal is to automate the process of determining the quality of goods in online stores through the development of a system using artificial intelligence.

The first section provides an overview of the available tools for the text sentiment analysis, described technical and functional requirements. The second section describes the types of neural networks and explained the choice of recurrent structure. The third section presents a description of the process of development and training of the neural network, as well as selected metrics for determining its accuracy. The fourth section described the process of collecting comments in online stores, describes the basic patterns of user behavior and the whole process of creating a software application. In the last section the norms and measures on labor protection and safety were considered.

Pages – 87, tables – 14, figures – 34, appendices – 3.

Keywords: ecommerce, artificial intelligence, neural networks, text analysis, Python.