

**С. О. Максимчук**, магістрант  
Університету митної справи та фінансів  
**Л. В. Кабак**, кандидат технічних наук,  
доцент кафедри програмного  
забезпечення комп'ютерних систем  
Національного технічного університету  
“Дніпровська політехніка”  
**Б. І. Мороз**, доктор технічних наук,  
професор кафедри програмного  
забезпечення комп'ютерних систем  
Національного технічного університету  
“Дніпровська політехніка”

### **ВИКОРИСТАННЯ СУЧАСНИХ ТЕХНІК DATA MINING В МИТНІЙ СПРАВІ УКРАЇНИ**

*Розглянуто техніки й алгоритми інтелектуального аналізу даних, проаналізовано актуальність їх використання для ведення бізнесу і державного управління за сучасних умов розвитку інформаційних технологій та їх впровадження з метою автоматизації різноманітних сфер людської діяльності. Проаналізовано сучасні тенденції застосування технології інтелектуального аналізу даних в митній справі, інтеграції цієї технології в системи аналізу ризиків для виявлення шахрайської діяльності, а також для отримання інших потенційно корисних прогнозів на основі аналізу раніше здобутих даних. Запропоновано приклад одного з можливих варіантів використання інтелектуального аналізу даних у системах контролю ризиків, а саме аномальний аналіз. Для цього використано базу даних Oracle з опцією Advanced Analytics, техніку Anomaly Detection, створено прототип таблиці з даними про операції експорту й імпорту товарів.*

*Ключові слова: база даних Oracle; інтелектуальний аналіз даних; виявлення аномалій; фіскальна служба; система контролю ризиків.*

*Рассмотрены техники и алгоритмы интеллектуального анализа данных, проанализирована актуальность их использования для ведения бизнеса и государственного управления в современных условиях развития информационных технологий и их внедрения с целью автоматизации различных сфер*

© С. О. Максимчук, Л. В. Кабак, Б. І. Мороз, 2019

---

человеческой деятельности. Рассмотрены и проанализированы современные тенденции использования технологии интеллектуального анализа данных в таможенной деятельности, интеграции данной технологии в системы анализа рисков с целью обнаружения мошеннической деятельности, а также для получения других потенциально полезных прогнозов на основе анализа ранее добытых данных. Предложен пример одного из возможных вариантов использования интеллектуального анализа данных в системах контроля рисков, а именно аномальный анализ. Для этого была использована СУБД Oracle с опцией *Advanced Analytics*, техникой *Anomaly Detection* и создан прототип таблицы с данными по операциям экспорта и импорта товаров.

Ключевые слова: база данных Oracle; интеллектуальный анализ данных; выявление аномалий; фискальная служба; система контроля рисков.

*Nowadays using of data mining analytics is actual question. Data mining techniques and their algorithms are becoming more common and necessary. Data warehouses stores very big amount of data at the moment and their volume increases every day. Performing analysis of data with data mining techniques and algorithms allows to detect new potentially helpful and beneficial patterns in data that can be used to make prediction. So getting results of predictive analytics helps to make right business decisions for entrepreneur and solve many issues for government institution.*

*Customs of developed countries started using data mining through integration of data mining tools in the risk management systems to detect fraud behavior and get other useful information from data as well. It is possible to find some study related to this issue.*

*Risk management systems which are used by Ukrainian customs does not support data mining tools. Simultaneously, Ukrainian customs uses Oracle Database as well. So using data mining techniques and algorithms build-in Oracle Database is obvious step to extend the available risk management system and keep up to date. So this work reveals modern tendencies and importance of integration data mining technology in business and public administration.*

*In this article have been considered data mining techniques, modern tendencies of the usage data mining technology by customs, integration of such a technology in the risk management systems so as to detect fraud behavior and get other potentially useful predictions based on data has been collected before. You can find brief descriptions of data mining techniques such as anomaly detection, classification, clustering, regression, anomaly detection in this article as well as real-world examples how each of this techniques can be applied for customs need. Also here have been proposed, designed, implemented and described real scenario of usage Oracle Data Mining in customs, in particular usage of anomaly detection*

---

*technic for analysis tables which store data of export and import operations in order to detect fraud behavior. The results have been described in this work and it have been represented graphically.*

**Key words:** *Oracle Database; data mining; anomaly detection; customs; risk management system.*

**Постановка проблеми.** Нині майже кожна фірма чи організація використовує базу даних для збереження, обробки й накопичення інформації, інколи навіть не одну. Аналізуючи раніше зібрану інформацію, можна отримувати корисні висновки і прогнози для використання у розвитку фірми або організації або для передбачення негативних явищ з метою запобігання їм. Для цього й застосовується інтелектуальний аналіз даних. Сучасні техніки, алгоритми й інструменти для його реалізації стають дедалі необхіднішими, розвинутими й гнучкими, а технологія – більш поширеною і запитуваною, ніж кілька років тому.

Зазначена технологія може успішно використовуватись як для ведення бізнесу і прийняття бізнес-рішень фірмами будь-якого розміру, так і для запровадження в державних установах. Фіскальна служба не виняток у цьому списку. Інтелектуальний аналіз даних можна використовувати для отримання різних прогнозів. Зокрема, така технологія інтегрується до системи контролю ризиків, наприклад, використовуючи техніку аномального аналізу та її алгоритми для виявлення шахрайської діяльності або інших аномалій. І це не єдиний можливий варіант використання технології інтелектуального аналізу даних. Розвинуті країни вже успішно використовують можливості інтелектуального аналізу даних, досягаючи необхідних результатів.

Фіскальна служба України має на меті відповідати світовим стандартам, для чого вона розвивається і реформується для досягнення цієї мети в нинішні часи. Вона також використовує систему контролю ризиків, яка побудована з використанням технології, відмінної від Data Mining. У сучасному світі той, хто перестає розвиватись, програє. Тому питання переходу на нові технології, техніки, алгоритми й інструменти інтелектуального аналізу даних залишається актуальним.

