

## АНАЛІЗ ВИКОРИСТАННЯ HS- ТА HTS-КОДІВ У СИСТЕМАХ МИТНОЇ КЛАСИФІКАЦІЇ: ВИКЛИКИ ТА МОЖЛИВОСТІ ІНТЕГРАЦІЇ ІТ-ТЕХНОЛОГІЙ

Степан Крупа<sup>1</sup>, Наталія Кунанець<sup>2</sup>

<sup>1</sup> Національний університет “Львівська політехніка”,  
кафедра систем штучного інтелекту, Львів, Україна,

<sup>2</sup> Національний університет “Львівська політехніка”,  
кафедра інформаційних систем та мереж, Львів, Україна,

<sup>1</sup> E-mail: [stepan.m.krupa@lpnu.ua](mailto:stepan.m.krupa@lpnu.ua), 0009-0000-2074-9762

<sup>2</sup> E-mail: [Natalia.E.Kunanets@lpnu.ua](mailto:Natalia.E.Kunanets@lpnu.ua), 0000-0003-3007-2462

© Крупа С., Кунанець Н., 2024

Проаналізовано особливості використання гармонізованої системи опису та кодування товарів, гармонізованої тарифної системи кодів у сучасних системах митної класифікації. Особливу увагу приділено викликам, що виникають при застосуванні цих кодів, зокрема через складність товарної номенклатури, а також різноманітність описів товарів. Окрім того, досліджуються можливості інтеграції ІТ-технологій, машинного навчання та методів штучного інтелекту для автоматизації та оптимізації процедур митної класифікації. Розглядаються перспективи підвищення точності та ефективності роботи митних органів завдяки впровадженню інноваційних рішень. Також важливо зазначити, що системи класифікації можуть відрізнятися між країнами, що ускладнює уніфікацію процесу на міжнародному рівні. Це стає серйозною перешкодою для ефективної митної діяльності. Завдяки машинному навчанню та аналізу великих обсягів даних, митні органи можуть ефективніше виявляти невідповідності та оптимізувати роботу з товарними кодами. Впровадження таких інноваційних рішень допоможе покращити точність і швидкість роботи митних служб, що своєю чергою сприятиме прозорості та ефективності міжнародної торгівлі.

**Ключові слова:** HS-коди, HTS-коди, митна класифікація, ІТ-технології, автоматизація, машинне навчання.

### Вступ

У сучасних умовах глобалізації та зростання міжнародної торгівлі митна класифікація відіграє ключову роль в ефективному управлінні товаропотоками між країнами. Однією з основних складових митної системи є використання кодів Гармонізованої системи (HS) та Гармонізованої тарифної системи (HTS), що забезпечують єдиний підхід до класифікації товарів на міжнародному рівні. Ці коди є важливим інструментом для регулювання митних тарифів, зборів та митних процедур.

Проте з розвитком торгівлі та постійною зміною товарних номенклатур, виникають певні виклики в адаптації цих систем до нових умов. Ці виклики зумовлені необхідністю точнішого і швидшого визначення кодів для нових товарів, автоматизацією процесів класифікації та інтеграцією ІТ-технологій у митні операції. Використання передових інформаційних технологій, як-от

системи на основі штучного інтелекту (ШІ) та великих даних, може значно підвищити ефективність і точність митної класифікації.

Актуальним є проведення аналізу особливостей використання HS- та HTS-кодів у системах митної класифікації, визначення основних викликів, з якими стикаються митні органи і бізнес у цих процедурах, та дослідження можливості інтеграції ІТ-технологій для вирішення наявних проблем.

### **Формулювання проблеми**

Митна класифікація товарів є основоположним елементом міжнародної торгівлі, оскільки вона визначає правильність нарахування митних платежів, відповідність торговим нормам та регулятивним вимогам. У цьому контексті використання Гармонізованої системи опису, кодування товарів (HS) та Гармонізованої тарифної системи (HTS) кодів є необхідним для забезпечення точності та ефективності процесу класифікації товарів. Проте складність товарної номенклатури та різноманітність описів часто стають причинами неточностей і помилок у класифікації.

З урахуванням глобалізації ринків і зростання обсягів міжнародної торгівлі стає необхідним вдосконалення процесів митної класифікації. Сучасні ІТ-технології, машинне навчання та методи штучного інтелекту пропонують нові можливості для автоматизації цього процесу, підвищуючи точність і швидкість визначення кодів. Однак впровадження цих технологій вимагає ретельного аналізу поточних систем, розуміння наявних викликів і визначення можливих напрямків розвитку [1].

У сучасних умовах міжнародної торгівлі митна класифікація товарів відіграє важливу роль у забезпеченні правильного обліку та дотриманні торговельних норм. Правильне визначення митних кодів безпосередньо впливає на дотримання регулятивних вимог, точність нарахування митних платежів, а також на швидкість опрацювання даних про вантажі на митниці. Проте традиційні методи визначення кодів товарів, що базуються на людському факторі, часто є недостатньо ефективними, оскільки можуть бути неточними та вимагати значних часових витрат. В умовах зростаючих обсягів торгівлі виникає необхідність у впровадженні нових підходів до автоматизації цього процесу. Машинне навчання та інноваційні ІТ-технології відкривають можливості для підвищення точності та швидкості визначення митних кодів, проте їхня інтеграція потребує глибокого аналізу та адаптації до специфічних завдань митної класифікації.

У контексті зростаючих обсягів міжнародної торгівлі та глобалізації ринків виникає потреба у вдосконаленні процесів митної класифікації товарів. Традиційні методи, що покладаються на ручну роботу, часто призводять до помилок і затримок, що ускладнює ефективне управління митними процедурами [2]. Інноваційні ІТ-технології, зокрема машинне навчання, мають потенціал для суттєвого покращення цих процесів, забезпечуючи більш точне і швидке визначення кодів товарів.

