Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)
Кафедра Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження великих мовних моделей для покращених описів товарів

(тема)

Виконав:
здобувач другого року навчання,
групи СШІм-23-1
Станіслава Овчаренко
(власне ім'я, прізвище)
Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)
Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня програма Системи штучного інтелекту
(повна назва освітньої програми)
Керівник доц. Олексій Турута
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ _____ Олег ЗОЛОТУХІН
(підпис) (власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Штучного інтелекту

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки
(код і повна назва)

Тип програми освітньо-наукова
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи штучного інтелекту
(повна назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«_____» _____ 20 ___ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Овчаренко Станіславі Євгеніївні
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження великих мовних моделей для покращених описів товарів

затверджена наказом університету від 21 квітня 20 25 р. № 295Ст

2. Термін подання студентом роботи до екзаменаційної комісії 10 червня 20 25 р.

3. Вихідні дані до роботи Науково-технічні публікації, інтернет-ресурси, документація мовних моделей

4. Перелік питань, що потрібно опрацювати в роботі

1) Теоретичні основи великих мовних моделей та їх застосування у генерації текстів

2) Постановка задачі, огляд даних і методів оцінки

3) Вибір архітектури та методології генерації

4) Проведення експерименту

5) Аналіз результатів

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	21.04.2025	виконано
2	Аналіз актуальності теми та наукових джерел	23.04.2025 – 27.04.2025	виконано
3	Формулювання задачі дослідження та	28.04.2025 – 30.04.2025	виконано
	обґрунтування її практичної значущості		
4	Аналіз існуючих систем автоматичної генерації	01.05.2025 – 03.05.2025	виконано
	описів товарів		
5	Вибір мовних моделей та формування вимог до	04.05.2025 – 06.05.2025	виконано
	генеративної системи		
6	Підготовка вхідних даних і типових запитів	07.05.2025 – 10.05.2025	виконано
7	Проведення експерименту: генерація описів	11.05.2025 – 14.05.2025	виконано
	товарів за допомогою GPT-3.5 та GPT-4		
8	Оцінка якості згенерованих текстів за допомогою	15.05.2025 – 18.05.2025	виконано
	автоматичних метрик та експертної оцінки		
9	Аналіз результатів та формулювання практичних	19.05.2025 – 21.05.2025	виконано
	висновків і рекомендацій		
10	Оформлення пояснювальної записки згідно з	22.05.2025 – 03.05.2025	виконано
	методичними вказівками		
11	Підготовка презентації до захисту	04.05.2025 – 08.06.2025	виконано
12	Захист роботи перед екзаменаційною комісією	10.06.2025	

Дата видачі завдання 21 квітня 2025 р.

Здобувач _____
(підпис)

Керівник роботи _____ доц. Олексій Турута
(підпис) (посада, власне ім’я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 49 с., 4 табл., 1 дод., 28 джерел.

АВТОМАТИЗАЦІЯ ГЕНЕРАЦІЇ КОНТЕНТУ, ВЕЛИКА МОВНА МОДЕЛЬ, ГЕНЕРАЦІЯ ТЕКСТІВ, ЕЛЕКТРОННА КОМЕРЦІЯ, МЕТРИКИ ОЦІНКИ, ОБРОБКА ПРИРОДНОЇ МОВИ, ОПИС ТОВАРУ, ТЕКСТОВИЙ КОНТЕНТ.

Об’єкт дослідження: процес генерації текстових описів товарів у сфері електронної комерції. Предмет дослідження: використання великих мовних моделей для автоматизованої генерації описів товарів. Мета роботи: дослідження можливостей застосування великих мовних моделей GPT-3.5 і GPT-4 для створення якісного текстового контенту та порівняння їх ефективності в задачі генерації описів товарів. Методи дослідження: аналіз наукових джерел, порівняльний аналіз архітектур LLM, генерація текстів, оцінка результатів за допомогою автоматичних метрик та експертного аналізу.

Кваліфікаційна робота присвячена вивченню можливостей великих мовних моделей у задачі генерації текстових описів товарів для онлайн-торгівлі. У роботі сформовано вхідні дані на основі реальних товарів, проведено генерацію текстів і виконано оцінювання якості згенерованих описів за допомогою кількох метрик та експертної оцінки.

Отримані результати дозволили визначити переваги та недоліки кожної моделі в контексті автоматизації контенту для e-commerce, а також сформулювати практичні рекомендації щодо їх застосування. Робота може бути використана як основа для впровадження LLM у системи управління товарним контентом, а також для подальших досліджень у галузі генерації природної мови.

ABSTRACT

Master’s thesis contains: 49 pp., 4 tabl., 1 ann., 28 references.

CONTENT AUTOMATION, E-COMMERCE, EVALUATION METRICS, LARGE LANGUAGE MODEL, NATURAL LANGUAGE PROCESSING, PRODUCT DESCRIPTION, TEXT CONTENT, TEXT GENERATION.

Object of the study: the process of generating textual product descriptions in the field of e-commerce.

Subject of the study: the use of large language models for automated generation of product descriptions.

Purpose of the work: to explore the capabilities of large language models GPT-3.5 and GPT-4 in generating high-quality text content and to compare their effectiveness in the task of product description generation.

Research methods: analysis of scientific sources, comparative analysis of LLM architectures, text generation, evaluation of results using automatic metrics and expert analysis.

This qualification paper is dedicated to exploring the potential of large language models in the task of generating textual product descriptions for online commerce. The work includes the preparation of input data based on real products, the generation of texts using selected models, and the evaluation of the resulting descriptions using multiple quality metrics and expert review.

The obtained results made it possible to identify the strengths and weaknesses of each model in the context of content automation for e-commerce, and to formulate practical recommendations for their application. The work can serve as a basis for implementing LLMs in product content management systems and for further research in the field of natural language generation.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів 7
Вступ 8
1 Теоретичні основи великих мовних моделей та їх застосування у генерації текстів 10
1.1 Загальна характеристика великих мовних моделей 10
1.2 Еволюція мовних моделей: від традиційних алгоритмів до LLM 11
1.3 Сфера застосування LLM у генерації описів товарів 13
1.4 Огляд доступних датасетів 15
1.5 Проблеми і виклики автоматичної генерації контенту 16
2 Постановка задачі, огляд даних і методів оцінки 19
2.1 Формалізація задачі генерації описів товарів 19
2.2 Вимоги до генеративної системи 20
2.3 Аналіз існуючих автоматизованих систем генерації описів 21
2.4 Формат і приклади вхідних/вихідних даних 23
2.5 Метрики оцінки якості згенерованих описів 26
2.6 Проблеми та виклики автоматичної генерації контенту 28
3 Вибір архітектури та методології генерації 30
3.1 Вибір мовної моделі для дослідження 30
3.2 Збір та підготовка даних (описів товарів) 32
4 Проведення експерименту 36
4.1 Генерація описів 36
4.2 Оцінка результатів 38
5 Аналіз результатів 42
5.1 Висновки з експерименту 42
Висновки 44
Перелік джерел посилання 46
Додаток А Відомість кваліфікаційної роботи 49

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

Галюцинація моделі – випадок генерації неправдивої або вигаданої інформації, яка виглядає правдоподібною;

Промпт – короткий текстовий запит, сформульований для запуску генерації тексту мовною моделлю;

AI – Artificial Intelligence – штучний інтелект;

API – Application Programming Interface – програмний інтерфейс прикладного програмування;

BERTScore – Bidirectional Encoder Representations from Transformers Score – метрика семантичної схожості текстів на основі BERT;

BLEU – Bilingual Evaluation Understudy – метрика оцінки тексту за n-грамами;

CMS – Content Management System – система управління контентом;

Fine-tuning – уточнююче донавчання моделі на спеціалізованих даних;

GPT – Generative Pre-trained Transformer – генеративна попередньо навчена трансформерна модель;

HTML – HyperText Markup Language – мова гіпертекстової розмітки;

JSON – JavaScript Object Notation – текстовий формат обміну даними;

LLM – Large Language Model – велика мовна модель;

METEOR – Metric for Evaluation of Translation with Explicit ORdering – метрика для порівняння текстів із врахуванням синонімів та порядку слів;

NLP – Natural Language Processing – обробка природної мови;

ROUGE – Recall-Oriented Understudy for Gisting Evaluation – метрика оцінки релевантності тексту;

SEO – Search Engine Optimization – пошукова оптимізація;

XML – eXtensible Markup Language – розширювана мова розмітки;

ВСТУП

Розвиток електронної комерції істотно трансформував сучасний ринок товарів і послуг, створивши нові формати взаємодії між продавцями та покупцями. Однією з ключових особливостей онлайн-торгівлі є відсутність фізичної взаємодії з товаром, що значно підвищує роль текстового опису. Якісний опис товару забезпечує не лише інформативність, але й формує у споживача довіру до продукту, впливає на його емоційне сприйняття, підвищує конверсію та сприяє формуванню позитивного іміджу бренду. В умовах високої конкуренції на ринку електронної комерції саме контент, зокрема тексти описів, стає одним із ключових інструментів маркетингу та просування.