**Аналіз останніх досліджень і публікацій.** Нині на ринку існує чимало програмного забезпечення, середовищ розробки та інструментів для здійснення інтелектуального аналізу даних. Вони дають змогу використовувати всі переваги технік і алгоритмів Data Mining як у вигляді графічного інтерфейсу, так і з використанням мов програмування. Прикладами таких продуктів є: Oracle Data Mining, Microsoft SQL Server Integration Services, Rapid Miner Studio, STATISTICA і багато інших [1; 2].

Технологія Data Mining може бути застосована у різних сферах. Її можуть використовувати підприємці, наукові організації і державні інституції, застосовуючи ту чи іншу техніку та її алгоритми для виконання відповідних завдань.

---

Компанія Oracle надає систему управління базами даних (далі – СУБД) із вбудованим у неї інструментарієм для інтелектуального аналізу даних. Компанії, які використовують засоби й послуги компанії Oracle, діляться своїми історіями про успіхи та описують варіанти використання засобів інтелектуального аналізу даних. Наприклад, Turkcell є найбільшою телекомунікаційною компанією в Туреччині, яка використовує Oracle Advanced Analytics та Oracle Data Mining для аналізу записів клієнтів із метою виявлення потенційних шахрайських викликів та їхніх проплачених клієнтів. Stubhub – це торговельний майданчик для продажу квитків. Компанія використовує Oracle Advanced Analytics для кращого розуміння клієнтів компанії та їхньої поведінки в Інтернеті. Національна лабораторія Argonne використовує Oracle Data Mining, щоб моделювати та прогнозувати схильність кристалізації білків із білкових послідовностей. Вчені лабораторії змогли використати Oracle Data Mining для ідентифікації набору атрибутів, що корелюють схильність білків до кристалізації, вони також використовують Oracle Data Mining для побудови необхідних моделей [3; 4].

Використання технології Data Mining в митній справі – не нове питання. Застосування технік і алгоритмів Data Mining описується в працях, автором яких є Бертран Лапорт (Bertrand Laporte). Наприклад, у дослідженні, що має назву “Inspecting less to inspect better: The use of data mining for risk management by customs administrations”, де він разом зі співавторами наводить модель системи контролю ризиків, яка побудована з використанням засобів Data Mining для виявлення шахрайської активності під час оформлення митних декларацій [5].

Використання техніки класифікації для виявлення шахрайської діяльності демонструють автори статті “A Novel Unsupervised Classification Method for Customs Fraud Detection”, наводячи відповідні розрахунки й результати [6].

Фіскальна служба використовує Автоматизовану систему аналізу та управління ризиками (далі – АСАУР), що базується на використанні алгоритмів нечіткої логіки (fuzzy logic) [7]. Водночас Єдина автоматизована інформаційна система ДФС побудована з використанням бази даних Oracle [8]. Отже, логічно було б використовувати інструменти інтелектуального аналізу даних, які входять до сучасних версій бази даних Oracle, і виконувати завдання контролю ризиками з використанням засобів Data Mining, які включені в СУБД.

Використовуючи наведені тенденції і керуючись досвідом попередніх дослідників, у статті проведено власне дослідження порушеного питання, розроблено прототип, який прозоро демонструє можливість застосування описуваної технології митницями України.

---

**Мета статті** – проаналізувати актуальність використання технології інтелектуального аналізу даних, її алгоритмів і технік, а також сучасні тенденції впровадження даної технології на митницях, її інтеграції до систем контролю ризиків митниць, розглянути можливість застосування цієї технології в системі контролю ризиків фіскальної служби України, запропонувавши можливі варіанти використання її технік та алгоритмів і продемонструвати розроблений власноруч приклад з конкретними даними, технікою аналізу, моделлю і результатом.

**Виклад основного матеріалу.** Щоб продемонструвати можливість застосування технік і алгоритмів Data Mining, було створено й налаштовано модель аналізу даних з використанням техніки Anomaly Detection та її алгоритму One-Class Support Vector Machine з використанням СУБД Oracle, котра містить опцію Advanced Analytics, частиною якої є Oracle Data Mining. Використовувалось також розширення Oracle Data Miner середовища Oracle SQL Developer для побудови моделі в графічному інтерфейсі. Було створено прототип бази даних фіскальної служби з цільовими даними для аналізу. Проектування здійснювалось із використанням середовища Star UML та Microsoft Visio. Перш ніж розглядати конкретну реалізацію, необхідно розібратися у фундаментальних поняттях.

Data Mining (інтелектуальний аналіз даних або отримання даних) – це процес виявлення в сирих, раніше накопичених даних, раніше не відомих, нетривіальних, практично корисних і доступних для інтерпретації знань, необхідних для прийняття рішень у різних сферах людської діяльності [9].

Технологія Data Mining об'єднує техніки, описані нижче й зображені на рис. 1, кожна з яких базується на одному або кількох алгоритмах:

- класифікація (classification) – зарахування об'єктів (спостережень, подій) до одного із заздалегідь відомих класів;
- регресія (regression) – установлення залежності безперервних вихідних від вхідних змінних;
- кластеризація – групування об'єктів (спостережень, подій) на основі даних (властивостей), що описують сутність цих об'єктів, у кластери. Об'єкти всередині кластера мають бути “схожими” один на одного й відрізнятися від об'єктів, що ввійшли в інші кластери. Чим більше схожі об'єкти усередині кластера й чим більше відмінностей між кластерами, тим точніша кластеризація;
- правила асоціації (Association Rules) – виявлення закономірностей між пов'язаними подіями. Прикладом такої закономірності є правило, яке вказує, що з події X впливає подія Y. Такі правила називаються асоціативними. Вперше це завдання було запропоноване для знаходження типових шаблонів покупок, що здійснюються в супермаркетах, тому іноді його ще називають аналізом ринкового кошика (market basket analysis);
- аналіз відхилень або виявлення аномалій (Anomaly Detection) – виявлення найбільш нехарактерних шаблонів (аномалій) у даних [9].

