

ПРОГРАМНИЙ МОДУЛЬ ВИЗНАЧЕННЯ КОДУ МИТНОЇ КЛАСИФІКАЦІЇ ЗА ОПИСОМ ТОВАРУ

Вінницький національний технічний університет

Анотація

Запропоновано інтелектуальний модуль автоматичного визначення коду митної класифікації УКТ ЗЕД за описом товару, що дозволив значно скоротити витрати брокерів на митне оформлення та дає можливості на повну автоматизацію процесу в майбутньому.

Ключові слова: Docker, Elasticsearch, Neo4j, LLM, повнотекстовий пошук, УКТ ЗЕД, ІІІ.

Abstract

An intelligent module for automatic determination of the UKT ZED customs classification code based on the description of the goods has been proposed, which has significantly reduced the costs of customs clearance for brokers and provides opportunities for full automation of the process in the future.

Keywords: Docker, Elasticsearch, Neo4j, LLM, full-text search, UKT ZED, AI.

Вступ

Процес визначення кодів митної класифікації УКТ ЗЕД є трудомістким та часовитратним процесом. Розвиток штучного інтелекту та можливості повнотекстового пошуку відкривають нові можливості для створення інтелектуальних модулів, котрі можуть аналізувати великі обсяги текстових описів. Використання LLM (великих мовних моделей) у поєднанні з механізмами повнотекстового пошуку (Elasticsearch) дозволить значно підвищити точність визначення кодів митної класифікації УКТ ЗЕД, скоротити часові витрати на цей процес та зменшити ймовірність помилок. Автоматизована система, побудована на таких принципах, не лише визначатиме код УКТ ЗЕД на основі опису товару, але й надаватиме користувачеві пояснення щодо вибору конкретного коду, а також його можливі альтернативи.

Метою роботи є розширення функціональних можливостей програмного модуля по автоматичному визначенню кодів УКТ ЗЕД за описом товару з використанням технологій повнотекстового пошуку можливостей штучного інтелекту а саме великих мовних моделей LLM.

Результати дослідження

Інтелектуальні системи автоматичного визначення кодів УКТ ЗЕД на основі текстового опису товару набувають практичного значення у зв'язку з необхідністю прискорення процесів класифікації та зменшення впливу людського фактора. Встановлено, що точність класифікації значною мірою залежить від структури опису товару та здатності системи працювати з неструктурованими текстами [5].

Запропонована система реалізована у вигляді монолітного веб-додатку на основі FastAPI, де взаємодія з користувачем, обробка запитів та виклики мовних моделей відбуваються в єдиному середовищі [1]. Для аналізу опису товару використовуються генеративні мовні моделі (LLM), які автоматично виділяють ключові характеристики (матеріал, призначення, тип, форма), необхідні для подальшої класифікації. Форма введення опису товару побудована за допомогою FastAPI наведена на рисунку 1.

Формування пошукового запиту для визначення УКТ ЗЕД здійснюється шляхом побудови ембедінгу опису, який далі порівнюється з ембедінгами записів у базі через Elasticsearch. Пошук відбувається за косинусною відстанню, що дозволяє враховувати смислову схожість навіть при різному формулюванні. Для забезпечення універсальності та підтримки декількох LLM використано абстрактний клас AIModel, реалізація якого в нащадках (ChatGPTModel, GeminiModel тощо) дозволяє застосовувати принцип поліморфізму [2]. Це спрощує зміну або додавання моделей без потреби модифікувати основну логіку застосунку [1].

Метод тестування базувався на вибірці описів, підібраних за стратегією максимального мінімального віддалення (Max-Min Distance Selection), що дозволило сформувати репрезентативну множину прикладів. Встановлено, що в межах підготовленої вибірки точність системи досягала високого рівня релевантності результатів класифікації.

Введіть опис товару

маска лінзи фари

Підібрати код
Очищення форми

Коди УКТЗЕД для товару за описом:

- (XVI) Машини, обладнання та механізми; електротехнічне обладнання; їх частини; звукозаписувальна та звуковідтворювальна апаратура, апаратура для запису або відтворення телевізійного зображення і звуку, їх частини та приладдя -> (85) Електричні машини, обладнання та їх частини; апаратура для запису або відтворення звуку, телевізійна апаратура для запису та відтворення зображення і звуку, їх частини та приладдя -> (8512) обладнання електроосвітлювальне або сигналізаційне (крім виробів товарної позиції 8539), склоочисники, пристрої, що запобігають обмерзанню та запотіванню, які використовуються на велосипедах або моторних транспортних засобах -> (851290) частини -> (85129090) інші -> (8512909090) інші
- (XV) Недорогоцінні метали та вироби з них -> (83) Інші вироби з недорогоцінних металів -> (8302) арматура кріплення, фурнітура та аналогічні вироби з недорогоцінних металів, що використовуються для меблів, дверей, сходів, вікон, віконниць, у кузовах транспортних засобів, для лимарських виробів, валіз, ящиків, скриньок та аналогічних виробів, вішалки для шляп, гачки для шляп, кронштейни та аналогічні вироби з недорогоцінних металів, ролики з арматурою з недорогоцінних металів, автоматичні пристрої з недорогоцінних металів для закривання дверей (доводчики) -> (83023000) інша арматура кріплення, фурнітура та аналогічні вироби для автотransпортних засобів -> (8302300090) інші

Рисунок 1 – Форма введення опису товару та результат підбору кодів за описом товару

Таким чином, запропоновано та реалізовано ефективний програмний метод визначення кодів УКТЗЕД на основі поєднання LLM, ембедінгів та повнотекстового пошуку, адаптований до практичних завдань митної класифікації.

Висновки

Встановлено, що запропонований підхід дозволяє підвищити точність автоматичного визначення кодів УКТЗЕД на основі текстового опису товару, забезпечуючи релевантність результатів навіть за умови нечітко структурованих вхідних даних, що відповідає практичним вимогам митної класифікації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Fast API Documentation, URL: <https://fastapi.tiangolo.com/> (date of access: 01.02.2025).
2. Elasticsearch Guide [8.13] | Elastic. Elastic – The Search AI Company | Elastic. URL: <https://www.elastic.co/guide/en/elasticsearch/reference/current/index.html> (date of access: 27.02.2025).
3. Neo4j Documentation. Neo4j. URL: <https://neo4j.com/docs/> (date of access: 21.02.2025).
4. Що таке Docker і як використовувати його з Python. DevZone. URL: <https://devzone.org.ua/post/shcho-take-docker-i-yak-vykorystovuvaty-yoho-z-python> (дата звернення: 12.03.2025).
5. Державна Митна служба «Методичні рекомендації щодо класифікації окремих товарів згідно з вимогами УКТЗЕД». URL: <https://customs.gov.ua/web/content/5470/%D0%A2%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D1%8F%201.docx> (дата звернення: 12.01.2025)

Хомин Назар Миколайович – студент групи 1KN-21б факультету ІІТА кафедри комп'ютерних наук.

Науковий керівник: *Сілагін Олексій Віталійович* — к.т.н, доцент кафедри КН.

Khomyn Nazar Mykolayovych – student of group 1KN-21b of the IITA Faculty of Computer Science.

Supervisor: *Sihagin Oleksii Vitaliiiovych* — Ph.D. in Engineering, Associate Professor of the Department of Computer Science., Vinnytsia National Technical University, Vinnytsia.