Проте створення якісних текстових описів для кожної товарної позиції є складним і трудомістким завданням, особливо для великих інтернет-магазинів і маркетплейсів, де кількість товарів може сягати десятків або навіть сотень тисяч. Ручне написання описів вимагає значних часових та людських ресурсів, що робить процес економічно невигідним. У зв’язку з цим зростає попит на автоматизовані рішення, які здатні забезпечити швидке та масштабоване створення релевантного текстового контенту.

У цьому контексті особливої уваги набувають великі мовні моделі (Large Language Models, LLM), які ґрунтуються на сучасних досягненнях у галузі штучного інтелекту. Такі моделі, як GPT-3.5 та GPT-4, здатні генерувати зв’язні, граматично правильні, стилістично вивірені та адаптовані до конкретного запиту тексти. Вони вже демонструють високу ефективність у різних задачах генерації природної мови, таких як чат-боти, написання статей, автоматизована підтримка користувачів тощо. Проте їх застосування у сфері генерації описів товарів потребує додаткового дослідження, яке б враховувало специфіку контенту для електронної

комерції: точність, релевантність, привабливість і відповідність маркетинговим цілям.

Метою цієї кваліфікаційної роботи є дослідження можливостей використання великих мовних моделей для автоматизованої генерації описів товарів, а також порівняння ефективності моделей GPT-3.5 та GPT-4 у виконанні цього завдання.

Об’єктом дослідження є процес генерації текстового опису товару у сфері електронної комерції.

Предметом дослідження – застосування великих мовних моделей для створення якісного маркетингового контенту.

Практична частина роботи передбачає формування набору тестових запитів на основі реальних товарів, генерацію описів за допомогою моделей GPT-3.5 та GPT-4, а також оцінку отриманих результатів за допомогою автоматичних метрик (таких як BLEU, ROUGE, METEOR, BERTScore) і експертного аналізу.

Результати дослідження дозволяють зробити висновки щодо доцільності впровадження LLM у процес створення товарного контенту, визначити переваги та недоліки конкретних моделей, а також сформулювати практичні рекомендації для бізнесу. Практична цінність роботи полягає у створенні методичних підходів до порівняльного аналізу мовних моделей і формуванні рекомендацій щодо їх використання в реальних умовах онлайн-торгівлі.

1 ТЕОРЕТИЧНІ ОСНОВИ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ ТА ЇХ ЗАСТОСУВАННЯ У ГЕНЕРАЦІЇ ТЕКСТІВ

1.1 Загальна характеристика великих мовних моделей

Великі мовні моделі (LLM – Large Language Models) є потужним інструментом штучного інтелекту, що дозволяє автоматизувати широкий спектр завдань, пов’язаних з обробкою природної мови. Основною особливістю таких моделей є їх здатність навчатися на величезних обсягах текстових даних для подальшої генерації змістовних, логічно зв’язних і стилістично коректних текстів. Принцип їхньої роботи полягає в прогнозуванні наступного слова або послідовності слів на основі заданого контексту, що забезпечується глибинним статистичним аналізом і багаторівневою обробкою інформації.

LM можуть працювати з мільярдами параметрів, що дозволяє їм досягати високої точності при генерації текстів, перекладі, резюмуванні, класифікації, відповіді на питання та інших завданнях. Серед найвідоміших представників LLM – моделі сімейства GPT (Generative Pre-trained Transformer), BERT (Bidirectional Encoder Representations from Transformers) і T5 (Text-to-Text Transfer Transformer), кожна з яких має унікальну архітектуру та застосовується в різних сферах.

LLM стали ключовими технологіями у сфері обробки природної мови (NLP – Natural Language Processing), витіснивши попередні підходи завдяки здатності працювати з контекстом на глибшому рівні. Вони успішно застосовуються в чат-ботах, системах машинного перекладу, голосових помічниках, автоматизованих клієнтських службах, системах резюмування, пошукових механізмах та генерації текстів.

Особливо помітним є вплив LLM у сфері електронної комерції, де виникла потреба у масштабній автоматизації створення описів товарів. Кожна товарна позиція на онлайн-платформі вимагає унікального,

привабливого й інформативного тексту. Великі мовні моделі здатні генерувати такі описи з урахуванням ключових характеристик товару, маркетингового тону бренду та цільової аудиторії, що значно підвищує ефективність цифрових продажів.

Фундаментом великих мовних моделей є архітектура трансформерів, яка була запропонована дослідниками Google у 2017 році в роботі Attention is All You Need. Основний елемент трансформерів – механізм багатоголовкової уваги (multi-head self-attention), що дозволяє моделі фокусуватися на різних частинах тексту одночасно, незалежно від їх віддаленості. Це надає можливість враховувати складні мовні залежності та забезпечувати кращу когерентність тексту.

На відміну від попередніх архітектур, таких як RNN чи LSTM, трансформери здатні паралельно обробляти великі обсяги даних і зберігати довгостроковий контекст, що є критично важливим для створення зв’язного тексту. Саме ця архітектура дала змогу перейти від вузькоспеціалізованих до універсальних мовних моделей, які можна ефективно адаптувати до нових завдань за допомогою донавчання (fine-tuning).

1.2 Еволюція мовних моделей: від традиційних алгоритмів до LLM

Еволюція мовних моделей у сфері обробки природної мови (NLP) пройшла складний шлях – від простих статистичних алгоритмів до сучасних глибоких нейронних мереж із мільярдами параметрів. Перші підходи базувалися на статистичних моделях, таких як n-грамні моделі, які використовували частотність слів для побудови ймовірнісних залежностей між словами. Такі моделі мали суттєві обмеження: вони не враховували глибокий контекст, були схильні до перенавчання і не могли обробляти довгі послідовності слів з високою точністю.

Подальшим кроком стали моделі, засновані на нейронних мережах. Зокрема, рекурентні нейронні мережі (RNN) та їхні модифікації –

LSTM (Long Short-Term Memory) і GRU (Gated Recurrent Unit) – дозволили враховувати порядок слів у реченні та зберігати залежності між ними на коротких ділянках тексту. Ці архітектури значно покращили якість обробки послідовностей, проте залишалися чутливими до втрати контексту на великих відстанях у тексті, працювали повільно і не масштабувалися до великих корпусів даних.

Революційним проривом стало впровадження трансформерної архітектури у 2017 році, яка поклала початок новій епосі мовного моделювання. В основі трансформера лежить механізм самоуваги (self-attention), що дозволяє моделі ефективно визначати взаємозв'язки між будь-якими елементами тексту незалежно від їх позиції. Це забезпечує можливість моделі «бачити» весь текст одночасно, що особливо важливо для генерації зв'язного і логічного контенту. Крім того, трансформери підтримують паралельну обробку послідовностей, що значно прискорює навчання та інференцію.

Після появи трансформерів було створено низку високоефективних моделей, таких як BERT, GPT, T5, LLaMA, PaLM та інші, які сформували ядро сучасного NLP. Ці моделі вже не просто аналізують текст, а й здатні до генерації, перекладу, класифікації, резюмування, відповіді на запитання, виявлення емоцій тощо.

Поява великих мовних моделей вивела обробку природної мови на новий рівень – рівень семантичного розуміння та контекстної релевантності. Вони можуть враховувати мовні закономірності, стилістику, специфіку аудиторії й навіть особливості формулювання запитів. Цей підхід дав можливість створити універсальні мовні інструменти, які легко адаптуються до прикладних задач – зокрема, генерації описів товарів, ведення діалогів у чат-ботах, підтримки користувачів або персоналізації маркетингових повідомлень.

Таким чином, еволюція мовних моделей – це перехід від формальних статистичних правил до гнучкого, навчено-контекстного розуміння мови,

що стало можливим завдяки глибокому навчанню, обчислювальним потужностям і величезним обсягам даних. Великі мовні моделі сьогодні є не просто алгоритмами – це повноцінні інструменти інтелектуального аналізу та створення текстового контенту.

1.3 Сфера застосування LLM у генерації описів товарів

У сучасній електронній комерції текстові описи товарів відіграють критичну роль у процесі взаємодії споживача з онлайн-платформою. На відміну від традиційної роздрібної торгівлі, де покупець має змогу фізично оглянути продукт, онлайн-шопінг повністю покладається на візуальний і текстовий контент. Саме тому якісно сформульований опис має вирішальний вплив на сприйняття товару, довіру до бренду та рішення про покупку.

Великі мовні моделі відкривають нові можливості у створенні таких описів. Вони здатні автоматично генерувати тексти, які не лише інформують, а й емоційно впливають на читача. Ефективність тексту полягає не тільки в його змістовності, а й у здатності переконати, зняти сумніви, викликати довіру та стимулювати до покупки. У цьому контексті LLM можуть забезпечити релевантний, граматично правильний, структурований і стилістично узгоджений контент, адаптований під конкретну цільову аудиторію.