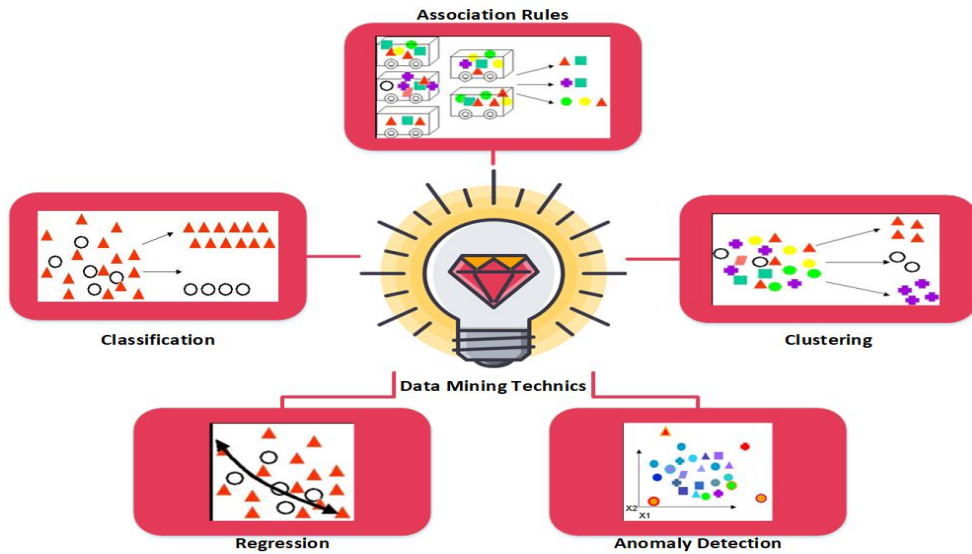


Рис. 1. Техніки Data Mining

Насправді кожен із названих технік можна адаптувати для використання в митній справі. Наприклад, правила асоціації застосувати так, як їх застосовують для аналізу ринкового кошика в супермаркетах. Аналізуючи раніше провезені товари, можна побудувати асоціацію й дізнатись, які товари провозитимуться у майбутньому і з якою вірогідністю цю інформацію можна використовувати для тарифного регулювання та встановлення коректного рівня податків. Класифікацію можна використовувати для зарахування нової операції на кордоні (експорту або імпорту) до одного зі встановлених класів. Скажімо, це може бути клас потенційно шахрайських операцій, безпечних операцій і проміжний клас між першими двома. Приблизно такий сценарій описано в статті “Classification Method for Customs Fraud Detection” [6]. Регресію можна використати для аналізу в часі показників митних зборів. Кластеризація може бути використана для групування декларантів у групи за ступенем ризику. Аналіз відхилень може бути використано для визначення шахрайських операцій – аномалій під час розрахунку суми податку і його сплати, коли така сума може бути скомпрометована. Виявлення подібних записів у базі дає змогу отримати інформацію про декларанта, посадових осіб, які ці операції здійснювали, та категорію товарів. Після цього можна вжити відповідних заходів, виконати дії для запобігання таким махінаціям у майбутньому. Саме цей сценарій розглянемо далі.

Цільовими користувачами запропонованої системи стануть працівники фіскальної служби, які виконують митні формальності. Розроблювана система

---

може зменшити обсяг паперової роботи, а також допомогти виявити порушення, використовуючи сучасні технології інтелектуального аналізу даних.

На діаграмі (рис. 2) зображено варіанти використання розробленої системи для інспектора фіскальної служби. Інспектор, виконуючи митні формальності, вносить інформацію з митної декларації до бази даних, які надалі можуть бути використані для інтелектуального аналізу даних з метою виявлення ризиків і порушень. Також інспектор може використовувати наявну в базі даних інформацію про декларантів, які вже здійснювали митні операції раніше, а також інформацію системи контролю ризиків, отриману в результаті інтелектуального аналізу даних або аналізу з використанням алгоритмів нечіткої логіки (fuzzy logic).

На діаграмі (рис. 3) зображено варіанти використання розробленої системи для аналітика фіскальної служби, який, зі свого боку, може будувати нові моделі інтелектуального аналізу даних, вносити правки в наявні моделі аналізу даних, аналізувати отримані результати, оцінювати ефективність моделей на основі цих результатів, представляти результуючі дані в зручній для подальшого використання формі – виконувати фільтрацію даних, налаштовувати автоматизацію запуску роботи моделей, коли цільовий набір даних для аналізу розширюється з внесенням нової інформації інспекторами, наприклад, аналітик може реалізувати автоматичний запуск моделі на виконання з певною періодичністю, налаштувавши Oracle Scheduler.

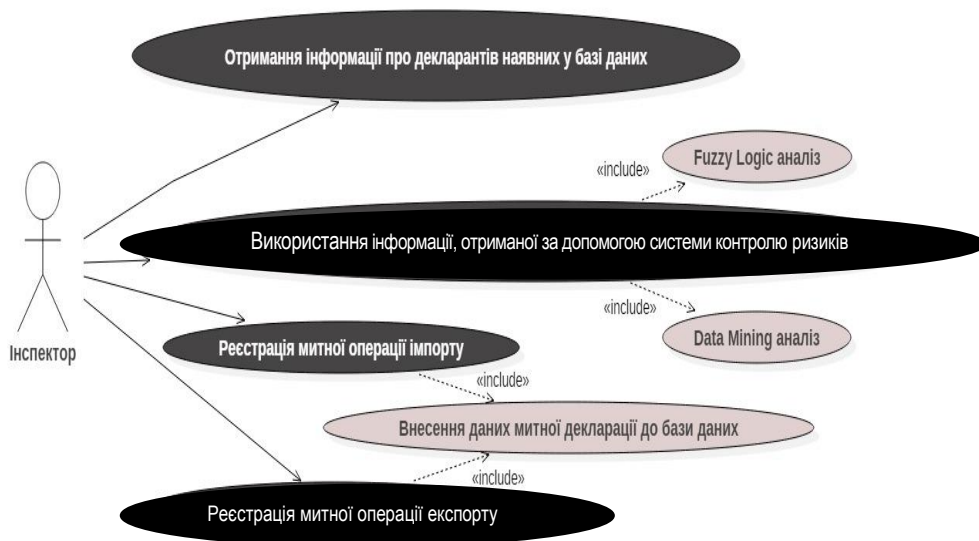


Рис. 2. Діаграма варіантів використання розробленої системи для інспектора фіскальної служби