Проте впровадження таких інноваційних підходів потребує ретельного аналізу. Необхідно оцінити наявні системи митної класифікації, виявити їхні недоліки та виклики, а також визначити потенційні шляхи для інтеграції нових технологій. Це передбачає вивчення алгоритмів машинного навчання, їхню адаптацію до конкретних завдань митної служби, а також оцінку впливу таких технологій на загальну ефективність митних операцій. Тож таке дослідження може закласти основу для модернізації митних процесів та сприяти розвитку більш ефективних і надійних методів класифікації товарів.

### **Аналіз останніх досліджень та публікацій**

В останні роки спостерігається значне зростання наукового інтересу до застосування інформаційних технологій, зокрема машинного навчання, у митній класифікації. Багато дослідників фокусуються на автоматизації процесу підбору HS- і HTS-кодів, що є основою для правильного нарахування митних платежів та забезпечення відповідності міжнародним торговельним стандартам.

У роботах таких авторів, як-от Liya Ding [3], проводиться глибокий аналіз алгоритмів машинного навчання, які можуть застосовуватися для класифікації товарів на основі текстових описів.

Дослідження показують, що традиційні методи класифікації часто не можуть задовольнити вимоги сучасного ринку через їхню обмежену точність і високу залежність від людського фактора. Наприклад, у роботі ZhenZhen Fan підкреслюють, що більшість помилок у класифікації товарів виникає через варіативність описів та складність товарної номенклатури, що значно ускладнює роботу митних органів.

Паралельно з цим дослідження, присвячені інтеграції штучного інтелекту в митні системи, демонструють перспективи значного підвищення ефективності та точності процесу класифікації. Зокрема, робота DongLiang Chen акцентує увагу на тому, що використання алгоритмів глибокого навчання дає змогу зменшити кількість помилок та підвищити швидкість опрацювання митної інформації, що особливо важливо в умовах зростаючих обсягів міжнародної торгівлі.

Досліджено можливість застосування машинного навчання для автоматизації процесу класифікації. Особливу увагу приділено оцінці різних алгоритмів машинного навчання, як-от нейронні мережі та опрацювання природної мови (NLP), для підвищення точності і швидкості підбору HS та HTS кодів [4].

До того ж низка досліджень вказує на виклики, пов'язані з впровадженням таких інноваційних підходів, відзначено складність адаптації алгоритмів машинного навчання до специфіки митних даних, а також необхідність забезпечення надійності та прозорості таких систем для їхньої успішної інтеграції в наявні митні процедури.

Отже, аналіз останніх досліджень та публікацій показує, що, хоча застосування машинного навчання в митній класифікації є перспективним напрямком, воно вимагає подальших досліджень і оптимізації для подолання наявних викликів та максимізації його потенціалу.

### Формулювання цілі статті

Метою статті є комплексний аналіз і вдосконалення процесів митної класифікації товарів за допомогою сучасних ІТ-технологій, зокрема машинного навчання. Перше завдання полягає в оцінці важливості використання HS-кодів (Harmonized System) та HTS-кодів (Harmonized Tariff Schedule) для забезпечення точності та ефективності митних процедур у міжнародній торгівлі. Це передбачає детальне дослідження ролі цих кодів у дотриманні торгових норм і оптимізації митних платежів.

Для досягнення **мети** необхідно вирішити такі завдання:

1. Проаналізувати основні принципи функціонування систем HS- та HTS-кодів у митній класифікації.
2. Визначити ключові виклики, з якими стикаються митні органи та бізнес у процесі використання цих систем.
3. Дослідити можливості інтеграції інформаційних технологій для автоматизації та покращення процесів митної класифікації.
4. Оцінити потенціал використання сучасних ІТ-рішень, як-от штучний інтелект та великі дані, для підвищення точності та ефективності класифікаційних процесів.
5. Запропонувати рекомендації щодо подальшого розвитку митної класифікації з використанням ІТ-технологій.

**Об'єктом** дослідження є системи митної класифікації, засновані на використанні кодів Гармонізованої системи (HS) та Гармонізованої тарифної системи (HTS).

**Предметом** дослідження є процеси застосування та вдосконалення систем HS- та HTS-кодів у митній класифікації з акцентом на інтеграцію інформаційних технологій для вирішення наявних викликів та підвищення ефективності митних операцій.

### Виклад основного матеріалу

Проведення аналізу наявних традиційних методів класифікації товарів та виявлення їхніх недоліків передбачає розгляд основних проблем, як-от значні витрати часу в процесі класифікації, неточності і складність управління великою кількістю товарних описів, які зумовлюють неефективність традиційних методів.

#### Важливість HS- та HTS-кодів у митній класифікації

Гармонізована система опису та кодування товарів (HS) – це міжнародна система класифікації товарів, розроблена та підтримувана Всесвітньою митною організацією (WCO). Вона широко використовується в митних операціях, торгівлі, а також для статистичних цілей. Офіційно введена 1988 року, HS на сьогодні використовується більш ніж 200 країнами й економічними об'єднаннями для класифікації товарів у процесі імпорту та експорту. HS-коди (Harmonized System) та HTS-коди (Harmonized Tariff Schedule) [5] є основними інструментами для класифікації товарів у міжнародній торгівлі. HS-коди забезпечують універсальну систему для опису та кодування товарів. Вони використовуються у всьому світі для спрощення та стандартизації митних процедур. HTS-коди є розширенням HS-кодів, що детально відображає митні ставки та тарифи для конкретних країн.

Кожен товар у системі HS класифікується за допомогою унікального шестизначного коду, який має три рівні:

1. Групи товарів (перші два цифри): визначають загальну категорію товару.
2. Підгрупи товарів (середні два цифри): уточнюють категорію товару в межах групи.
3. Підкатегорії товарів (останні два цифри): деталізують характеристики товару.