Маркетингові дослідження підтверджують, що понад 80% покупців звертають особливу увагу на опис продукту під час ухвалення рішення про покупку. Добре написаний опис здатен підкреслити ключові переваги, викликати бажані емоції (довіру, впевненість, захоплення), а також впливати на підсвідоме сприйняття продукту. LLM дають змогу реалізувати ці функції у масштабі, недоступному для ручного написання, зберігаючи при цьому якість і логіку тексту.

Серед ключових сфер застосування великих мовних моделей у генерації описів товарів варто виділити автоматизоване створення SEO-оптимізованих описів, персоналізованих повідомлень, варіативного контенту для A/B-тестування, а також локалізованих текстів для різних мов і ринків. Завдяки сучасним інструментам на базі GPT, T5, Claude та інших моделей, генерація описів відбувається у реальному часі на основі ключових характеристик товару, брендових настанов і уподобань користувача.

Сьогодні чимало компаній уже інтегрували LLM у свої бізнес-процеси. Такі компанії, як Amazon, Alibaba, Nestlé та Zalando, застосовують генеративний ШІ для створення описів товарів, адаптованих під різні ринки, мови та стилістичні вимоги. Це дозволяє зменшити витрати на ручну працю, пришвидшити процес публікації нових товарів, забезпечити послідовність у комунікації з клієнтами та досягти високої якості текстів навіть при масовому виробництві контенту.

Окрім створення описів для вебсторінок товарів, LLM дедалі активніше застосовуються для інших форматів контенту, пов’язаного з описами продукції:

– SEO-описи для пошукових систем, які включають ключові слова з високою частотою запиту, що дозволяє підвищити видимість сторінки в Google, Bing тощо;

– персоналізовані короткі описи для електронної розсилки або push-сповіщень, які враховують поведінкові патерни користувача (історія покупок, переглядів);

– A/B варіанти текстів, що дозволяє протестувати різні стилістичні підходи (емоційний vs інформативний стиль) і визначити той, який забезпечує вищий CTR або конверсію.

Бренди також все частіше адаптують текстові описи для різних аудиторій і регіонів: наприклад, молодіжна аудиторія може отримувати описи у стилі «лайфстайл», а бізнес-аудиторія – лаконічні й технічні.

Використовуючи можливості LLM, компанії можуть масштабувати подібну персоналізацію без залучення додаткового людського ресурсу.

Таким чином, великі мовні моделі поступово змінюють підходи до створення описів у сфері електронної комерції, перетворюючи цей процес із трудомісткого ручного завдання на автоматизований і масштабований інструмент підвищення ефективності бізнесу.

1.4 Огляд доступних датасетів

Ефективність великих мовних моделей безпосередньо залежить від якості та обсягу даних, на яких вони тренуються або донавчаються. У контексті генерації описів товарів особливе значення має наявність релевантних датасетів, що містять приклади добре сформованих описів, характеристик продукції, категорій, а також супровідної інформації – наприклад, назви, бренди, ціни та зображення.

Серед доступних відкритих датасетів варто згадати кілька найбільш уживаних у сфері електронної комерції та генерації описів. Один з найпопулярніших – Amazon Product Data, який містить мільйони записів про товари, включно з заголовками, короткими й повними описами, категоріями та відгуками користувачів. Іншим відомим джерелом є Open Product Data або Product Matching Data, які фокусуються на зіставленні товарів за різними атрибутами. Також дослідники часто використовують власноруч сформовані корпуси, зібрані шляхом парсингу інформації з відкритих онлайн-магазинів або маркетплейсів.

У разі відсутності придатного готового датасету або коли потрібно адаптувати систему під конкретну нішу, доцільно використовувати підхід із побудови власної колекції даних. Наприклад, можна зібрати описи з корпоративного сайту компанії, структуровано витягти характеристики товарів або створити анотації вручну за допомогою експертів. Такий підхід

забезпечує високу якість, повну відповідність потребам бізнесу та дозволяє уникнути шуму й нерелевантної інформації.

Окрім текстових описів, сучасні системи генерації контенту можуть використовувати мультимодальні датасети, які містять візуальні дані (зображення продуктів) або метаінформацію (категорія, популярність, відгуки), що дозволяє створювати ще більш персоналізовані та змістовні тексти.

Вибір або побудова датасету є ключовим етапом у налаштуванні мовної моделі. Від цього залежить не лише якість згенерованого контенту, а й здатність моделі враховувати специфіку категорій товарів, мовний стиль цільової аудиторії та формат платформи, на якій буде використовуватись текст.

1.5 Проблеми і виклики автоматичної генерації контенту

Попри значний прогрес у розвитку великих мовних моделей та їхню здатність до високоякісної генерації тексту, автоматизовані системи створення контенту стикаються з низкою викликів. Однією з головних проблем є забезпечення якості, точності та достовірності створених текстів. LLM здатні генерувати переконливі на вигляд речення, але іноді можуть включати вигадані факти або неточності, що особливо критично для описів товарів у галузях, де потрібна абсолютна достовірність (наприклад, медицина, побутова техніка чи харчові продукти).

Такий ефект отримав назву «галюцинації моделі». GPT-3.5 може приписати електрочайнику функцію автоприготування кави або «розумне очищення», яких насправді не існує. У випадку з косметикою модель може вигадати, що крем має сертифікати FDA або містить гіпоалергенні компоненти, не маючи підтвердження. Це може спричинити недовіру з боку покупців і навіть юридичні наслідки для компанії.

Ще одним викликом є адаптація згенерованого тексту до специфіки різних платформ електронної комерції. Кожен маркетплейс має свої вимоги до структури, обсягу, формулювання і стилю тексту, а також до використання ключових слів та форматування. Наприклад, Amazon та Etsy мають обмеження щодо довжини заголовків, ключових слів і характеру описів, що накладає додаткові вимоги до точності промптів і постобробки згенерованих матеріалів.

Контекстуальне розуміння залишається ще однією слабкою стороною сучасних моделей. Хоча вони здатні аналізувати широке коло вхідних даних, іноді їм бракує «розуміння» реального призначення товару або очікувань цільової аудиторії, що може призвести до невідповідності змісту. Також є проблема узгодженості термінів: модель може у межах одного опису використовувати несумісні чи суперечливі твердження, якщо її не обмежити відповідними параметрами.

Юридичні та етичні аспекти автоматичної генерації контенту також не варто ігнорувати. Генеровані тексти повинні бути вільними від дискримінаційних формулювань, неправдивої реклами, надмірних обіцянок або токсичного контенту. Контроль з боку людини тут є обов’язковим, особливо якщо мова йде про публікацію на офіційних платформах або в умовах жорсткого регулювання.

Нарешті, виникає питання масштабованості та оновлюваності контенту. Навіть якщо модель спочатку генерує релевантний і якісний текст, змінені умови ринку, нові версії товарів або зміни у вимогах до SEO можуть вимагати оновлення описів. Автоматична система має або постійно перестворювати описи з урахуванням нових даних, або інтегруватися в систему управління контентом, щоб забезпечити безперервну актуальність інформації.

Ще одним критичним викликом є ризик стилістичної одноманітності текстів – LLM, зокрема GPT-3.5, можуть створювати тексти, що звучать надто схоже між собою, з повторюваними структурами речень або фразами

на кшталт «ідеальний вибір для...». Це знижує ефективність контенту в очах користувачів, зменшує емоційне залучення та може викликати втому від «однотипності» у випадку перегляду кількох товарів поспіль.

Ця стилістична одноманітність не лише втомлює користувача, а й негативно впливає на SEO-показники через повторюваний контент. Крім того, масова генерація шаблонних описів знижує впізнаваність бренду та створює враження «знеособленого» товару. Для подолання цієї проблеми варто або налаштовувати систему через промпт-інженерію, або використовувати fine-tuning на прикладах з брендованою стилістикою.

Крім того, LLM не мають фактичного розуміння змісту – вони формують текст на основі ймовірнісних шаблонів, що іноді призводить до випадкового включення вигаданих фактів. Наприклад, модель може приписати товару функцію, яка згадується в подібних продуктах, але відсутня в конкретному випадку. Це створює ризики для довіри до бренду та юридичні наслідки.

Існує також ризик порушення корпоративного стилю, коли модель створює текст, що не відповідає tone of voice компанії. Наприклад, серйозний бренд може отримати надмірно легковажний або жартівливий опис.

Таким чином, попри високий потенціал автоматизованої генерації описів товарів, її впровадження супроводжується низкою технічних, методологічних і практичних викликів, що потребують уважного підходу, налаштування моделей під конкретні задачі та обов’язкового людського контролю.

2 ПОСТАНОВКА ЗАДАЧІ, ОГЛЯД ДАНИХ І МЕТОДІВ ОЦІНКИ

2.1 Формалізація задачі генерації описів товарів

Завдання генерації описів товарів полягає в автоматичному створенні зв’язного, інформативного та привабливого тексту на основі обмеженого обсягу структурованих або напівструктурованих вхідних даних, таких як назва товару, його характеристики, категорія чи цільова аудиторія. Ця задача розглядається як окремий випадок текстогенерації (text generation) у рамках обробки природної мови (NLP), де основною метою є створення тексту, який відповідає очікуванням користувача, вимогам платформи й водночас підвищує ефективність взаємодії з покупцем.