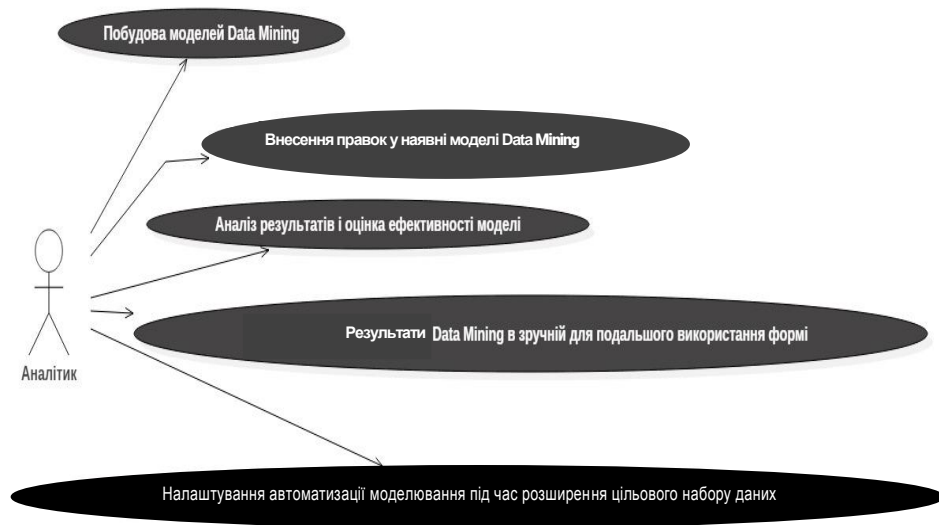


Рис. 3. Діаграма варіантів використання для аналітика фіскальної служби

Скориставшись функціональною моделлю SADT, можна відобразити структуру розроблюваної системи. Схеми в нотації IDEF0 дають змогу відобразити функції, які система дозволяє автоматизувати, зв'язки між ними, а також вхідні параметри, параметри управління, механізми, що впливають на виконання функції, результатами якої є певні вихідні параметри.

На рис. 4 зображено контекстну діаграму оформлення митної процедури, яка описує процес у цілому.

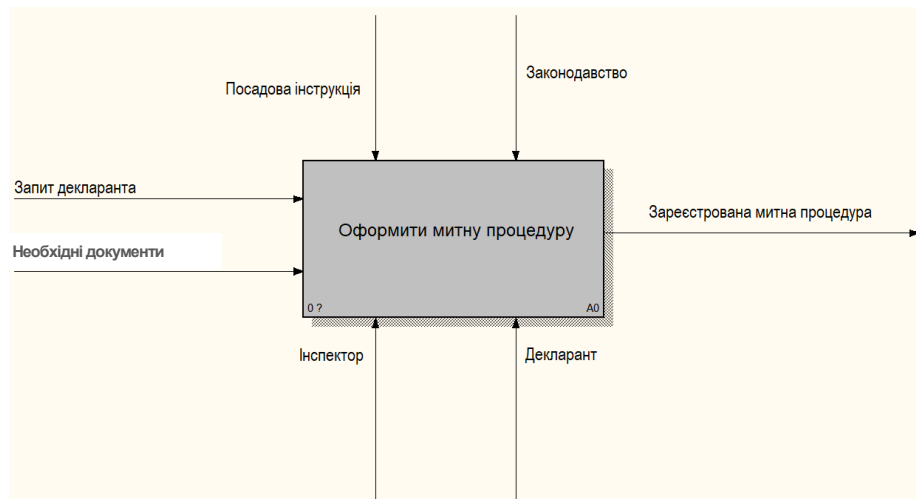


Рис. 4. Контекстна діаграма оформлення митної процедури



На рис. 5 зображено діаграму декомпозиції процесу (рис. 4) для більшої деталізації та чіткості. З метою реалізації процесу оформлення митної процедури виконуються такі функції:

- перевірка декларанта – використовується Єдина автоматизована інформаційна система ДФС і Система контролю ризиками, які включають в себе інструменти й техніки інтелектуального аналізу даних;
- виконання додаткових дій – залежно від результатів перевірки виконуються певні дії для покращання процесу митного оформлення або для запобігання правопорушенням;
- реєстрація митної процедури – оформлення митної декларації, інших необхідних документів, внесення даних до бази, яка є частиною ЄАІС. Внесені дані надалі можуть бути використані для інтелектуального аналізу даних.