ГАРМОНІЗОВАНА СИСТЕМА	ГАРМОНІЗОВАНА СИСТЕМА	ГАРМОНІЗОВАНА СИСТЕМА	НОМЕНКЛАТУРА ЄС	УКРЗЕД
08	02	32	00	00
Група (HS Chapter)	Товарна позиція (HS Heading)	Підпозиція товару (HS Subheading)	УКРЗЕД - товарна категорія (CN Code)	Товарна підпозиція (Taric code)

Рис. 1. Розподіл HS-коду

Перші два знаки коду 08 описують назву товарної групи. У цьому випадку це їстівні плоди та горіхи, шкірки цитрусових або динь. Чотири знаки дають детальніший опис на рівні товарної позиції.

Ось кілька прикладів товарів, що входять до цієї групи:

1. 0801 – Кокоси, бразильські горіхи та кеш'ю.
2. 0802 – Інші горіхи (мигдаль, фундук, волоські горіхи, пекан, фісташки).
3. 0803 – Банани, включно з плантанами.
4. 0804 – Фініки, інжир, ананаси, авокадо, гуава, манго.
5. 0805 – Цитрусові фрукти (апельсини, мандарини, лимони, лайми).
6. 0806 – Виноград, свіжий або сушений (ізюм).
7. 0807 – Дині, кавуни та інші плоди гарбузових.
8. 0808 – Яблука, груші та айва.
9. 0809 – Абрикоси, вишні, черешні, персики, сливи та терен.

Товари, що належать до цієї групи, охоплюють як свіжі, так і сушені плоди та горіхи.

Наступний крок – підпозиція товару на рівні шести знаків 080232, що означає “інші”, далі переходимо до класифікації на рівні останніх восьми і десяти знаків 0802320000 відповідно товарної категорії і товарної підкатегорії, які дають остаточний детальний опис продукції, детальна назва якої “Інші горіхи, свіжі або сушені, очищені від шкаралупи або неочищені, зі шкіркою або без

шкірки: горіхи волоські: без шкаралупи.” Пунктуація також є достатньо важливим елементом у позначенні товарної номенклатури. До неї можна віднести кому, крапку з комою та дві крапки, що зазвичай використовуються у товарній номенклатурі при описі товарних угруповань.

Правильне визначення та використання цих кодів критично важливо для забезпечення відповідності торговим нормам, правильного розрахунку митних платежів і уникнення торгових бар'єрів. Неправильна класифікація товарів може призвести до помилок у митних розрахунках, штрафів і затримок у митних процедурах.

### ***Традиційні методи класифікації товарів***

Використання HS- та HTS-кодів дає змогу опрацювати митні декларації, оскільки коди забезпечують точну ідентифікацію товару, а також оптимізувати процеси збору статистичних даних про імпорт та експорт. HS- та HTS-коди використовуються для визначення ставок митних зборів, ПДВ, акцизів та інших зборів на товари під час їх переміщення через кордон. HS-коди допомагають урядам встановлювати та контролювати імпорتنі обмеження, квоти та нетарифні бар'єри. Вони також використовуються для регулювання контролю за експортом товарів, які можуть підпадати під санкції або експортні обмеження. Дані, зібрані з використання стандартизованих кодів, дають змогу зменшити складність при роботі з різними митними системами різних країн, забезпечуючи узгодженість у класифікації товарів. Це полегшує роботу митних брокерів, імпортерів, експортерів і перевізників. За допомогою HS- та HTS-кодів використовуються для аналізу обсягів міжнародної торгівлі, економічного прогнозування та визначення торговельних тенденцій. HS- та HTS-коди є основним інструментом для стандартизації та регулювання світової торгівлі, а їхнє правильне використання дає змогу забезпечити ефективність митного контролю та економічної політики.

Традиційні методи класифікації товарів за допомогою HS- та HTS-кодів переважно передбачають ручний процес. Митні агенти або інші фахівці аналізують товарні описи і визначають відповідні коди, проводячи пошук за таблицями. Цей процес є повільним і може містити помилки, оскільки залежить від людського фактора і передбачає багато рутинної роботи. Поява помилок у класифікації може призвести до штрафів або затримок на митниці.

Основні недоліки традиційних методів:

- висока ймовірність помилок в класифікації;
- значні часові витрати;
- складність опрацювання великого обсягу даних;
- часте оновлення системи;
- різне тлумачення кодів;
- обмеженість кодування.

Помилки можуть призвести до неправильного підбору кодів, що впливає на точність митних платежів. Ручна класифікація займає багато часу, що затримує митні процеси. Велику кількість товарів важко класифікувати без автоматизації, що ускладнює управління даними. Для деяких товарів класифікація за кодами HS та HTS може бути непростою, особливо якщо товар належить до нових технологій або має кілька призначень. Щорічно WCO оновлює коди HS, що може вимагати змін у програмному забезпеченні та методологіях класифікації. У різних країнах може бути власна інтерпретація окремих кодів, що створює проблеми для точності митних процедур. Шестизначна структура не завжди дає змогу детально описати товари, особливо нові або специфічні продукти.

### ***Потенціал машинного навчання в митній класифікації***

Підвищенню ефективності, точності та швидкості опрацювання даних для оптимізації роботи з HS- та HTS-кодами сприяє використання інформаційних технологій.

Застосування інформаційних технологій для оптимізації роботи з HS- та HTS-кодами доцільно здійснювати декількома шляхами:

- автоматизація процесів класифікації з використанням інформаційних технологій може значно полегшити процес призначення HS-кодів на основі описів товарів;
- інтеграція інформаційних систем з митними базами даних сприятиме спрощенню ведення бізнесу, даючи змогу автоматично виявляти зміни в HS-кодах і оновлювати внутрішні каталоги товарів;
- використання методів штучного інтелекту дасть змогу аналізувати великі масиви даних та підвищити точність при класифікації нових або складних товарів.