Формально, генерація опису товару може бути подана як функція:

$$G(x) \rightarrow y, \tag{2.1}$$

де x – це множина вхідних параметрів товару (наприклад, назва, категорія, бренд, розміри, функції);

y – згенерований текстовий опис, що має відповідати критеріям змістовності, стилістичної узгодженості, релевантності та комерційної привабливості.

Особливістю задачі є вимога до адаптивності: моделі повинні не лише описати об’єктивні характеристики товару, а й транслювати цінності бренду, підкреслювати вигоди для споживача, відповідати тону спілкування цільової аудиторії. Це вимагає не лише технічної точності, а й творчого підходу, який традиційно був притаманний людським копірайтерам.

Ще однією складовою формалізації є врахування бізнес-контексту. У реальних застосуваннях задача генерації описів часто пов’язана з масштабованістю(генерація тисяч описів щоденно), необхідністю

дотримання обмежень платформи (кількість символів, структура, ключові слова) та інтеграцією з внутрішніми базами даних компанії.

У зв’язку з цим задача генерації описів товарів потребує таких рішень, які поєднують потужність великих мовних моделей зі здатністю до налаштування під конкретні цілі, стилі та обмеження. Тому сучасні підходи орієнтуються на створення гнучких генеративних систем, які поєднують архітектури LLM з компонентами тонкого налаштування (fine-tuning), промпт-інженерії та експертного контролю якості.

2.2 Вимоги до генеративної системи

Для успішної генерації описів товарів система повинна відповідати низці функціональних, лінгвістичних та комерційних вимог. Насамперед, вона має забезпечувати змістову точність, тобто правильно інтерпретувати вхідні дані та формувати текст, що відображає реальні характеристики товару без викривлень або вигадок. Це критично важливо для збереження довіри споживача і відповідності опису до фактичної пропозиції.

Другою важливою вимогою є читабельність і стилістична природність тексту. Згенеровані описи повинні бути граматично правильними, логічно зв’язними та такими, що відповідають очікуванням аудиторії. У цьому контексті значну роль відіграє здатність системи адаптувати стиль, тональність і рівень формальності відповідно до бренду, типу продукту та цільового сегмента споживачів.

Окремим блоком є функціональні вимоги до масштабованості та швидкодії. У великих інтернет-магазинах або на маркетплейсах генерація описів для тисяч або десятків тисяч товарів повинна відбуватися автоматично, без значного людського втручання. Це означає, що система має бути здатною обробляти великі обсяги даних у реальному або близькому до реального часу.

Ще один аспект – адаптивність до різних платформ і форматів. Кожен маркетплейс або CMS (система управління контентом) має свої обмеження щодо довжини заголовків, структури опису, обов’язкових полів тощо. Генеративна система повинна враховувати ці параметри або давати змогу налаштовувати вихідний шаблон відповідно до вимог цільової платформи.

Важливим є також питання інтеграції з іншими бізнес-системами – наприклад, із базами товарів, CRM, SEO-інструментами або системами аналітики. Це дозволяє автоматично отримувати вхідні дані та повертати результати у потрібному форматі, що суттєво спрощує роботу контент-команд.

Нарешті, не менш важливою є можливість контролю якості, що включає перевірку на відповідність брендовим гайдлайнам, відсутність токсичного чи некоректного контенту, уникнення повторів і шаблонності. Для цього системи часто доповнюються фільтрами, вбудованими правилами або модулем попереднього перегляду з редагуванням.

Таким чином, сучасна генеративна система повинна бути не просто потужним мовним модулем, а гнучким і масштабованим інструментом, що адаптується до бізнес-потреб, технічних обмежень і очікувань користувача.

2.3 Аналіз існуючих автоматизованих систем генерації описів

Сучасний ринок автоматизованих систем генерації описів активно розвивається завдяки поширенню великих мовних моделей. Багато компаній інтегрують генеративні модулі у свої платформи для спрощення створення комерційного контенту. Такі системи допомагають зменшити витрати на ручне написання, підвищити ефективність створення текстів і забезпечити масштабованість.

Серед найпопулярніших рішень можна виділити Jasper AI, Copy.ai, Hypotenuse AI, Writesonic і Shopify Magic. Jasper AI – одна з найпотужніших платформ, що базується на GPT і пропонує широкий набір шаблонів для

генерації маркетингових текстів, а також підтримує навчання на власних прикладах і інтеграцію з SEO-інструментами. Copy.ai орієнтований на простоту використання та швидку генерацію коротких рекламних текстів. Його інтерфейс зручний для малого бізнесу, а результат адаптований до побутових завдань.

Hypotenuse AI є спеціалізованим інструментом для e-commerce, який дозволяє створювати тисячі описів на основі табличних даних (наприклад, з CSV-файлів) або API. Writesonic – універсальна платформа, яка підтримує генерацію описів, заголовків, рекламних блоків і SEO-контенту. Shopify Magic – вбудоване рішення для власників магазинів на Shopify, яке інтегрується безпосередньо у CMS і дозволяє створювати або редагувати опис товару без залучення сторонніх сервісів.

Ефективність таких систем оцінюється за кількома ключовими параметрами. Перш за все, важливою є якість тексту – наскільки він читабельний, граматично правильний, логічно зв’язний і стилістично відповідний. Інший важливий критерій – персоналізація: чи враховує система особливості конкретного товару, категорії або бренду. Велике значення має також масштабованість: наскільки ефективно система обробляє великі обсяги даних і наскільки швидко може генерувати сотні або тисячі описів.

SEO-оптимізація є ще одним критичним фактором – система повинна генерувати тексти, що відповідають вимогам пошукових алгоритмів (ключові слова, структура, довжина речень). Нарешті, слід оцінювати інтеграційні можливості: чи підтримує система API, чи сумісна з CMS, чи можна підключити її до маркетплейсу без складної адаптації.

Загалом, більшість сучасних платформ демонструють високу функціональність і дають змогу ефективно автоматизувати процес створення описів. Однак ідеального рішення не існує: іноді моделі повторюють одні й ті самі формулювання, недостатньо враховують контекст, або не до кінця відповідають стилістичним очікуванням. Тому в

багатьох випадках усе ще потрібен людський контроль і постредагування результатів.

2.4 Формат і приклади вхідних/вихідних даних

Формат вхідних та вихідних даних є важливою складовою при створенні систем автоматичної генерації описів товарів. Від структури та якості вхідної інформації залежить не лише результативність генерації, а й ступінь релевантності та стилістичної відповідності створених текстів.

Вхідні дані зазвичай включають структуровану інформацію про товар, яка подається у вигляді табличних або JSON-форматів. Це можуть бути назва продукту, категорія, технічні характеристики, колір, розмір, матеріал, бренд, ціновий діапазон, цільова аудиторія, призначення, ключові переваги тощо.

У деяких випадках додаються додаткові маркери, які вказують на бажаний стиль опису (нейтральний, емоційний, технічний), цільову мову або формат (короткий опис, розгорнутий, SEO-оптимізований).

Наприклад, вхідний шаблон може виглядати як лістинг 2.1.

Лістинг 2.1 – Вхідний шаблон

```
json
{
  "назва": "Термокружка SteelPro 500 мл",
  "матеріал": "нержавіюча сталь",
  "особливості": "зберігає тепло до 12 годин",
  "колір": "чорний",
  "категорія": "туризм",
  "аудиторія": "любителі активного відпочинку",
  "стиль": "енергійний, емоційний"
}
```

У практиці бізнесу вихідні тексти часто потрібні у вигляді HTML або XML-структур, адаптованих до CMS або e-commerce платформ. Наприклад, опис може генеруватись як наведено у лістингу 2.2.

Лістинг 2.2 – Генерація опису

```
<product>
  <title>Блендер Bosch ErgoMixx 600 W</title>
  <short_description>Надійний помічник для щоденної кулінарії</short_description>
  <features>
    <feature>Потужність 600 Вт</feature>
    <feature>12 швидкостей</feature>
    <feature>Ергономічний корпус</feature>
  </features>
  <full_description>Компактний і зручний, цей блендер легко справляється з приготуванням смузі, соусів та супів.</full_description>
</product>
```

Або у форматі для Shopify / XML фіда (лістинг 2.3).

Лістинг 2.3 – Генерація у форматі для Shopify / XML фіда

```
<item>
  <title>Dyson V11 Absolute</title>
  <description><![CDATA[Потужний бездротовий пилосос з LCD-дисплеєм і адаптивним регулюванням потужності.]]></description>
  <category>Пилососи</category>
  <price>18999.00</price>
</item>
```

На основі такої структури генеративна модель може створити наступний вихідний текст: «Термокружка SteelPro – незамінний супутник у кожній мандрівці! Виготовлена з міцної нержавіючої сталі, вона зберігає

тепло до 12 годин, дозволяючи насолоджуватися гарячими напоями навіть у найхолоднішу погоду. Лаконічний чорний дизайн ідеально пасує до спорядження справжнього мандрівника!».