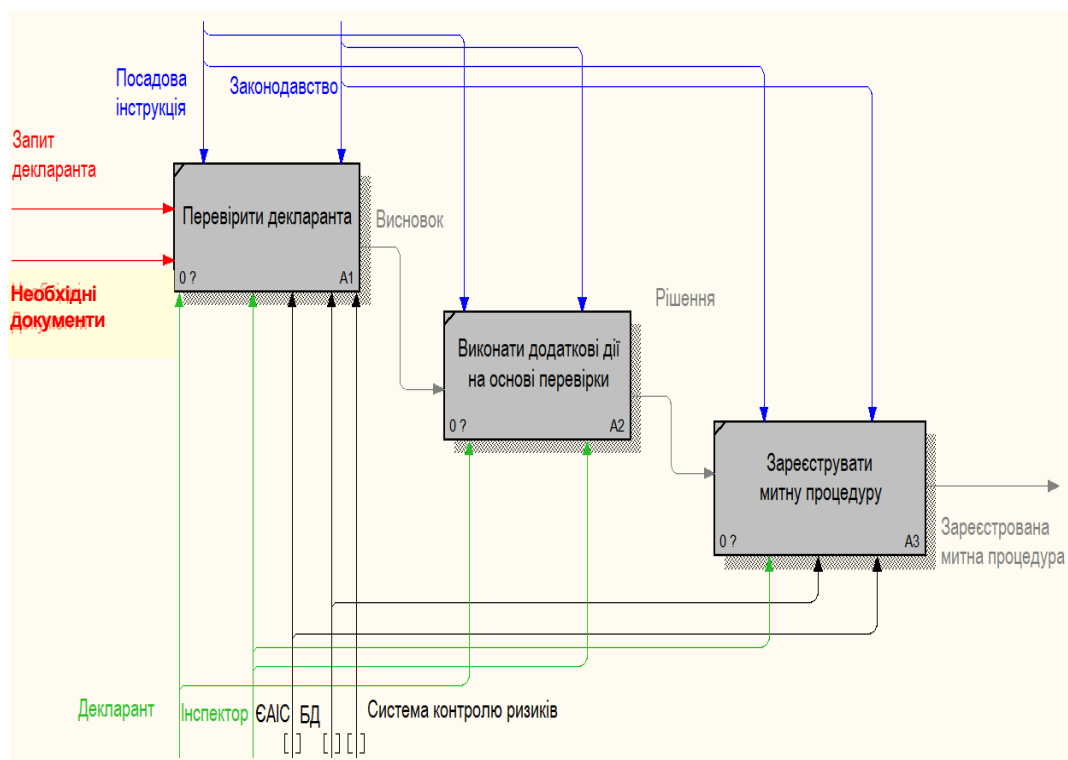


Рис. 5. Декомпозиція A0 процесу оформлення митної процедури

На рис. 6 зображено діаграму Data Flow Diagram (далі – DFD) для функції перевірки декларантів. Вона детально описує вказану функцію і відображає зв'язок виконуваних дій із даними системи.

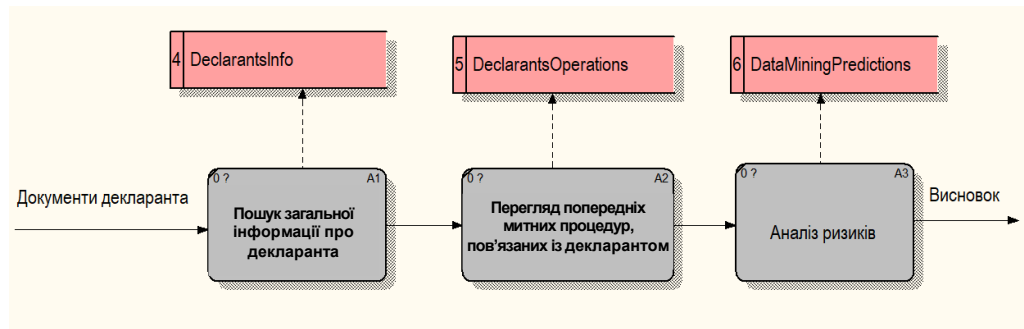


Рис. 6. Діаграма DFD перевірки декларантів

Для демонстрації використання техніки Anomaly Detection було створено невелику базу даних з декількох таблиць. ER-модель зображено на рис. 7.

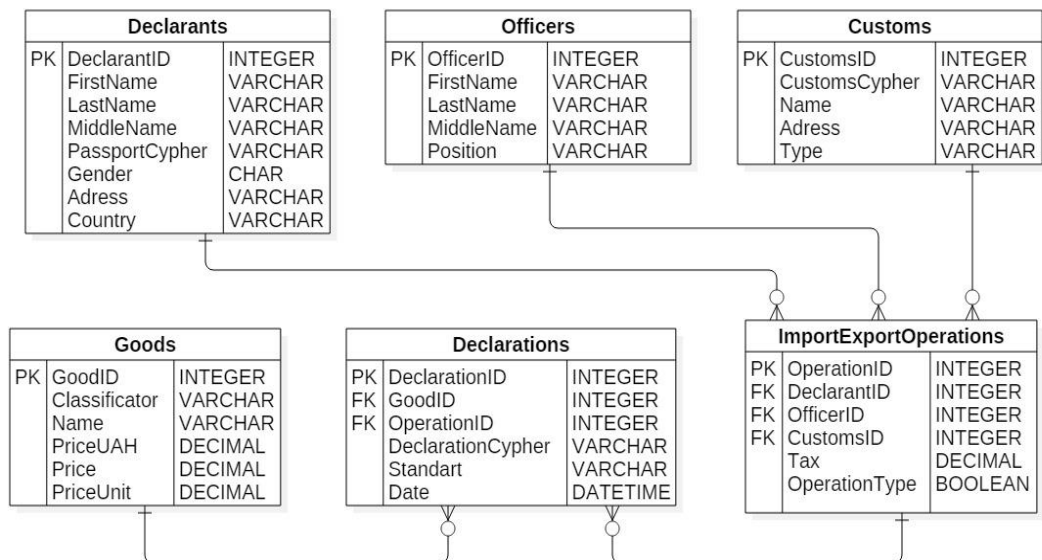


Рис. 7. ER-модель бази даних експортних та імпортних операцій виконано за принципом зірки (star schema)

Дані таблиці пов'язані таким чином, що відповідають принципу зірки (star schema), коли в центрі є головна таблиця, з якою пов'язана решта, за допомогою механізму зовнішніх ключів. На рис. 7 зображено шість таблиць:

---

• Import Export Operations – містить дані про операції імпорту та експорту, дані щодо сплати податків. Така таблиця є цільовим джерелом даних для аналізу;

- Declarants – зберігає відомості про декларантів;
- Officers – зберігає відомості про посадових осіб митниці;
- Customs – зберігає інформацію про митні пости України;
- Declarations – містить відомості про всі оформлені декларації;
- Goods – зберігає відомості про товари, зокрема УКТ ЗЕД, ціну товару в іноземній і національній валюті.