Інформаційні технології, що використовують алгоритми автоматизації, можуть значно спростити процес класифікації товарів за кодами HS та HTS. Завдяки застосуванню таких технологій можна автоматизувати вибір правильного коду для певного товару на основі введених параметрів та характеристик товару, здійснити швидкий пошук відповідних кодів у базі даних HS- та HTS-кодів, зменшити кількість помилок у класифікації. Інформаційні технології також можуть бути використані для забезпечення дотримання всіх вимог митного контролю, перехресної перевірки даних на предмет порушень, як-от неправильно класифіковані товари, невідповідність тарифів або неправильно розраховані податки, відстеження та моніторингу товарних потоків, що допомагає боротися з контрабандою та порушеннями митного законодавства. Застосування інформаційних технологій у системах класифікації товарів за допомогою HS- та HTS-кодів є не лише доцільним, а й необхідним у сучасній глобальній економіці. Використання ІТ-технологій сприятиме підвищенню точності, швидкості та ефективності опрацювання товарів, зменшенню ризику помилок та зручнішому управлінню митними процедурами.

Оскільки коди HS та HTS регулярно оновлюються Всесвітньою митною організацією (WCO) та національними урядами, щоб системи класифікації були завжди в актуальному стані, використання інформаційних технологій може забезпечити автоматичні оновлення бази даних кодів, приведення їх у відповідність до кожної ітерації змін, що знизить ризик використання застарілих кодів. Використання хмарних сервісів надає швидкий доступ до нових даних.

Завдяки ІТ-рішенням можна інтегрувати системи класифікації HS- та HTS-кодів із митними, логістичними та фінансовими платформами. Технології швидкої передачі даних між різними системами дають змогу оптимізувати процес митного декларування, експорту та імпорту товарів. Синхронізацію інформації про тарифи, квоти та торговельні угоди спрощує процеси планування та прийняття рішень у сфері міжнародної торгівлі.

Інформаційні технології можуть значно підвищити ефективність збору, зберігання та аналізу даних про товарні потоки та використання HS- та HTS-кодів. Аналітичні системи можуть опрацьовувати великі обсяги даних, щоб виявити тенденції у світовій торгівлі, оцінити ефективність використання тарифних пільг та визначити основні проблеми митної системи. Автоматизована генерація звітів дає змогу відстежувати поточні операції та статистику, що значно прискорює аналіз та прийняття рішень.

Використання інформаційних систем для класифікації товарів за HS- та HTS-кодами допомагає уникнути помилок, пов'язаних із ручним опрацюванням даних. Використання інформаційних систем сприяє підвищенню ефективності перевірки відповідності кодів законодавчим вимогам та міжнародним стандартам, забезпечують точні розрахунки митних зборів та тарифів, що особливо важливо при складних комбінаціях товарів або змішаних партіях товарів.

Впровадження ІТ-рішень дає змогу скоротити час на ручну класифікацію товарів і зменшити витрати, пов'язані з людським фактором, оптимізувати процедуру використання ресурсів митних служб та підприємств завдяки автоматизації повторюваних завдань, прискорити опрацювання митних декларацій, що дає змогу пришвидшити процеси імпорту та експорту товарів.

Інформаційні системи дозволяють гармонізувати процедури міжнародної торгівлі та сприяти кращій співпраці між країнами:

Методи машинного навчання як складові штучного інтелекту пропонують нові можливості для автоматизації процесу митної класифікації. Основні технології машинного навчання, які можуть бути застосовані в цій сфері, передбачають:

- нейронні мережі;
- технології опрацювання природної мови (NLP);
- алгоритми класифікації.

Глибокі нейронні мережі можуть опрацьовувати складні текстові дані та забезпечувати високу точність класифікації завдяки здатності розпізнавати складні патерни. Технології NLP аналізують текстові описи товарів і витягують ключову інформацію для точного визначення кодів. NLP може допомогти зменшити неоднозначність у описах і полегшити автоматичний підбір кодів.

Детальніше зупинимося на використанні алгоритмів класифікації для автоматичного призначення HS- та HTS-кодів на основі текстового опису товару. Ці алгоритми можуть навчатися на великих наборах даних і покращувати точність класифікації з часом. Класифікації HS складні і їхнє використання для присвоєння коду товару складний рутинний процес. Наведемо приклади, що ілюструють суперечливість класифікацій HS. Багато компаній виробляють спеціалізовану продукцію, і вони оголошують усі за частини та їхні складові частинами основного продукту (це дає змогу знизити ставки мит). Проте гвинти, пружини, шайби, гайки, кільця ущільнювачів тощо можна вважати деталями “загального призначення” і класифікувати як певні типи товарів, що потребує обкладання їх вищими ставками мита.

Класифікаційна рубрика 9505 10 охоплює вироби, які широко визнані традиційно використовуваними під час Різдвяних свят і виробляються та проєктуються як вироби для Різдвяних свят. Однак він не охоплює предмети “зимового сезону”, як-от: бурульки, снігові кристали, зірки, північні олені або сніговики, тому що вони використовуються не тільки для Різдвяних свят, але і як прикраси протягом усього зимового сезону. У Великій Британії ставка мита становить 0 % для різдвяних прикрас та 2,7 % для зимових прикрас.

Використання алгоритмів класифікації для оптимізації роботи з HS- та HTS-кодами може бути зосереджене на автоматизації процесів класифікації товарів під час їхнього митного оформлення. Розглянемо приклад, де використання машинного навчання та алгоритмів класифікації допомагає підприємствам автоматизувати та оптимізувати цей процес.