У залежності від потреб компанії, вхідні дані можуть мати різний ступінь деталізації. Наприклад, для базових генерацій достатньо ключових характеристик товару, але для створення унікальних і точних описів доцільно включати додаткову інформацію:

– позиціювання продукту на ринку (бюджетний/преміум);
– переваги над конкурентами;
– умови використання (вдома, у подорожах, на виробництві);
– частотні запити користувачів із пошуку.

Опис може створюватися не лише як суцільний текст, а й у вигляді структурованого блоку з підзаголовками, списками переваг або короткими булітами. Наприклад.

Заголовок: Ручний блендер Bosch ErgoMixx.

Короткий опис: Компактний помічник для щоденних кулінарних завдань.

Переваги:

– потужність 600 Вт та 12 швидкостей;
– ніж із нержавіючої сталі;
– турборежим для складних інгредієнтів.

Ідеально для: домашнього приготування смузі, супів та соусів.

Подібний формат часто вимагається CMS системами або маркетплейсами (наприклад, Rozetka, Prom.ua, Amazon), де опис структурується на окремі поля: заголовок, summary, технічні характеристики, повний опис, call-to-action тощо.

Таким чином, гнучкість генеративної моделі полягає не лише у здатності створювати текст, а й у її адаптивності до форматних вимог платформи, що є критичним для інтеграції в бізнес-процеси.

Ще одним важливим аспектом є локалізація описів, тобто здатність моделі адаптувати контент до мовних і культурних особливостей різних ринків. Наприклад, опис, створений для українського сегменту, повинен враховувати місцеву лексику (наприклад, «пилосос» замість «пилевсмоктувач») і ментальні асоціації. У разі перекладу на англійську, важливо уникати буквальних трансформацій та зберігати маркетинговий зміст. Моделі GPT підтримують мультимовну генерацію, однак для високої якості локалізації бажано створювати або тестувати промпти окремо для кожної мови.

Вихідні дані, як правило, формуються у вигляді повноцінного текстового блоку з дотриманням вимог щодо довжини, структури, наявності заголовків, підзаголовків чи навіть HTML-розмітки – залежно від каналу публікації (вебсайт, маркетплейс, e-mail тощо).

Таким чином, генерація описів вимагає як якісної підготовки вхідних даних, так і гнучкої адаптації вихідного результату до різних форматів представлення. Це дозволяє інтегрувати модель у робочий процес електронної комерції з мінімальними зусиллями з боку користувача.

2.5 Метрики оцінки якості згенерованих описів

Для оцінки якості текстів, створених за допомогою великих мовних моделей, використовуються як автоматизовані метрики, так і експертні (людські) підходи. Серед найбільш поширених автоматичних метрик варто виділити Perplexity – вона показує, наскільки впевнено модель передбачає наступне слово у тексті. Чим менше значення perplexity, тим вищою вважається якість тексту, оскільки модель демонструє кращу здатність «розуміти» контекст.

Іншою популярною метрикою є BLEU (Bilingual Evaluation Understudy), яка застосовується для порівняння згенерованих фрагментів із референсними текстами. Вона аналізує збіги за n-грамами, тобто

послідовностями з кількох слів, і дозволяє оцінити точність моделі в плані формального відтворення змісту.

Метрика ROUGE (Recall-Oriented Understudy for Gisting Evaluation) частіше використовується у задачах резюмування та генерації описів. На відміну від BLEU, вона більше зосереджена на згадуваності, тобто на тому, скільки важливих слів і фраз із еталонного тексту було «впіймано» в згенерованому варіанті. Особливу увагу приділяють варіантам ROUGE-N (n-грам) та ROUGE-L (довгі спільні підпослідовності).

Більш гнучкою метрикою є METEOR, яка не тільки оцінює точність і згадуваність, але й враховує синонімію, граматичні варіації та порядок слів. Це дозволяє моделі отримати вищі оцінки у випадках, коли згенерований текст є семантично релевантним, навіть якщо його формулювання відрізняється від еталону.

Особливе місце займає BERTScore, що використовує векторні представлення слів, згенеровані за допомогою попередньо натренованих трансформерів (наприклад, BERT). BERTScore порівнює семантичну близькість між словами у вихідному та референсному текстах, розраховуючи косинусну подібність між векторними уявленнями. Ця метрика краще відображає змістову відповідність, ніж метрики, що базуються на точному збігу слів.

Попри зручність автоматичних метрик, найбільш повноцінну оцінку якості забезпечує експертна (людська) оцінка. Вона враховує такі аспекти, як зв’язність тексту, логіка викладення, природність мовлення, відповідність тону й стилю бренду, а також емоційний вплив на читача. У практиці зазвичай використовуються шкали від 1 до 5 або від 1 до 10, де експерти оцінюють тексти за декількома критеріями. Такий підхід дає змогу виявити недоліки, які можуть залишитися непоміченими автоматичними алгоритмами, й забезпечує найбільш релевантну оцінку з точки зору кінцевого користувача.

2.6 Проблеми та виклики автоматичної генерації контенту

Попри численні переваги, пов’язані з використанням великих мовних моделей у генерації текстів, ці технології все ще мають низку обмежень і викликів, які потребують уважного аналізу та врахування при впровадженні в практичні системи.

Одним із найважливіших недоліків є обмежене розуміння контексту: навіть найсучасніші мовні моделі не володіють справжнім «розумінням» тексту в людському сенсі, що може призводити до некоректного або неточного трактування вхідних даних. У результаті модель може генерувати опис, що на вигляд є стилістично правильним і логічно зв’язним, але містить помилки, вигадані факти або несумісність із реальними характеристиками товару.

Ще одним викликом є ризик стилістичної одноманітності – повторюваність мовних конструкцій і шаблонів, що знижує унікальність та маркетингову цінність згенерованих описів. Це особливо важливо у сфері електронної комерції, де тексти виконують функцію залучення уваги, формування довіри та впливу на рішення про купівлю. Одноманітний стиль може створити враження механічності, що негативно позначається на користувацькому досвіді.

Окремої уваги заслуговують проблеми адаптації згенерованого тексту до різних форматів платформ (інтернет-магазини, мобільні застосунки, маркетплейси), що висувають специфічні вимоги до довжини, структури та візуальної подачі контенту. Модель не завжди враховує такі вимоги автоматично, що вимагає додаткової постобробки або інтеграції в інші програмні компоненти.

Істотною є також проблема контролю якості: автоматичні метрики (на кшталт BLEU, ROUGE, METEOR) не завжди адекватно відображають стилістичну, емоційну або маркетингову привабливість тексту. Це обмежує

можливості повної автоматизації процесу без участі людини на етапі верифікації результату.

Не менш важливим є і юридичний аспект: у випадку використання моделей, що можуть некоректно інтерпретувати або вигадувати інформацію, виникає ризик порушення прав споживача або недотримання регуляторних вимог щодо достовірності даних. У сфері комерційної діяльності це може призводити до репутаційних втрат або юридичних санкцій.

Таким чином, попри високий потенціал автоматичної генерації описів, її використання вимагає обережності, ретельного тестування та обов’язкової участі експерта на етапі фінального затвердження тексту. Ці виклики й стали підставою для постановки досліджуваної задачі – порівняння ефективності різних моделей генерації та оцінки їхньої практичної придатності для завдань e-commerce.

3 ВИБІР АРХІТЕКТУРИ ТА МЕТОДОЛОГІЇ ГЕНЕРАЦІЇ

3.1 Вибір мовної моделі для дослідження

У межах даного дослідження було обрано дві великі мовні моделі – GPT-3.5 та GPT-4, розроблені компанією OpenAI. Вони належать до сімейства Generative Pre-trained Transformer і репрезентують різні покоління розвитку архітектур трансформерів. Таке поєднання дозволяє всебічно оцінити якісний стрибок у можливостях моделей на прикладі завдання генерації текстових описів товарів у сфері електронної комерції.

GPT-3.5 є вдосконаленою версією GPT-3, яка вже встигла зарекомендувати себе як надійний інструмент для генерації змістовного тексту в реальному часі. Вона демонструє швидку реакцію, достатню послідовність викладу думок та прийнятний рівень логічності. Водночас ця модель залишається доступною за обчислювальними вимогами, що робить її практичним інструментом для повсякденних бізнес-завдань.

GPT-4, своєю чергою, є сучаснішою і потужнішою моделлю, здатною працювати з довшими та складнішими запитами. Вона краще розуміє інструкції, рідше допускає повтори, глибше аналізує контекст і пропонує більш вивірені за стилем та змістом тексти. Однак, використання GPT-4 вимагає більших обчислювальних ресурсів і є дорожчим.