Перед використанням Data Mining було створено Data Mining репозиторій. Інформацію на сервері організовано таким чином: одна схема слугує для накопичення даних митниці і зберігає описані на ER-моделі вище таблиці та їх дані, друга схема слугує для маніпуляцій під час інтелектуального аналізу даних і збереження таблиць з результатами прогнозування (рис. 8). Модель використовує цільові дані в одній схемі та зберігає отримані результати (прогнози), а також таблиці з проміжними даними – в іншій схемі. Така архітектура забезпечує стабільну і безпечну роботу без втрати цінних даних.

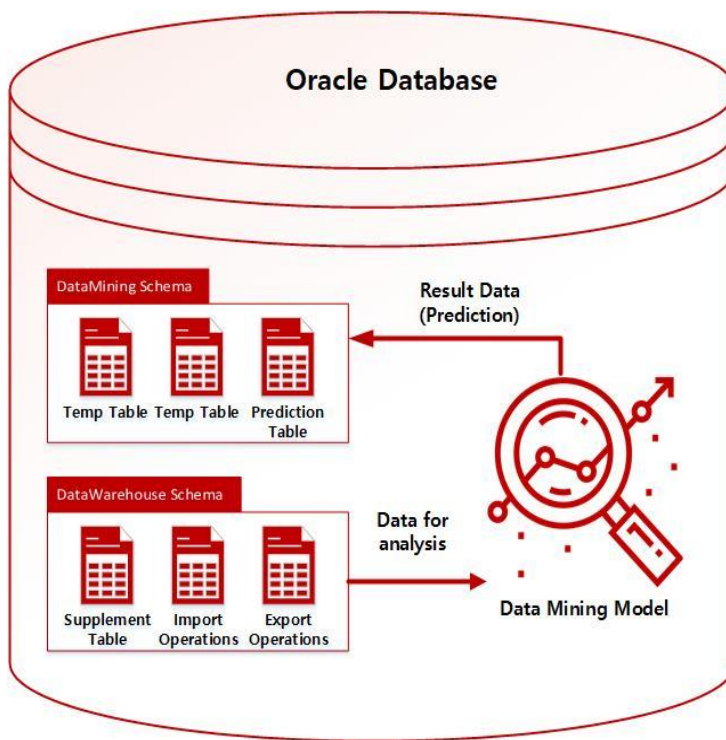


Рис. 8. Структура створеної бази даних

У процесі використання графічного інтерфейсу Oracle Data Miner було створено модель інтелектуального аналізу даних з операцій імпорту (рис. 9).

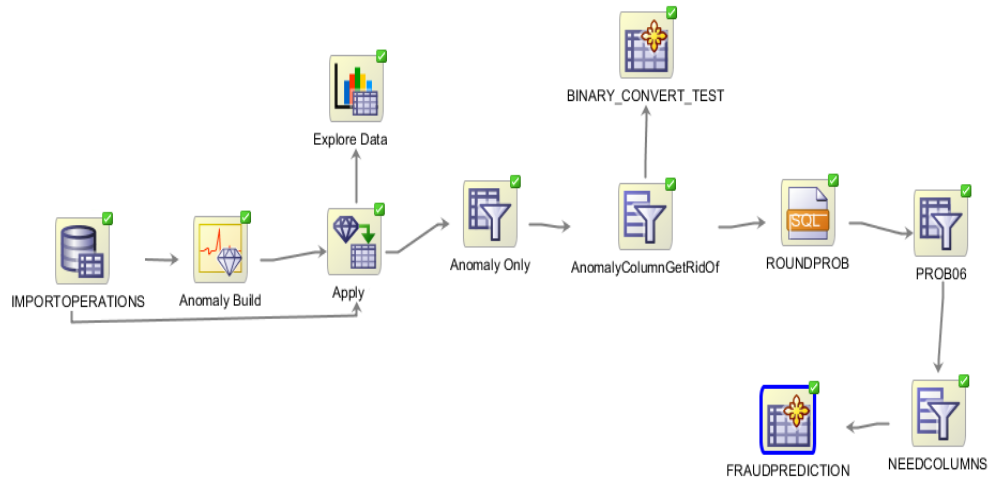


Рис. 9. Виявлення аномалій під час імпортних та експортних операцій

Джерело даних – таблиця з інформацією про дані покупців Import Export Operations. Цільовий атрибут аналізу OperationId – ідентифікатор покупця. Після того як аналіз аномалій виконано, результати аналізу накладаються на дані в таблиці Import Export Operations і, зрештою, отримуємо таблицю, де наведено інформацію про аномальні операції, процентна ймовірність того чи іншого висновку (рис. 10).

Виявити аномальні записи нескладно, найскладнішим і важливим завданням під час аналізу аномалій є встановлення чинника аномалій. Виконавши аналіз даних в результуючій таблиці, було встановлено, що аномалію викликають атрибути, які стосуються суми сплачених податків. Для отримання таблиці, яку можна використовувати для подальшого прийняття рішень, проводиться необхідна фільтрація даних. Так, можна обрати лише аномальні записи, після чого зменшити їхню кількість за процентним показником, який вказує, наскільки аномалія ймовірна в записі. Далі можна обрати необхідну інформацію, пов'язану з аномальними записами, наприклад, контактні дані декларанта, дані про інспектора, який здійснював оформлення, місце й час – таблицю з організованими таким чином даними можна використовувати надалі в роботі фіскальної служби для виявлення ризиків і порушень та проведення відповідних заходів.

