УДК 004.85:004.912:004.738.5

АВТОМАТИЗОВАНА КАТЕГОРИЗАЦІЯ ТОВАРІВ ІНТЕРНЕТ-МАГАЗИНУ З ВИКОРИСТАННЯМ ВЕЛИКОЇ МОВНОЇ МОДЕЛІ

Нестеренко А.В.[1]; Новікова Т.О.[2]
[1,2] Національний університет «Одеська політехніка»
[1,2] Україна, Одеса
[1] arseniy.nesterenko234@gmail.com; [2] novikova.t.o@op.edu.ua

Анотація. *Розроблено вебзастосунок для автоматизованої категоризації товарів за допомогою великої мовної моделі Gemini. Реалізовано вибір категорії з трирівневої ієрархії, перевірку результату та підтвердження адміністратором. Використано Firebase Firestore для зберігання даних. Рішення спрямоване на спрощення наповнення товарного каталогу.*

Ключові слова: *категоризація товарів; великі мовні моделі; інтернет-магазин; Firebase; Gemini.*

Вступна частина. Розширення асортименту інтернет-магазинів ускладнює підтримання узгодженої структури товарного каталогу. Назви й описи продукції можуть бути неповними або неоднозначними, а схожі товари часто належать до різних кінцевих категорій. Ручне віднесення нової позиції до відповідного розділу потребує часу та створює ризик помилок. У наукових працях запропоновано методи класифікації товарних описів [1], застосування дистрибутивної семантики [2] та ієрархічні моделі з урахуванням структури категорій [3].

Великі мовні моделі (Large Language Models, LLM) дозволяють аналізувати назви та описи товарів без створення окремих правил для кожної категорії. Дослідження класифікації без додаткового навчання (zero-shot) демонструють можливості моделей GPT і Claude для цієї задачі [4]. Водночас модель може помилитися у виборі кінцевої категорії або запропонувати неіснуючий клас. Тому автоматичну пропозицію доцільно поєднувати із серверною перевіркою та рішенням адміністратора.

Серед наявних платформ автоматизований добір категорій застосовує Shopify: система пропонує категорію зі стандартної таксономії, а продавець може перевірити результат [5]. На відміну від такого підходу, у розроблюваному рішенні пропозиція формується для власної трирівневої структури каталогу.

Мета роботи. Спрощення процесу категоризації товарів шляхом розроблення вебзастосунку, який формує пропозицію щодо категорії за текстовими даними товару та забезпечує її перевірку адміністратором. Об’єктом розгляду є процес наповнення й упорядкування товарного каталогу, а основним практичним завданням - реалізація узгодженої взаємодії інтерфейсу, бази даних і сервісу штучного інтелекту.

Основна частина. Розроблено клієнт-серверну архітектуру: публічний та адміністративний інтерфейси реалізуються засобами Next.js, серверна логіка - Spring Boot, зберігання записів товарів і категорій - Firebase Firestore. Для автентифікації адміністратора застосовано Firebase Authentication, а для зображень товарів - Cloudinary. Відвідувач без авторизації може переглядати лише опубліковані товари. Адміністратор створює чернетки, редагує описи й характеристики, керує зображеннями, призначає категорії та публікує товари.

Каталог організовано як трирівневу ієрархію категорій. Кожна категорія має ідентифікатор, рівень та посилання на батьківський вузол. Товар зберігає посилання на кінцеву категорію, а не дубльований текстовий шлях. Це дозволяє підтримувати цілісність зв'язків і перевіряти належність результату класифікації до наявного дерева. Для прототипу підготовлено каталог із 78 категорій та набір зі 144 товарів; ці дані можуть слугувати матеріалом для подальшого оцінювання точності.

Автоматизоване визначення категорії ініціюється адміністратором для збереженої чернетки товару. Сервер формує один запит до моделі Gemini, передаючи назву, опис, характеристики та ієрархію категорій. Структурована відповідь містить ідентифікатори категорій трьох рівнів і коротке пояснення. На відміну від двоетапного підходу, який об'єднує спеціалізований класифікатор і велику мовну модель [6], тут повний шлях пропонується за один виклик Gemini.