Скажімо міжнародна логістична компанія має великий обсяг товарних операцій і опрацьовує тисячі вантажів щомісяця. Кожен вантаж вимагає правильного класифікування згідно з митними нормами, що вимагає використання HS- та HTS-кодів для визначення тарифів, податків і торгових обмежень. Як вже зазначалося, процес класифікації товарів є трудомістким. Логістична компанія вирішує впровадити інформаційну систему на основі методів машинного навчання для класифікації товарів з використанням HS- та HTS-кодів, зокрема Support Vector Machines (SVM), для порівняння товарів зі схожими характеристиками і присвоєння правильного коду на основі аналізу наявної бази даних.

Метод **Support Vector Machines** – це один із найпотужніших та найпопулярніших методів машинного навчання, який використовується для задач класифікації та регресії. Основна ідея алгоритму полягає в пошуку гіперплощини, яка максимально розділяє дані на дві або більше категорії.

SVM намагається знайти гіперплощину (лінію в 2D, площину в 3D або площину вищої розмірності), яка розділяє точки даних різних класів так, щоб максимізувати відстань (маржу) між найближчими точками різних класів. Ці точки, які визначають маржу, називаються **опорними векторами**. Якщо дані є лінійно розділеними, SVM знаходить пряму або гіперплощину, яка їх розділяє. Якщо дані не можна розділити лінійною гіперплощиною, SVM використовує **ядрові методи (kernel tricks)**. Ядро дає змогу перенести дані в простір вищих розмірностей, де їх можна

лінійно розділити. Найпоширеніші ядра: поліноміальне (Polynomial Kernel), радіальне базисне функціональне ядро (RBF Kernel), лінійне ядро (Linear Kernel).

SVM шукає максимальну маржу між класами, що дає змогу знайти найбільш надійну гіперплощину для розділення класів. Це підвищує стійкість до помилок і забезпечує кращу узагальненість для нових даних. Вектори, які знаходяться на краю маржі або близько до неї, називаються **опорними векторами**. Тільки ці точки впливають на положення гіперплощини. Це дає змогу зробити алгоритм ефективним, оскільки він ігнорує всі інші точки, що не лежать на межі класифікації.

Параметр  $C$  дає змогу контролювати баланс між максимальною маржею та кількістю помилок класифікації. Високе значення  $C$  надає перевагу коректній класифікації над великою маржею, а низьке значення навпаки дає змогу створювати велику маржу, навіть якщо деякі точки будуть неправильно класифіковані.

Опорні вектори допомагають алгоритму знаходити найбільш важливі рубрики класифікатора для точного класифікування.

Support Vector Machines (SVM) ґрунтується на оптимізації задачі пошуку гіперплощини, яка розділяє дані. Основне завдання SVM полягає в максимізації відстані (маржі) між класами за умови правильного розділення. Формально це можна подати так:

Гіперплощина, що розділяє два класи, визначається рівнянням:

$$w^T x + b = 0, \quad (1)$$

де  $w$  – це вектор ваг (коефіцієнти, які визначають орієнтацію гіперплощини);  $x$  – це вектор вхідних ознак;  $b$  – це зміщення (bias), яке визначає положення гіперплощини.

Для коректної класифікації точок, які належать до різних класів, використовуються такі умови:

$$\begin{aligned} w^T x_i + b &\geq 1 \text{ для } y_i = +1, \\ w^T x_i + b &\leq -1 \text{ для } y_i = -1, \end{aligned} \quad (2)$$

де  $y_i$  – це мітка класу (+1 або -1) для точки  $x_i$ .

Відстань між двома найближчими точками різних класів (маржа) дорівнює:

$$\text{Маржа} = \frac{2}{\|w\|}. \quad (3)$$

Щоб максимізувати маржу, необхідно мінімізувати  $\|w\|$ . Формулюється задача оптимізації:

$$\min_{w,b} \left( \frac{1}{2} \right) \|w\|^2. \quad (4)$$

Забезпечується обмеження, що кожна точка класифікується коректно:

$$y_i (w^T x_i + b) \geq 1 \quad \forall_i. \quad (5)$$

Отже, класична задача оптимізації для лінійного SVM полягає в мінімізації такого функціонала:

$$\min_{w,b} \left( \frac{1}{2} \right) \|w\|^2 \text{ за умови } y_i (w^T x_i + b) \geq 1. \quad (6)$$

Для нелінійно розділених даних використовуються ядрові методи, що дають змогу перенести дані в простір вищих вимірів. Ядрова функція  $K(x_i, x_j)$  обчислює скалярний добуток у цьому просторі:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j), \quad (7)$$

де  $\phi(x)$  – це відображення даних у простір вищих вимірів.

Для реальних задач часто додають параметр  $C$  для контролю за кількістю неправильно класифікованих точок. Задача оптимізації у разі м'якої маржі виглядає так:



$$\min_{w,b,\varepsilon} \left( \frac{1}{2} \right) \|w\|^2 + C \sum_{i=1}^n \xi_i, \quad (8)$$

де  $\xi_i$  – це змінні для штрафування помилок, а  $C$  – параметр, що контролює рівень регуляризації (баланс між правильною класифікацією і шириною маржі).

Компанія збирає велику базу даних товарів. Ця база містить опис товарів, специфікації, походження, історію класифікації, митні декларації та інші характеристики. Збираємо дані про товари, тобто їхній опис, специфікації. Приводимо дані до формату, зручного для SVM, векторизуючи текстові описи товарів у числові характеристики.

