

РОЗРОБКА БОТА-КОНСУЛЬТАНТА ДЛЯ ІНТЕРНЕТ-МАГАЗИНУ НА ОСНОВІ НЕЙРОМЕРЕЖІ ТРАНСФОРМЕРА З ІНТЕГРАЦІЄЮ ДО БАЗИ ДАНИХ

Вінницький національний технічний університет

Анотація

У даній роботі досліджується можливість створення інтелектуального бота-консультанта для потреб інтернет-магазину, який базується на архітектурі нейромережі трансформера та забезпечує інтеграцію зі структурованою базою даних через спеціалізовані програмні інструменти. Розроблена система спрямована на обробку різноманітних запитів користувачів, таких як пошук товарів за заданими параметрами чи надання персоналізованих рекомендацій, із подальшим формуванням відповідей у природній мовній формі, що імітує консультацію реального продавця. У ході дослідження проаналізовано сучасні підходи до автоматизації взаємодії з клієнтами в електронній комерції, визначено ключові принципи проектування архітектури бота, а також оцінено його ефективність і перспективи подальшого розвитку з урахуванням специфіки локальних торговельних платформ.

Ключові слова: інтелектуальний бот, нейромережа трансформера, електронна комерція, база даних, обробка природної мови, автоматизоване консультування.

Abstract

This paper explores the possibility of creating an intelligent chatbot consultant for the needs of an online store, based on the transformer neural network architecture and ensuring integration with a structured database through specialized software tools. The developed system is aimed at processing various user requests, such as product searches based on given parameters or providing personalized recommendations, with subsequent generation of responses in natural language form that mimics a real salesperson's consultation. The study analyzes modern approaches to automating customer interactions in e-commerce, identifies key principles for designing the chatbot architecture, and evaluates its efficiency and prospects for further development, taking into account the specifics of local trading platforms.

Keywords: intelligent chatbot, transformer neural network, e-commerce, database, natural language processing, automated consulting.

Вступ

Сучасний етап розвитку електронної комерції характеризується стрімким зростанням кількості онлайн-покупців, що створює нові виклики для торговельних платформ у забезпеченні якісної, швидкої та персоналізованої взаємодії з клієнтами [4]. Збільшення обсягів онлайн-торгівлі, як зазначається в літературі, зумовлює потребу в інноваційних технологічних рішеннях, які здатні автоматизувати процеси консультування та підвищувати зручність для користувачів [1]. Традиційні чат-боти, що базуються на фіксованих сценаріях або простих алгоритмах обробки запитів, часто виявляються недостатньо гнучкими для відповідей на складні чи нестандартні запити, такі як "Який пілосос обрати для 3-кімнатної квартири до 12000 грн?" або "Скільки телевізорів до 20000 грн у наявності?". У цьому контексті нейромережі трансформери, які вирізняються здатністю до глибокого аналізу контексту й генерації природної мови завдяки механізму уваги [2], відкривають нові перспективи для створення інтелектуальних систем. Проте їхня ефективність значною мірою залежить від інтеграції з базами даних, які є джерелом актуальної інформації про товари, що вимагає розробки спеціальних інструментів для взаємодії між нейромережею та структурованими даними [6].

Метою даного дослідження є розробка концепції бота-консультанта для інтернет-магазину, який поєднує можливості нейромережі трансформера з механізмами доступу до бази даних, забезпечуючи гнучкість у відповідях і адаптивність до потреб користувачів. Наразі проект перебуває на початковій стадії, тому акцент зроблено на теоретичному обґрунтуванні, аналізі існуючих проблем і пропонуванні можливих підходів до їх вирішення. Зокрема, розглядається проблема відсутності прямої взаємодії асистента з базою даних, що може призводити до збоїв у пошуковій системі, а також

досліджується ідея використання багатоетапного пошуку для підвищення точності відповідей. Оскільки остаточні результати ще не сформовані, робота залишає простір для подальшого уточнення архітектури й експериментальних перевірок.