Перед збереженням пропозиції сервер перевіряє існування всіх зазначених категорій, відповідність їхніх рівнів та правильність зв'язків між батьківськими й дочірніми вузлами. Адміністратор може прийняти пропозицію або самостійно вибрати іншу категорію. Зміна назви, опису чи характеристик товару скасовує раніше отриману пропозицію як неактуальну. Якщо Gemini API недоступний, ручний вибір категорії залишається можливим; помилка ШІ не блокує керування каталогом.

У програмній реалізації розділено публічні й захищені операції, перевірено статуси товарів і встановлено правила публікації: товар повинен мати назву, опис, щонайменше одне зображення та чинну кінцеву категорію. Для серверної частини виконано модульні та компонентні перевірки, зокрема перевірку прав доступу, структури категорій, умов публікації та обробки некоректних відповідей AI-сервісу. Водночас кількісне оцінювання точності Gemini і часу обробки на тестовому наборі потребує окремих експериментів, тому відповідні числові показники в цій роботі не заявляються.

Висновки. У результаті роботи розроблено вебзастосунок, що спрощує процес наповнення та впорядкування товарного каталогу завдяки автоматизованому визначенню категорій на основі назв, описів і характеристик товарів. Використання моделі штучного інтелекту Gemini дозволяє формувати пропозиції щодо категоризації, а перевірка їх відповідності наявній ієрархії та можливість підтвердження адміністратором забезпечують контроль отриманих результатів. Таким чином, розроблене рішення спрямоване на зменшення обсягу ручної роботи адміністратора та підвищення ефективності управління товарним каталогом інтернет-магазину.

Використання інструментів ШІ. Під час підготовки тез використано ChatGPT (OpenAI, GPT-6) для перевірки пунктуації, граматики та правопису українського тексту, а також для перекладу анотації англійською мовою. Перевірку змісту та остаточне редагування виконано автором.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Vandic, D., Frasincar, F., & Kaymak, U. (2018). A framework for product description classification in e-commerce. *Journal of Web Engineering, 17*(1–2), 1–27. https://doi.org/10.26421/JWE17.1-2

[2] Gupta, V., Karnick, H., Bansal, A., & Jhala, P. (2016). Product classification in e-commerce using distributional semantics. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 536–546). https://aclanthology.org/C16-1052/

[3] Cevahir, A., & Murakami, K. (2016). Large-scale multi-class and hierarchical product categorization for an e-commerce giant. In *Proceedings of COLING 2016, the 26th International*

Conference on Computational Linguistics: Technical Papers (pp. 525–535). https://aclanthology.org/C16-1051/
[4] Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2025). LLMs for product classification in e-commerce: A zero-shot comparative study of GPT and Claude models. *Natural Language Processing Journal, 11*, Article 100142. https://doi.org/10.1016/j.nlp.2025.100142
[5] Shopify. (n.d.). *Product category*. Shopify Help Center. https://help.shopify.com/en/manual/products/details/product-category
[6] Cheng, Z., Zhang, W., Chou, C.-C., Jau, Y.-Y., Pathak, A., Gao, P., & Batur, U. (2024). E-commerce product categorization with LLM-based dual-expert classification paradigm. In *Proceedings of the 1st Workshop on Customizable NLP (CustomNLP4U)* (pp. 294–304). https://doi.org/10.18653/v1/2024.customnlp4u-1.22

Nesterenko A.V., Novikova T.O.

Automated categorization of online store products using a large language model

Abstract. *A web application for automated product categorization using the Gemini large language model has been developed. The system supports category selection from a three-level hierarchy, result validation, and administrator confirmation. Firebase Firestore is used for data storage. The solution aims to simplify product catalog management.*

Keywords: *product categorization; large language models; e-commerce; Firebase; Gemini.*