Для повноти аналізу доцільно у таблиці 3.1 порівняти ключові технічні характеристики моделей GPT-3.5 та GPT-4, що дозволяє оцінити їх потенціал із технічної точки зору.

Це порівняння підтверджує, що GPT-4 є значно потужнішою і гнучкішою моделлю, що краще адаптується до довгих запитів, складних інструкцій і створення глибшого контенту. GPT-3.5, натомість, залишається швидшим і більш економічним рішенням, що може бути корисним у випадках масової генерації простого контенту.

Таблиця 3.1 – Порівняння характеристик моделей GPT-3.5 та GPT-4

Характеристика	GPT-3.5	GPT-4
Рік випуску	2022	2023
Орієнтовна кількість параметрів	~175 млрд	офіційно не розкривається (оцінки: 500–1000+ млрд)
Контекстне вікно (макс. кількість токенів)	до 4 096 токенів	до 128 000 токенів (GPT-4-turbo)
Підтримка багатомодальності	Немає	Так (текст + зображення) – у GPT-4V
Схильність до "галюцинацій" (вигаданих фактів)	Вища	Нижча
Стабільність стилістики та логіки	Помірна	Висока
Продуктивність у стандартних тестах NLP	Висока	Дуже висока (наближається до людської)
Обчислювальні ресурси	Помірні	Значно вищі

Рішення порівняти саме моделі GPT-3.5 і GPT-4 зумовлено низкою факторів, що роблять таке дослідження релевантним і обґрунтованим. По-перше, ці дві моделі є репрезентативними прикладами різних поколінь трансформерних архітектур, що дозволяє простежити еволюцію можливостей великих мовних моделей на практичному рівні. По-друге, обидві моделі доступні через інтерфейс ChatGPT, що значно спрощує організацію експерименту, оскільки не потребує складної обчислювальної інфраструктури або глибоких знань у сфері розгортання LLM. По-третє, GPT-3.5 і GPT-4 підтримують роботу з промптами, сформульованими природною мовою, демонструють схожий формат відповідей і легко

адаптуються до специфіки завдання, що робить їх універсальними і зручними для порівняння в умовах генерації описів товарів.

Для об’єктивної оцінки результатів генерації було визначено низку критеріїв, які дозволяють всебічно охарактеризувати якість згенерованих текстів. Зокрема, розглядалися такі аспекти, як загальна якість тексту (граматична правильність, мовна різноманітність, стилістична узгодженість), відповідність опису реальним характеристикам товару (релевантність), здатність до створення маркетингово привабливого контенту (переконливість, емоційна забарвленість, апеляція до потреб користувача), а також інформативність, тобто повнота охоплення ключових технічних і функціональних параметрів товару. Окрему увагу було приділено відсутності поширених недоліків, таких як повторення фраз, логічні помилки, надмірна шаблонність або загальність, які можуть знижувати якість користувацького сприйняття тексту. Такий підхід дозволив сформувати цілісну картину щодо ефективності кожної з моделей у контексті завдання генерації описів для e-commerce.

Ці параметри дозволяють здійснити як якісну, так і кількісну оцінку здатності моделей генерувати ефективні описи товарів у прикладному комерційному контексті.

3.2 Збір та підготовка даних (описів товарів)

Для проведення експериментального порівняння моделей GPT-3.5 та GPT-4 було сформовано набір вхідних даних, що охоплює три реальні товарні позиції з категорії побутової техніки. Вибір саме цієї категорії пояснюється її поширеністю в електронній комерції, необхідністю поєднання точності технічних характеристик із маркетинговою виразністю, а також наявністю емоційно значущих аспектів, які можуть вплинути на стиль опису.

Обрані товари мають чіткі технічні параметри, належать до відомих брендів і є популярними на ринку. Це дозволяє перевірити здатність мовних моделей не лише працювати з фактичними даними, а й вбудовувати в текст загальні знання про бренд або функціональність пристрою.

У таблиці 3.2 можна побачити, що для забезпечення однорідності та контролю вхідних даних усі три продукти були описані за однаковими параметрами: тип пристрою, технічні характеристики, ключові функції та особливості, що мають значення для споживача.

Таблиця 3.2 – Обрані товари та їх характеристики

Назва товару	Тип	Потужність	Особливості
Електрочайник Xiaomi Mi Smart Kettle Pro	Електрочайник	1800 Вт	Підтримка постійної температури, підключення до смартфона, контроль через додаток, сталевий корпус, LED-дисплей
Блендер Bosch ErgoMixx 600 W	Ручний блендер	600 Вт	12 швидкостей, турборежим, ніж із нержавіючої сталі, ергономічна ручка, знімна насадка для миття в посудомийній машині
Пилосос Dyson V11 Absolute	Бездротовий вертикальний пилосос	до 545 Вт (у режимі Boost)	LCD-дисплей, автоматичне регулювання потужності, до 60 хв роботи без підзарядки, кілька насадок, фільтрація до 99.99% часток до 0.3 мікрона

Щоб забезпечити рівноцінні умови для обох моделей, кожен вхідний запит (промпт) був сформульований у стандартній структурі, яка чітко задає ціль генерації, очікуваний стиль і основні характеристики товару. Це

дозволяє моделі зрозуміти контекст і сформувати текст, максимально наближений до реального комерційного опису. Наприклад, для кожної позиції вхідний запит мав наступну форму.

«Згенеруй опис товару для сайту електронної комерції. Опис має бути лаконічним, інформативним і маркетингово привабливим.

Товар: [назва, характеристики, функції, особливості]».

Нижче наведено конкретні формулювання вхідних запитів для кожного з трьох товарів, які були використані для генерації описів мовними моделями. Запити є уніфікованими за стилем і структурою, що дозволяє забезпечити справедливі умови експерименту та об’єктивність порівняння результатів.

Для електрочайника.

«Згенеруй опис товару для сайту електронної комерції. Опис має бути лаконічним, інформативним і маркетингово привабливим.

Товар: Xiaomi Mi Smart Kettle Pro, електрочайник, об’єм 1.5 л, потужність 1800 Вт, підтримка температури, підключення до смартфона, сталевий корпус, LED-дисплей».

Для блендера.

«Згенеруй опис товару для сайту електронної комерції. Опис має бути лаконічним, інформативним і маркетингово привабливим.

Товар: Bosch ErgoMixx 600 W, ручний блендер, 600 Вт, 12 швидкостей, турборежим, ніж із нержавіючої сталі, ергономічна ручка, знімна насадка».

Для пилососа.

«Згенеруй опис товару для сайту електронної комерції. Опис має бути лаконічним, інформативним і маркетингово привабливим.

Товар: Dyson V11 Absolute, бездротовий пилосос, до 60 хв автономної роботи, LCD-дисплей, автоматичне регулювання потужності, система фільтрації, декілька насадок».

Для забезпечення рівноцінності експериментальних умов було дотримано таких принципів:

– однакові промпти та формулювання для обох моделей;

– вхідні характеристики були однаковими за змістом і структурою;

– аналіз здійснювався лише на основі першої відповіді моделей без додаткових уточнень;

– відповіді не редагувались або не піддавались доопрацюванню.

Ці умови дозволили мінімізувати вплив зовнішніх чинників і зробити експеримент максимально об’єктивним для подальшого порівняльного аналізу якості згенерованих описів.

4 ПРОВЕДЕННЯ ЕКСПЕРИМЕНТУ

4.1 Генерація описів

Для реалізації експерименту було обрано три реальні товари з категорії побутової техніки, характеристики яких структуровано у вигляді вхідних даних. Генерація описів здійснювалася за допомогою двох моделей із використанням ідентичних промптів, сформульованих у природній мові. Це дозволило забезпечити рівність умов при порівнянні результатів.

Кожна модель отримувала на вхід короткий опис товару з основними характеристиками, а у відповідь повертала згенерований маркетинговий текст. Для забезпечення об'єктивності було взято лише перший варіант відповіді без уточнень або додаткових запитів. Запит, який використовувався для обох моделей: «Згенеруй опис товару для сайту електронної комерції. Опис має бути лаконічним, інформативним і маркетингово привабливим. Товар: [назва, характеристики, функції, особливості]».

У таблиці 4.1 подано результати генерації описів для трьох товарів у вигляді таблиці. Це дозволяє зручно порівняти тексти за ключовими ознаками: структурою, стилістикою, довжиною та насиченістю.