Результати дослідження

Розробка бота-консультанта передбачає створення системи, яка базується на трирівневій архітектурі, що включає клієнтську частину, серверну логіку та базу даних. Клієнтська складова, реалізована у вигляді веб-інтерфейсу або інтеграції з месенджерами, забезпечує зручний спосіб введення запитів користувачами у довільній формі, що є важливим для забезпечення природності взаємодії. Серверна частина, яка є основою системи, спирається на нейромережу трансформера, принципи роботи якої ґрунтуються на механізмах уваги й описані в літературі [2]. Ця нейромережа відповідає за аналіз тексту запитів, витягування ключових параметрів, таких як категорія товару, бюджет чи специфічні характеристики, а також за генерацію відповідей, що відповідають контексту запиту. Для забезпечення зв'язку з базою даних серверна логіка доповнена модулем взаємодії через REST API, розробленим із використанням мови Python, що дозволяє формувати запити до реляційної бази даних, наприклад, MySQL [6].

На даному етапі дослідження особлива увага приділяється проблемі інтеграції нейромережі з базою даних, оскільки попередні спроби реалізації показали, що без такої взаємодії система не здатна надавати точні відповіді на запити, які потребують актуальних даних про товари. Наприклад, запит "Скільки телевізорів до 20000 грн у наявності?" вимагає простого SQL-запиту до бази даних, такого як `SELECT COUNT(*) FROM products WHERE category='TV' AND price<=20000 AND availability='in_stock'` [5]. Проте складніші запити, такі як "Який пілосос обрати для 3-кімнатної квартири до 12000 грн?", потребують не лише витягування параметрів (категорія, бюджет), а й аналізу контексту (розмір квартири, можливі додаткові потреби), що ускладнює обробку без належної інтеграції. Для вирішення цієї проблеми пропонується підхід, за яким нейромережа спочатку аналізує запит, формує параметри для пошуку, передає їх системі управління базою даних, а потім отримує результати й перевіряє їхню відповідність початковому запиту. Якщо даних недостатньо, система може ініціювати додатковий запит до бази або уточнення у користувача, наприклад, "Чи є у вас тварини в квартирі?".

Такий багатоетапний пошук, концепція якого частково базується на сучасних підходах до автоматизації інформаційних систем [3], дозволяє підвищити точність відповідей і зробити систему більш адаптивною. Наприклад, у випадку запиту про пілосос нейромережа може спочатку отримати список моделей у заданому ціновому діапазоні, а потім, проаналізувавши характеристики (потужність, наявність спеціальних насадок), обрати оптимальний варіант і сформулювати відповідь: "Для 3-кімнатної квартири рекомендую пілосос моделі X за 11500 грн, який має високу потужність і підходить для великих приміщень." Якщо потрібних характеристик у базі немає, бот може звернутися до користувача за уточненнями або виконати повторний запит із розширеними параметрами, наприклад, врахувавши теги типу `large_area` чи `pet_friendly`. Цей підхід, хоч і потребує додаткових обчислювальних ресурсів, забезпечує гнучкість, яку не можуть запропонувати традиційні чат-боти.

Для реалізації системи розглядається кілька технічних аспектів. По-перше, серверна частина може бути розгорнута в хмарному середовищі, що дозволяє оптимізувати обробку запитів і масштабувати систему в майбутньому [7]. По-друге, для забезпечення стабільності роботи пропонується впровадити перевірку доступності API та бази даних: у разі збоїв користувач отримуватиме повідомлення типу "Вибачте, наразі не можу отримати дані через технічні проблеми." По-третє, як допоміжний варіант, розглядається можливість використання локальної легкої нейромережі, наприклад, спрощеної моделі на основі DistilBERT, яка могла б виконувати базовий аналіз запитів у разі обмеженого доступу до основної моделі чи хмарного сервера. Час обробки запитів наразі планується обмежити 10 секундами, хоча в процесі тестування цей параметр може бути скоригований залежно від складності запитів і продуктивності системи.

Порівняно з існуючими аналогами, такими як Mindbody чи ShopBot, розроблювана система має потенційні переваги у вигляді гнучкості обробки нестандартних запитів і прямої інтеграції з базою даних, що усуває залежність від сторонніх платформ. Водночас на етапі розробки вже виявлено низку викликів. Зокрема, якість відповідей залежить від повноти й структурованості даних у базі: якщо характеристики товарів не містять достатньої інформації (наприклад, тегів про придатність для певних умов), бот може потребувати додаткових уточнень у користувача, що ускладнює взаємодію.