ANOM_SVM_1_6_PRED	ANOM_SVM_1_6_PROB	CUST_ID	CUST_STATE_PROVINCE	CUST_LAST_NAME	COUNTRY_ID
261	1 0.5541455367317341	39 676	Maharashtra	Rhodes	52 771
262	1 0.5692326370479609	43 232	Parana	Rhodes	52 775
263	1 0.6108113617840054	25 514	South Africa	Rhodes	52 791
264	1 0.5709694436915912	47 010	England - Greater Manc...	Rhodes	52 789
265	0 0.512745159735057	50 565	Kuala Lumpur	Rhodes	52 769
266	1 0.6028507792024513	4 121	Wroclaw	Rhodes	52 786
267	0 0.5150788637473982	7 677	Andhra Pradesh	Rhodes	52 771
268	1 0.5199066252649458	11 232	Selangor	Rhodes	52 769
269	1 0.5704633666849682	14 788	Krakow	Rhodes	52 786
270	1 0.5326732174194448	18 343	Noord-Holland	Rhodes	52 770
271	1 0.5413924427666915	46 557	England - Norfolk	Rhodes	52 789
272	0 0.5004497541202058	21 898	Friesland	Rhodes	52 770
273	1 0.5191042741326525	25 455	Zuid-Holland	Rhodes	52 770
274	1 0.5389196413469113	29 010	Noord-Holland	Rhodes	52 770
275	1 0.517657796895787	32 565	Zuid-Holland	Rhodes	52 770
276	1 0.5865691757507908	37 010	Kanagawa	Ryan	52 782
277	1 0.5501512948000599	40 565	Santa Catarina	Ryan	52 775
278	1 0.5417402340112204	44 120	Zuid-Holland	Ryan	52 770
279	1 0.5636864059357586	34 403	Victoria	Ryan	52 774
280	1 0.5537763139068412	47 899	Zuid-Holland	Ryan	52 770
281	1 0.536966685649189	1 455	Noord-Holland	Ryan	52 770
282	1 0.5139799695331775	5 010	Selangor	Ryan	52 769
283	1 0.549732960702248	8 565	Rio de Janeiro	Ryan	52 775
284	1 0.5325904745946002	12 122	Gelderland	Ryan	52 770
285	1 0.5511124372937262	15 677	England - Avon	Ryan	52 789
286	1 0.5330667157633009	19 232	Limburg	Ryan	52 770

Рис. 10. Результуючі дані після застосування техніки аномального аналізу (без фільтрації)

Планувальник завдань Oracle допомагає автоматизувати роботу, тепер можна запускати моделі інтелектуального аналізу через установлені проміжки часу й отримувати оновлені дані, коли цільовий набір даних для аналізу розширюється із внесенням нової інформації інспекторами. Діаграму діяльності, що описує даний процес, зображено на рис. 11.



Рис. 11. Діаграма діяльності, автоматизація роботи з моделями Data Mining під час розширення цільового набору даних

Результати інтелектуального аналізу можна використовувати для створення програмних додатків зі зручним інтерфейсом користувача для подальшої обробки отриманих даних та їх візуалізації (рис. 12).

ID	Anomaly %	Name	Last Name	Phone	Email	Postal Code	State Province	City	Address	Birth Year	Gender
1059	0.68	Thayer	Fort	(521) 628-2723	thayer.fort@company2.example.com	57008	Berlin	Berlin	3267 Exchangeable Turnpike	1922	M
944	0.92	Everett	Lovett	(765) 344-1825	everett.lovett@company2.example.com	53574	Flevopolder	Almere	1952 Color Street	1948	M
1066	0.61	Flint	Race	(751) 593-5380	flint.race@company2.example.com	84037	Brandenburg	Cottbus	8166 Chisel Drive	1949	M
900	0.64	Franklin	Lyles	(778) 589-6802	franklin.lyles@company2.example.com	90074	Bremen	Bremen	3545 Meteorite Turnpike	1931	M
268	0.67	Gemaine	Diehl	(555) 496-6819	gemaine.diehl@company2.example.com	40804	Barcelona	Badalona	5300 Curtained Lane	1928	F
1034	0.83	Winfred	Henry	(780) 385-2563	winfred.henry@company2.example.com	47876	Aquitaine	Bordeaux	8131 Reactionaries Lane	1926	M
263	0.61	Winfred	Wensel	(771) 430-2077	winfred.wensel@company2.example.com	49147	Almeria	Almeria	2930 Sequoia Avenue	1966	F
905	0.67	Gordon	Spence	(665) 610-1481	gordon.spence@company2.example.com	53574	Flevopolder	Almere	9567 Donated Boulevard	1932	M
287	0.61	Yvonne	Lauderdale	(207) 348-4062	yvonne.lauderdale@company2.example.com	42859	Drenthe	Assen	4899 Albrecht Boulevard	1923	F
207	0.66	Idette	Forster	(407) 716-9734	idette.forster@company2.example.com	40804	Barcelona	Badalona	2034 Maddest Street	1933	F
461	0.65	Iris	Spooner	(527) 778-3205	iris.spooner@company2.example.com	88512	Dublin	Blackrock	11340 Supematural Boulevard	1954	F
945	0.62	Jason	Barnes	(715) 976-1258	jason.barnes@company2.example.com	39444	North Jutland	Aalborg	1437 Antiquy Street	1928	M
247	0.61	Jocelyn	Jewell	(401) 316-2020	jocelyn.jewell@company2.example.com	35228	Schleswig-Holstein	Bad Schwartau	782 Enduring Turnpike	1943	F
752	0.67	Joel	Mane	(532) 766-4741	joel.mane@company2.example.com	73071	Kerala	Alleppey	3046 Twigs Street	1926	M
1113	0.62	Keenan	Day	(455) 410-1127	keenan.day@company2.example.com	30568	Nordrhein-Westfalen	Aschen	8472 Barrack Boulevard	1930	M
931	0.61	Kelby	Krebs	(371) 646-9050	kelby.krebs@company2.example.com	36598	Cork	Baltimore	5028 Hermann Street	1957	M
150	0.61	Leah	Snodgrass	(628) 622-7305	leah.snodgrass@company2.example.com	34646	Girona	Girona	9899 Stockpile Street	1930	F
173	0.68	Lisette	Whitehead	(423) 210-1800	lisette.whitehead@company2.example.com	35559	Australian Capital Territory	Canberra	8783 Specifically Street	1930	F

Рис. 12. Приклад використання отриманих прогнозів у програмних додатках

---

Структуру такого додатка зображено на рис. 13. Ця архітектура відповідає сучасним тенденціям у розробці програмного забезпечення. Програмний додаток складається з трьох основних рівнів:

- Presentation Layer – рівень інтерфейсу користувача;
- Business Logic Layer – рівень, у якому міститься основна логіка додатка;
- Data Layer – рівень, у якому міститься логіка взаємодії з базою даних на програмному рівні.