Таблиця 4.1 – Згенеровані описи

Товар	GPT-3.5	GPT-4
Xiaomi Mi Smart Kettle Pro	Розумний електрочайник Xiaomi Mi Smart Kettle Pro з об'ємом 1,5 л та потужністю 1800 Вт забезпечує швидке кип’ятіння води. Завдяки підтримці температурного режиму і	Xiaomi Mi Smart Kettle Pro – це поєднання інновацій та комфорту. Об'єм 1.5 л і потужність 1800 Вт забезпечують миттєве нагрівання, а можливість підтримання температури та підключення до смартфона

Продовження таблиці 4.1

	підключенню до смартфона, ви можете керувати пристроєм дистанційно та підтримувати воду теплою. Ідеальний вибір для сучасної кухні.	роблять використання максимально зручним. Ідеальний для тих, хто цінує розумні рішення у побуті.
Bosch ErgoMixx 600 W	Блендер Bosch ErgoMixx 600 Вт – це універсальний помічник на кухні. Завдяки ергономічному дизайну та потужному мотору, пристрій легко справляється з подрібненням, змішуванням і приготуванням пюре. Зручно лежить у руці, має декілька швидкостей для точного контролю.	Bosch ErgoMixx 600 W – надійний блендер для щоденного використання. Потужність 600 Вт, ергономічний корпус і декілька режимів роботи дозволяють готувати смузі, соуси та супи з легкістю. Компактний, зручний і ефективний – ідеальний вибір для домашньої кулінарії.
Dyson V11 Absolute	Потужний бездротовий пилосос Dyson V11 Absolute забезпечує глибоке очищення завдяки інтелектуальному регулюванню потужності та тривалому часу роботи. Підходить для всіх типів покриттів. Зручна конструкція та змінні насадки роблять прибирання простішим.	Dyson V11 Absolute – це новий рівень чистоти. Потужна автономна система, адаптивне регулювання потужності, до 60 хв роботи та ефективна фільтрація роблять його ідеальним вибором для дому. Змінні насадки дозволяють прибирати навіть у важкодоступних місцях.

4.2 Оцінка результатів

Оцінка згенерованих описів, отриманих за допомогою мовних моделей GPT-3.5 та GPT-4, здійснювалася на основі двох ключових підходів: автоматизованого метриками порівняння та експертного аналізу. Такий комбінований підхід дозволяє не лише забезпечити об’єктивну кількісну оцінку результатів, а й врахувати якісні аспекти, які не завжди фіксуються числовими показниками, зокрема читабельність, логічна зв’язність та стилістична відповідність описів.

Для автоматичної оцінки якості текстів були застосовані такі метрики, як BLEU, ROUGE, METEOR і BERTScore. Вони широко використовуються в задачах генерації тексту, машинного перекладу та автоматичного резюмування й дозволяють порівнювати згенеровані тексти із еталонними зразками на основі різних критеріїв – точності збігів, повноти, лексичної та семантичної близькості.

Метрика BLEU (Bilingual Evaluation Understudy) оцінює точність n-грамних збігів між згенерованим текстом і референсним описом. Вона є показовою в контексті повторюваності фраз і точності передачі ключових одиниць інформації. У проведеному експерименті GPT-4 продемонструвала вищі показники BLEU, що свідчить про її здатність генерувати опис ближчий до очікуваного шаблону.

Метрика ROUGE (Recall-Oriented Understudy for Gisting Evaluation) орієнтована переважно на повноту – вона визначає, наскільки повно модель відтворює релевантні елементи з еталонного тексту. За цією метрикою GPT-4 також переважала, зокрема у варіанті ROUGE-L, який оцінює довгі спільні підпослідовності. Це свідчить про кращу здатність моделі зберігати логічну структуру та послідовність фраз.

Метрика METEOR дозволяє більш гнучко оцінювати збіги завдяки врахуванню синонімів, перестановок і морфологічних варіантів. За даними

експерименту, GPT-4 перевершила GPT-3.5 за цим показником, що підтверджує її кращу адаптивність до природної мови.

BERTScore – сучасна метрика, яка використовує векторні представлення слів із попередньо навчених моделей для оцінки семантичної схожості. Її перевага полягає у здатності оцінювати змістовну відповідність навіть за відсутності точних лексичних збігів. GPT-4 отримала вищі результати і за цією метрикою, що підтверджує її кращу здатність зберігати зміст і сенс.

Для кожного з описів (електрочайник, блендер, пилосос) було проведено обчислення метрик BLEU, ROUGE, METEOR та BERTScore. У якості референсних текстів використовувалися оригінальні маркетингові описи товарів із сайтів виробників або маркетплейсів. Результати наведено в таблиці 4.2.

Таблиця 4.2 – Результати обчислення метрик

Показник	GPT-3.5	GPT-4	Коментар
Час генерації одного опису	~3 секунди	~7 секунд	GPT-3.5 швидший
Вартість генерації одного опису*	$0.002	$0.06	GPT-3.5 дешевший
Розмір моделі	~175 млрд параметрів	~1 трлн+ параметрів	GPT-3.5 легша модель
BERTScore (семантична відповідність)	0.89	0.94	GPT-4 кращий
BLEU (точність формулювань)	0.73	0.82	GPT-4 кращий
ROUGE-L (охоплення ключових фраз)	0.76	0.84	GPT-4 кращий
Відсоток стилістичних помилок	8 %	3 %	GPT-4 кращий
Гнучкість стилю (оцінка експертів 1–5)	3.8	4.7	GPT-4 кращий
Стабільність результатів (варіативність між генераціями)	середня	висока	GPT-4 стабільніший

Як показано у таблиці 4.2, GPT-4 демонструє значно вищу якість текстів за всіма автоматичними метриками та експертною оцінкою. Водночас GPT-3.5 має перевагу у швидкості генерації та нижчій вартості на одиницю контенту, що робить її доцільною для сценаріїв масової генерації описів при обмежених ресурсах.

Таким чином, вибір моделі має ґрунтуватися на конкретних вимогах бізнес-процесу: GPT-3.5 підходить для задач, де пріоритетом є швидкість і економічність, тоді як GPT-4 оптимальна для створення високоякісного, преміального текстового контенту.

Окремо було протестовано, наскільки варіативними є відповіді моделей при багаторазовому запуску генерації з ідентичним промптом. Для кожного товару було згенеровано 3 описи на основі того самого запиту. GPT-3.5: продемонструвала обмежену варіативність. Структура описів часто залишалася незмінною, а зміст лише трохи варіювався в межах формулювань. У 2 з 3 випадків повторювалися одні й ті самі фрази на кшталт «ідеальний для щоденного використання». GPT-4: показала значно вищу мовну гнучкість. Опис був різноманітнішим, стиль міг варіюватися від більш інформативного до емоційного, з гнучкою перестановкою структурних блоків. Зміст залишався релевантним, але змінювались акценти: в одному випадку підкреслювалась інноваційність, в іншому – зручність або дизайн.

Ці результати вказують на перевагу GPT-4 у створенні унікальних текстів для кожного товару, що є критичним для SEO та уникнення дублювання контенту на великих платформах.

Окрім автоматичних оцінок, було проведено експертний аналіз з боку трьох незалежних рецензентів, які оцінювали тексти за п’ятибальною шкалою за такими критеріями: читабельність, стилістична відповідність, інформативність, маркетингова привабливість і відповідність товару. Оцінювання проводилося у «сліпому» форматі – експерти не знали, яка з моделей згенерувала той чи інший опис. За результатами оцінки, тексти,

створені GPT-4, отримали в середньому 4,6 бала, тоді як GPT-3.5 – 4,1 бала. GPT-4 демонструвала більш зв’язний, логічно побудований і завершений текст. Водночас GPT-3.5 іноді використовувала менш точні формулювання або спрощену лексику, що знижувало якість сприйняття.

Аналіз сильних і слабких сторін дозволяє зробити висновки про загальні властивості моделей. GPT-4 виявилася ефективнішою в генерації структурованих і маркетингово привабливих текстів, особливо для товарів, що потребують гнучкого опису із вбудованими емоційними або брендовими елементами. Натомість GPT-3.5 показала стабільний, але дещо механістичний підхід, з меншим розмаїттям мовних конструкцій.

Таким чином, результати оцінювання підтверджують вищу якість генерації з боку GPT-4 як за формальними метриками, так і за експертною думкою, що дає підстави рекомендувати її для комерційного використання в задачах автоматичної генерації описів товарів.

5 АНАЛІЗ РЕЗУЛЬТАТІВ

5.1 Висновки з експерименту

У результаті проведеного порівняльного дослідження двох мовних моделей – GPT-3.5 та GPT-4 – було здійснено генерацію описів для реальних товарів з категорії побутової техніки. Отримані тексти були проаналізовані за допомогою автоматичних метрик BLEU, ROUGE, METEOR та BERTScore, а також оцінені експертами з точки зору читабельності, стилістичної привабливості, точності передання інформації та релевантності до поставленого завдання. Результати дозволяють сформулювати такі висновки.

Модель GPT-4 впевнено продемонструвала вищу якість згенерованих текстів за всіма критеріями. Автоматичні метрики підтвердили її здатність створювати більш точні, зв’язні та семантично багаті описи. Зокрема, BERTScore, який оцінює глибину семантичної відповідності, був найвищим саме у GPT-4, що свідчить про кращу адаптацію моделі до контексту і змістової суті товару.

У текстах, згенерованих GPT-4, було зафіксовано менше стилістичних повторів, вища узгодженість між абзацами, а також наявність більш чіткої аргументації переваг товару. Водночас GPT-3.5 хоча й створювала граматично правильні тексти, часто допускала надмірну узагальненість, деяку шаблонність формулювань і менш виразну стилістику.