Крім того, обробка складних запитів може створювати значне навантаження на сервер, що потребує подальшої оптимізації, наприклад, через асинхронні запити чи кешування популярних результатів [7]. Оскільки проект ще не завершений, ці аспекти залишаються предметом подальшого дослідження й експериментів.

Висновки

Дослідження можливості створення бота-консультанта для інтернет-магазину на основі неймережі трансформера з інтеграцією до бази даних показало перспективність такого підходу для автоматизації взаємодії з клієнтами в умовах сучасної електронної комерції [4]. Запропонована концепція системи передбачає гнучке опрацювання запитів, точний пошук інформації та формування природних відповідей, що може підвищити зручність для користувачів і посилити конкурентні позиції торговельної платформи. Використання багатоетапного пошуку та інтеграція з базою даних через спеціалізовані інструменти дозволяють потенційно подолати обмеження традиційних чат-ботів, пов'язані з недостатньою контекстуальністю чи неточністю відповідей [3].

На даному етапі розробки акцент зроблено на теоретичному обґрунтуванні й аналізі можливих шляхів реалізації, що залишає простір для експериментів і уточнень у майбутньому. Серед перспектив вдосконалення системи розглядаються оптимізація продуктивності через кешування чи асинхронні запити, розширення функціоналу шляхом інтеграції з аналітичними модулями для аналізу поведінки користувачів, а також адаптація до ширшого асортименту товарів [1]. Водночас подальші дослідження мають зосередитися на вирішенні таких проблем, як залежність від якості даних у базі та потреба в балансі між швидкістю обробки й точністю відповідей. Отримані наразі результати свідчать про доцільність поєднання неймереж трансформерів із інструментами доступу до структурованих даних як основи для створення інтелектуальних систем у сфері онлайн-торгівлі, однак остаточні висновки залежатимуть від практичної реалізації й тестування.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. E-Commerce Platforms and Their Role in Today's Online Shopping Landscape. URL: <https://www.shopify.com/blog/ecommerce-platforms> (дата звернення: 19.03.2025).
2. The Rise of AI Chatbots in Customer Service. URL: <https://www.zendesk.com/blog/ai-chatbots/> (дата звернення: 19.03.2025).
3. Attention Mechanism in Neural Networks. URL: <https://ai.googleblog.com/2017/06/attention-and-augmented-recurrent.html> (дата звернення: 20.03.2025).
4. What is Retrieval-Augmented Generation (RAG)? URL: <https://developer.nvidia.com/blog/what-is-retrieval-augmented-generation-rag/> (дата звернення: 20.03.2025).
5. SQL SELECT COUNT() Statement. URL: https://www.w3schools.com/sql/sql_count.asp (дата звернення: 21.03.2025).
6. What is Robotic Process Automation? URL: <https://www.automationanywhere.com/rpa/robotic-process-automation> (дата звернення: 21.03.2025).
7. What is Cloud Computing? URL: <https://aws.amazon.com/what-is-cloud-computing/> (дата звернення: 19.03.2025).

Мірошниченко Станіслав Олександрович – студент групи 2КН-216, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, email: vhtrkmlbb@gmail.com.

Колесницький Олег Костянтинович – професор, к.т.н., кафедра комп'ютерних наук ВНТУ, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця, e-mail: kolesnytskiy@vntu.edu.ua.

Stanislav Oleksandrovych Miroshnychenko – Faculty of Intelligent Information Technology and Automation, Vinnytsia National Technical University, Vinnytsia, email: vhtrkmlbb@gmail.com.

Oleg Kostiantynovych Kolesnytskiy – Professor, PhD in Technical Sciences, Department of Computer Sciences, Vinnytsia National Technical University, Faculty of Intelligent Information Technologies and Automation, Vinnytsia, Ukraine, e-mail: kolesnytskiy@vntu.edu.ua.