Якщо є тільки описи товарів без кодів HS, завдання полягає у створенні моделі, яка автоматично визначить HS-код на основі текстових описів. Для цього потрібен навчальний набір даних, де для частини товарів є відповідні HS-коди, а для решти товарів – тільки описи. Цей процес передбачає кілька ключових кроків:

**Крок 1. Збір навчальних даних.**

**Крок 2. Попереднє опрацювання даних.**

**Крок 3. Навчання моделі SVM.**

**Крок 4. Прогнозування кодів для нових товарів.**

**На першому етапі відбувається формування навчальних даних.** Для тренування моделі потрібні зразки товарів з уже відомими HS-кодами та їхні описи. Якщо у вас немає кодів для всіх товарів, то це означає, що деякі товари будуть використовуватися тільки для прогнозування (класифікації), коли модель буде готова. **Розділяємо дані на тренувальний і тестовий набори** – 70–80 % для тренування та 20–30 % для тестування. Використовуємо тренувальний набір для навчання моделі SVM. На наступному кроці дані очищуються та стандартизуються для подальшого використання в алгоритмах. Опис товарів проходить лінгвістичне опрацювання для виділення ключових характеристик. Шляхом **токенізації** розділяємо текст на окремі слова або фрази. Перетворення тексту в числові вектори відбувається за допомогою методів TF-IDF (Term Frequency-Inverse Document Frequency) або Word2Vec.

**Далі проводимо навчання моделі** на тих даних, де є відповідні HS-коди, і зможе використовувати цю інформацію для прогнозування кодів на нових товарах.

**Присвоєння кодів HS чи HTS для нових товарів** відбувається після тренування моделі з можливістю використовувати модель для прогнозування HS-кодів для товарів, опис яких є, але код невідомий. **Для присвоєння кодів HS** використовуються наявні описані товари з відомими кодами для навчання класифікаційної моделі.

Метод Support Vector Machines (SVM) застосовуємо для аналізу та класифікації товарів за кодами HS та HTS у системах митної класифікації. Використовуючи SVM, моделі можуть навчатися на великих обсягах даних, що містять характеристики товарів і їхні відповідні коди. Це дає змогу ефективно виявляти схожості та відмінності між товарами, зокрема в складних випадках, коли товар може підпадати під декілька категорій. SVM допомагає мінімізувати помилки в класифікації, оптимізуючи процес митного опрацювання.

Модель навчається класифікувати товари за їхніми описами та іншими характеристиками. Використовуючи історичні дані, алгоритми класифікації навчаються на вхідних текстових і числових характеристиках товарів. Моделі аналізують зв'язок між текстовим описом та відповідним HS-кодом. Коли новий товар надходить до системи, алгоритм автоматично генерує пропозицію щодо HS- або HTS-коду на основі опису та інших характеристик. Якщо система невпевнена у результаті, вона може надати кілька варіантів коду для подальшого вибору користувачем. Результати автоматичної класифікації перевіряються митними експертами. Алгоритми можуть самонавчатися на основі нових випадків, що покращує точність з часом. Чим більше описів товарів із відомими кодами буде використано для навчання, тим кращі результати отримає модель для нових даних.

### Аналіз прикладу для прогнозування HS-кодів лише на основі опису товару

Припустімо, у нас є набір даних, який містить товари з відомими HS-кодами і великий масив товарів, який містить лише їхні описи без кодів. Ми можемо розділити дані так, щоб модель навчилася на відомих кодах і потім передбачила відсутні.

#### Наведемо код мовою python:

```
import pandas as pd
# Приклад даних
data_with_hs = {'description': ["Пластикові деталі для автомобілів",
                                "Алюмінієвий профіль для вікон",
                                "Літієві батареї для смартфонів",
                                "Частини і компоненти двигунів"],
                'hs_code': [3926, 7604, 8507, 8409]}
data_without_hs = {'description': ["Комп'ютерні миші бездротові",
                                    "Пластикові коробки для зберігання",
                                    "Батареї для електроінструментів"]}
```

```
df_with_hs = pd.DataFrame(data_with_hs)
```

```
df_without_hs = pd.DataFrame(data_without_hs)
```

Далі векторизуємо текстовий опис товарів, щоб перетворити їх у числові дані, придатні для SVM.

Наведемо код мовою python

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
# Об'єднуємо дані
X = df_with_hs['description']
y = df_with_hs['hs_code']
# Векторизація тексту
tfidf = TfidfVectorizer()
X_tfidf = tfidf.fit_transform(X)
```

# Розподіл на тренувальний і тестовий набір

```
X_train, X_test, y_train, y_test = train_test_split(X_tfidf, y, test_size=0.3, random_state=42)
```

**Проведемо навчання моделі SVM.** Модель SVM навчиться на товарних описах з відомими HS-кодами.

Наведемо код мовою python

```
from sklearn.svm import SVC
# Створюємо і навчаємо модель SVM
svm_model = SVC(kernel='linear')
svm_model.fit(X_train, y_train)
```

**Для визначення кодів для нових товарів** використовуємо навчену модель для прогнозування HS-кодів для товарів, де є тільки описи.

Наведемо код мовою python