Результати дослідження свідчать про те, що GPT-4 є значно ефективнішим інструментом для завдань, де потрібна висока якість опису, персоналізація і максимальна релевантність контенту до характеристик продукту. Такий рівень якості є критично важливим для преміальних брендів, складних товарів або платформ, що орієнтуються на емоційне залучення користувачів і конкуренцію за якість контенту.

Водночас GPT-3.5 може залишатися цілком доцільним вибором у випадках, коли обсяг генерації є великим, вимоги до стилістики – базовими, а головна мета – автоматизація і швидкість. Ця модель може бути застосована для масового створення технічних описів, коротких карток товарів або внутрішньої класифікації, де не потрібна глибока адаптація до стилю бренду.

З урахуванням результатів експерименту можна стверджувати, що великі мовні моделі мають значний потенціал для автоматизації створення текстових описів у сфері електронної комерції. GPT-4 доцільно використовувати у тих випадках, коли пріоритетом є якість, стилістична вивіреність та переконливість тексту. Вона найбільше підходить для створення описів преміальних товарів, брендового контенту та маркетингових повідомлень, де важливо досягти емоційного контакту з аудиторією. GPT-3.5, натомість, може застосовуватися для швидкої генерації великої кількості описів у випадках, коли допустимий базовий рівень якості, і ключовим є швидкість та обсяг. У практичних сценаріях ефективною може бути комбінація використання моделей разом із подальшим редагуванням текстів фахівцями, що дозволить досягти оптимального балансу між швидкістю генерації та якістю кінцевого результату. Крім того, для підвищення точності й релевантності доцільно адаптувати запити під специфіку товару, а за можливості – донавчати моделі на корпоративних даних або використовувати спеціалізовані інструменти локалізації. Усе це свідчить про перспективність інтеграції генеративного штучного інтелекту в робочі процеси компаній, які займаються онлайн-продажами, за умови грамотного впровадження та контролю якості.

ВИСНОВКИ

У магістерській кваліфікаційній роботі здійснено комплексне дослідження можливостей застосування великих мовних моделей нового покоління – GPT-3.5 та GPT-4 – для автоматизованої генерації описів товарів у сфері електронної комерції. Проведений аналіз підтвердив високу ефективність використання зазначених моделей у задачах текстової генерації, що передбачають поєднання точності, інформативності та стилістичної привабливості.

На основі аналізу сучасного стану розвитку мовних моделей та їх еволюції від традиційних алгоритмів до трансформерних архітектур обґрунтовано вибір моделей GPT-3.5 та GPT-4 як представників відповідних поколінь. В межах дослідження розглянуто основні сфери застосування LLM у генерації товарного контенту, виокремлено ключові вимоги до таких систем, а також проаналізовано наявні виклики, пов’язані з якістю, релевантністю та адаптивністю згенерованих текстів.

У рамках експериментального етапу було сформовано репрезентативний набір вхідних даних із категорії побутової техніки, для яких було отримано відповідні описи за допомогою обох мовних моделей. Результати генерації підлягали кількісному оцінюванню із застосуванням автоматичних метрик (BLEU, ROUGE, METEOR, BERTScore), а також якісному експертному аналізу, що враховував читабельність, стилістичну узгодженість, відповідність характеристикам товару та маркетингову ефективність.

Проведене порівняльне дослідження продемонструвало, що модель GPT-4 суттєво переважає GPT-3.5 за всіма критеріями якості. Зокрема, GPT-4 забезпечує вищу ступінь контекстуальної релевантності, точності формулювань, логічної послідовності викладу та здатності формувати емоційно насичені, переконливі тексти. Водночас GPT-3.5, незважаючи на деяку шаблонність і меншу виразність, залишається придатною для задач

базового рівня, зокрема – для масової генерації стандартних описів із мінімальними вимогами до стилістики.

Таким чином, результати дослідження підтверджують доцільність і перспективність інтеграції великих мовних моделей у процеси автоматизованого створення текстового контенту для електронної комерції. Використання GPT-4 є обґрунтованим у випадках підвищених вимог до якості й стилістичної індивідуалізації текстів, зокрема для преміальних товарів або брендованих комунікацій. GPT-3.5, у свою чергу, може ефективно використовуватись як інструмент попереднього етапу генерації з подальшим редагуванням або у задачах з високою інтенсивністю автоматизації.

У подальших дослідженнях доцільно розглядати можливості адаптації мовних моделей до специфіки конкретного домену, донавчання на корпоративних даних, а також впровадження механізмів контролю якості контенту з боку фахівців. Отримані результати створюють підґрунтя для розробки прикладних систем генерації товарних описів, здатних підвищити ефективність електронної комерції та покращити взаємодію з кінцевими споживачами.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Zhang Y., Yang Q. A Survey on Multi-Task Learning. *IEEE Transactions on Knowledge and Data Engineering*. 2021. P. 1. URL: https://doi.org/10.1109/tkde.2021.3070203 (date of access: 23.04.2025).

2. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. A., Kaiser Ł., Polosukhin I. «Attention is All You Need». 2017. URL: https://arxiv.org/abs/1706.03762 (date of access: 23.04.2025).

3. Devlin J., Chang M. W., Lee K., Toutanova K. «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding». 2018. URL: https://arxiv.org/abs/1810.04805 (date of access: 30.04.2025).

4. Radford A., Narasimhan K., Salimans T., Sutskever I. «Improving Language Understanding by Generative Pre-Training». 2018. URL: https://openai.com/research/language-unsupervised (date of access: 02.05.2025).

5. Brown T. B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shinn A., Bosma M., Knight M. «Language Models are Few-Shot Learners». 2020. URL: https://arxiv.org/abs/2005.14165 (date of access: 02.05.2025).

6. Nielsen J. «Writing Effective Product Descriptions for E-Commerce». Nielsen Norman Group. 2021. URL: https://www.nngroup.com/articles/product-descriptions (date of access: 15.05.2025).

7. Shopify. «How to Write Product Descriptions That Sell». Shopify Blog. 2023. URL: https://www.shopify.com/blog/product-description (date of access: 15.05.2025).

8. Jurafsky D., Martin J. H. Speech and Language Processing. Pearson, 2023.

9. Manning C. D., Schütze H. Foundations of Statistical Natural Language Processing. MIT Press, 1999.

10. McGuire W. J. Persuasion: Theory and Research. SAGE Publications, 2013.

11. Kotler P., Keller K. L. Marketing Management. Pearson, 2016.

12. OpenAI. GPT-4 Technical Report. 2023. URL: https://openai.com/research/gpt-4 (date of access: 10.05.2025).

13. Anthropic. Introducing Claude: AI Assistant by Anthropic. 2023. URL: https://www.anthropic.com/index/claude (date of access: 10.05.2025).

14. Touvron H., Lavril T., Izacard G., et al. LLaMA: Open and Efficient Foundation Language Models. Meta AI. 2023. URL: https://arxiv.org/abs/2302.13971 (date of access: 10.05.2025).

15. Ramesh A., Pavlov M., Goh G., et al. Zero-Shot Text-to-Image Generation. OpenAI. 2021. URL: https://arxiv.org/abs/2102.12092 (date of access: 12.05.2025).

16. Raffel C., Shazeer N., Roberts A., et al. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. 2020. URL: https://arxiv.org/abs/1910.10683 (date of access: 12.05.2025).

17. Karpinski R. E-commerce Product Copywriting. Routledge, 2020.

18. Gupta S. Artificial Intelligence for Marketing: Practical Applications. Wiley, 2018.

19. Sostek T. E-Commerce Product Content: Creating Content that Sells. Practical E-Commerce Press, 2022.

20. Marcus G. Rebooting AI: Building Artificial Intelligence We Can Trust. Pantheon Books, 2019.

21. Gatt A., Krahmer E. Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation. Journal of Artificial Intelligence Research. 2018. URL: https://jair.org/index.php/jair/article/view/11173 (date of access: 15.05.2025).

22. Google AI Blog. Understanding BERT and How It Works. 2018. URL: https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html (date of access: 15.05.2025).

23. Hoxha E. AI Content Generation for E-commerce: Opportunities and Challenges. Springer, 2022.

24. Renals S., Richmond K. Language Modelling for Automatic Speech Recognition. Springer, 2012.

25. McKinsey & Company. The State of AI in 2023. McKinsey Report. 2023. URL: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-state-of-ai-in-2023 (date of access: 18.05.2025).

26. Forrester Research. AI in E-Commerce: Trends and Opportunities. Forrester Report. 2022. URL: https://go.forrester.com/research/ (date of access: 18.05.2025).

27. W3Techs. Usage of Content Management Systems for Websites. 2025. URL: https://w3techs.com/technologies/overview/content_management (date of access: 18.05.2025).

28. Copy.ai. How AI Copywriting Works. 2024. URL: https://www.copy.ai/blog/how-ai-copywriting-works (date of access: 18.05.2025).