```
# Векторизація нових товарних описів без HS-кодів
new_descriptions_tfidf = tfidf.transform(df_without_hs['description'])
# Прогнозування HS-кодів
predicted_hs_codes = svm_model.predict(new_descriptions_tfidf)
# Додаємо передбачені коди до таблиці з товарами
df_without_hs['predicted_hs_code'] = predicted_hs_codes
print(df_without_hs)
```

#### Виводимо результат:

##### description predicted\_hs\_code

0	Комп'ютерні миші бездротові	8507
1	Пластикові коробки для зберігання	3926
2	Батареї для електроінструментів	8507

У цьому прикладі модель подала такий результат класифікації, що:

- комп'ютерні миші відносяться до групи 8507 (літієві батареї),
- пластикові коробки – до групи 3926 (пластикові вироби),
- батареї для електроінструментів – також до 8507 (батареї).

Якщо навчальний набір буде містити більше даних і деталізовані описи, точність моделі значно покращиться.

**Використання моделі сприяє зниженню кількості помилок** у класифікації товарів завдяки автоматичному пошуку та використанню актуальних даних, **зменшенню часу на класифікацію** товарів, оскільки процес, який раніше займав кілька годин, тепер виконується за кілька хвилин або навіть секунд, **підвищенню точності** завдяки використанню алгоритмів, які здатні аналізувати велику кількість характеристик товару одночасно. **Гнучкість системи** дає змогу легко інтегрувати нові товари та швидко адаптуватися до змін у митних правилах.

Інформаційна система, що використовує алгоритми класифікації, може бути інтегрована з національними та міжнародними митними платформами, що дає змогу автоматично оновлювати дані про зміни в HS- та HTS-кодах та надсилати автоматизовані митні декларації, зменшуючи час на ручну обробку документів.

**Переваги використання SVM** можна окреслити як **високу точність класифікації**, особливо для роздільних класів, **ефективність** для великих розмірностей та складних даних, **чітка узагальненість** для нових даних, оскільки маржа забезпечує мінімізацію помилок, **гнучкість** завдяки ядрам, що дає змогу працювати з лінійно і нелінійно розділеними даними.

### Виклики при впровадженні машинного навчання

Серед недоліків використання методу можемо відзначити такі:

- велика обчислювальна складність при роботі з великими наборами даних;
- чутливість до вибору ядра та його параметрів;
- не коректно працює у робочих середовищах з багатьма шумовими даними, оскільки схильний до перенавчання при занадто високих значеннях параметра C.

SVM є потужним інструментом для задач класифікації, який забезпечує високу точність та ефективність завдяки оптимізації маржі та використанню ядрових методів. Хоча він має певні обмеження, його здатність до узагальнення на нові дані робить його ідеальним вибором для багатьох реальних застосувань, зокрема текстове опрацювання, розпізнавання зображень, класифікацію біологічних даних тощо. Хоча машинне навчання пропонує значні переваги, його інтеграція в митні системи не є безпроблемною. Основні виклики пов'язані з таким:

- алгоритми машинного навчання повинні бути адаптовані до специфічних завдань митної класифікації, що може вимагати значних зусиль у розробці і налаштуванні;
- для ефективного навчання машинних алгоритмів потрібні високоякісні дані. Неповні або неточні дані можуть вплинути на точність класифікації;
- інвестиції в розробку та впровадження технологій машинного навчання можуть бути значними, що може стати бар'єром для деяких митних органів.

### Перспективи розвитку і рекомендації

Використання машинного навчання та алгоритмів класифікації відкриває можливості для більш широкої автоматизації митних процедур, залучаючи:

- прогнозування митних ризиків на основі аналізу історичних даних;
- виявлення аномалій та можливих порушень на етапі класифікації товарів.

Алгоритми класифікації на основі машинного навчання та опрацювання природної мови значно спрощують та оптимізують процес роботи з HS- та HTS-кодами, даючи змогу логістичним компаніям підвищувати точність і ефективність класифікації товарів, знижувати витрати та час на виконання митних процедур, а також підвищувати гнучкість у взаємодії з митними органами. У майбутньому використання машинного навчання в митній класифікації має потенціал для подальшого вдосконалення [5]. Можливі напрями розвитку:

- покращення алгоритмів: розроблення нових алгоритмів машинного навчання та їхня адаптація до специфічних потреб митної класифікації;
- інтеграція з наявними системами: поступове впровадження машинного навчання в наявні митні системи для покращення ефективності;
- підвищення якості даних: збір та опрацювання даних високої якості для покращення результатів машинного навчання.

Рекомендації містять необхідність проведення додаткових досліджень, тестування різних алгоритмів і забезпечення підтримки з боку державних і приватних організацій для реалізації нових технологій у митній сфері [6]. Отже, стаття має на меті внести значний внесок у розуміння і вдосконалення процесу митної класифікації через застосування сучасних IT-технологій, що дасть змогу підвищити точність, швидкість і ефективність митних операцій.

### Висновки

Аналіз використання алгоритмів підбору HS-кодів вказує на значний потенціал технологій машинного навчання для покращення процесу митної класифікації. Виявлено, що різні підходи, такі як класифікаційні алгоритми на основі тексту, методи машинного навчання, опрацювання природної мови (NLP) та гібридні методи, можуть суттєво підвищити точність і швидкість визначення правильних HS-кодів для товарів.

Оцінка переваг і недоліків інтеграції IT-технологій у митну класифікацію є наступним етапом. Це передбачає аналіз, як нові технології можуть підвищити ефективність процесів, а також виявлення потенційних труднощів, що можуть виникнути при їх впровадженні.

У статті розроблено рекомендації для вдосконалення процесів митної класифікації, заснованих на результатах проведеного аналізу. Пропозиції будуть спрямовані на оптимізацію митних процедур через інтеграцію машинного навчання і подолання виявлених проблем.

У статті проведено аналіз можливостей використання методу Support Vector Machines (SVM) для автоматизації класифікації товарів за HS-кодом (Harmonized System) та HTS-кодом (Harmonized Tariff Schedule) на основі текстових описів. Основними результатами дослідження є такі висновки:

1. Метод опорних векторів (SVM) показав свою високу ефективність у вирішенні задач класифікації товарів за кодами HS та HTS на основі їхніх текстових описів. Це стало можливим завдяки його здатності знаходити гіперплощину, яка оптимально розділяє різні категорії товарів.

2. Одним з ключових чинників успіху моделі є якісна підготовка навчальних даних. Використання текстових описів товарів потребує ретельної опрацювання, як-от: токенизація, векторизація (зокрема TF-IDF) та видалення зайвих елементів. Чим більш деталізовані описи товарів, тим вищу точність показує модель.

3. Використання методів машинного навчання, наприклад SVM, відкриває можливості для автоматизації митних операцій. Автоматична класифікація товарів за допомогою SVM може значно скоротити час опрацювання митних декларацій і знизити ймовірність людських помилок.

4. Впровадження методів машинного навчання в системи митної класифікації може підвищити ефективність опрацювання товарних описів та збільшити рівень точності присвоєння кодів HS / HTS. Це дає змогу не лише пришвидшити процеси, але й поліпшити дотримання митних вимог завдяки більш точній класифікації.

5. Незважаючи на позитивні результати, метод SVM може потребувати значних обчислювальних ресурсів для великих наборів даних. Окрім того, точність моделі залежить від якості та кількості вхідних даних, що вимагає регулярного оновлення та вдосконалення навчальних наборів.

Отже, метод Support Vector Machines має значний потенціал для покращення процесів митної класифікації товарів, однак для повноцінної реалізації цього потенціалу необхідна подальша оптимізація та розширення навчальних моделей.

#### СПИСОК ЛІТЕРАТУРИ

1. Крупа, С., Кривенчук, Ю. (2023). *Засоби покращення автоматизованого підбору HS-коду*, 87. [https://science.lpnu.ua/qm-2023/proceedings\\_](https://science.lpnu.ua/qm-2023/proceedings_)
2. Крупа, С., Кривенчук, Ю. (2024). *Огляд можливості покращення автоматизованого підбору HS-коду з використанням методів машинного навчання для оптимізації митного класифікаційного процесу*, 46–149. <https://doi.org/10.31891/2307-5732-2024-333-2-23>
3. Ding, L., Fan, Z., Chen, D (2015). *Auto-categorization of HS code using background net approach*, 1462–1471. <https://doi.org/10.1016/j.procs.2015.08.224>
4. Mohammed, M., Khan, M. B., Bashier, E. B. M. (2020). *Machine learning – Algorithms and applications*. CRC Press. <https://doi.org/10.1201/9781315371658>
5. World Customs Organization (WCO). (n.d.). *What is the Harmonized System (HS)*. <http://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx>
6. Chary Deekshith P., Singh R. P. (2020). Review on Advanced Machine Learning Model: Scikit-Learn. *International Journal of Scientific Research and Engineering Development (IJSRED)*, 3(4), 526–529. <https://www.researchgate.net/publication/344285242>

#### REFERENCES

1. Krupa, S., Krivenchuk, Yu. (2023). *Means of improving automated selection of HS code*, 87. [https://science.lpnu.ua/qm-2023/proceedings\\_](https://science.lpnu.ua/qm-2023/proceedings_)
2. Krupa, S., Krivenchuk, Yu. (2024). *Review of the possibility of improving automated HS code selection using machine learning methods to optimize the customs classification process*, 46–149. <https://doi.org/10.31891/2307-5732-2024-333-2-23>
3. Ding, L., Fan, Z., Chen, D. (2015). *Auto-categorization of HS code using background net approach*, 1462–1471. <https://doi.org/10.1016/j.procs.2015.08.224>
4. Mohammed, M., Khan, M. B., Bashier, E. B. M. (2020). *Machine learning – Algorithms and applications*. CRC Press. <https://doi.org/10.1201/9781315371658>
5. World Customs Organization (WCO). (n.d.). *What is the Harmonized System (HS)*. <http://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx>
6. Chary Deekshith P., Singh R. P. (2020). Review on Advanced Machine Learning Model: Scikit-Learn. *International Journal of Scientific Research and Engineering Development (IJSRED)*, 3(4), 526–529. <https://www.researchgate.net/publication/344285242>

**ANALYSIS OF THE USE OF HS AND HTS CODES  
IN CUSTOMS CLASSIFICATION SYSTEMS: CHALLENGES  
AND OPPORTUNITIES OF INTEGRATION OF IT TECHNOLOGIES**

**Stepan Krupa<sup>1</sup>, Nataliia Kunanets<sup>2</sup>**

<sup>1</sup> National University “Lviv Polytechnic”,  
Artificial Intelligence Systems, Lviv, Ukraine,

<sup>2</sup> National University “Lviv Polytechnic”,  
Information Systems and Networks Department, Lviv, Ukraine,

<sup>1</sup> E-mail: [stepan.m.krupa@lpnu.ua](mailto:stepan.m.krupa@lpnu.ua), 0009-0000-2074-9762

<sup>2</sup> E-mail: [Nataliia.E.Kunanets@lpnu.ua](mailto:Nataliia.E.Kunanets@lpnu.ua), 0000-0003-3007-2462

© *Krupa S., Kunanets N., 2024*

The peculiarities of the use of the harmonized system of description and coding of goods, the harmonized tariff system of codes in modern customs classification systems are analyzed. Special attention is paid to the challenges that arise when applying these codes, in particular due to the complexity of the product nomenclature, as well as the variety of product descriptions. In addition, the possibilities of integrating IT technologies, machine learning and artificial intelligence methods to automate and optimize customs classification procedures are being explored. Prospects for increasing the accuracy and efficiency of the work of customs authorities due to the implementation of innovative solutions are considered. It is also important to note that classification systems may differ between countries, making it difficult to unify the process internationally. This becomes a serious obstacle to effective customs activity. Thanks to machine learning and analysis of large volumes of data, customs authorities can more effectively detect discrepancies and optimize work with commodity codes. The implementation of such innovative solutions will help to improve the accuracy and speed of work of customs services, which, in turn, will contribute to the transparency and efficiency of international trade.

**Keywords:** HS codes, HTS codes, customs classification, IT technologies, automation, machine learning.