Object-Relational Mapping (ORM) – засіб Entity Framework, слугує для взаємодії моделей таблиць у додатку з таблицями, що зберігаються в базі даних.

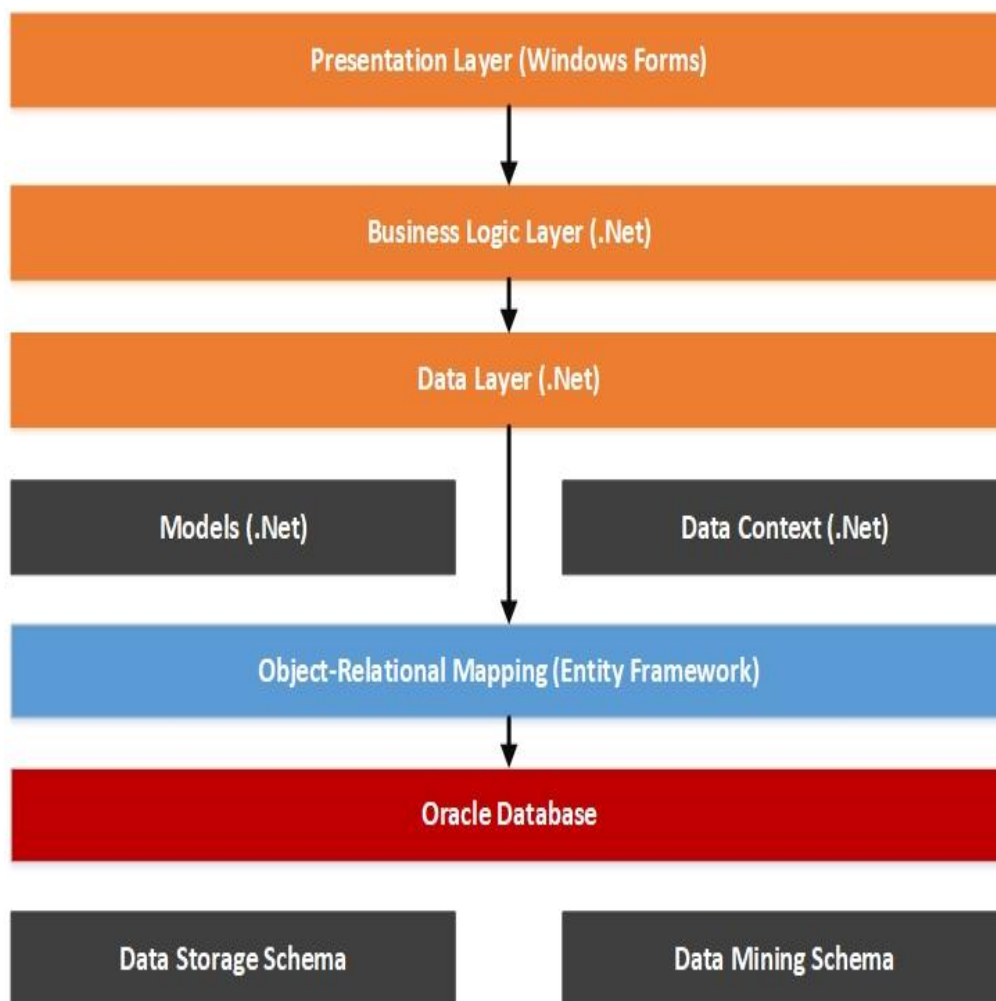


Рис. 13. Тривінева архітектура програмного додатка

---

**Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі.** У статті проаналізовано й досліджено:

- останні тенденції розвитку інформаційних технологій для інтелектуального аналізу даних;
- структуру, особливості та можливості реалізації даної технології в СУБД Oracle;
- популярність використання засобів інтелектуального аналізу підприємствами, державними установами та іншими організаціями;
- тенденції щодо запровадження використання даної технології в митній справі, а саме для вдосконалення роботи системи контролю ризиків.

У зв'язку з цим набула подальшого розвитку можливість використання технології Data Mining в митних підрозділах фіскальної служби України для аналізу операцій експорту та імпорту з метою виявлення потенційно шахрайської діяльності. Результати цього дослідження можуть бути використані для подальших досліджень і розробок, а також для запровадження використання технології митницями України, розширення можливостей АСАУР.

#### **Список використаних джерел:**

1. Top 33 Data Mining software. URL: <https://www.predictiveanalyticstoday.com/top-data-mining-software>
2. 50 top free data mining software. URL: <https://www.predictiveanalyticstoday.com/top-free-data-mining-software/#>
3. Tierney B. Predictive Analytics Using Oracle Data Miner: Develop & Use Data Mining Models in Oracle Data Miner, SQL & PL/SQL: Oracle Press, 2014. 429 p.
4. Oracle Advanced Analytics Customer Success Stories. URL: <https://www.oracle.com/technetwork/database/options/advanced-analytics/odm/odm-customers-086483.html>
5. Geourjon A.-M., Laporte B., Coundoul O., Gadiaga M. Inspecting less to inspect better: The use of data mining for risk management by customs administrations measurement Applied to Customs and Tax Administrations in Developing Countries ed. by T. Cantens, R. Ireland and G. Raballand. Washington, DC: World Bank, 2012. November.
6. A Novel Unsupervised Classification Method for Customs Fraud Detection / Habibollah R. R., Saeed A., Rahbar F., Ruhollah Rahmani and oth. // Indian Journal of science and Technology. 2015. Vol. 8(35). DOI: 10.17485/ijst/2015/v8i35/87306
7. В ДФС функціонує сучасна автоматизована система митного контролю та оформлення. URL: <http://zp.sfs.gov.ua/media-ark/news-ark/print-249074.html>



---

8. Електронна митниця. URL: [https://www.mdoffice.com.ua/pdf/\\_jeais\\_dfs\\_1602281713.pdf](https://www.mdoffice.com.ua/pdf/_jeais_dfs_1602281713.pdf)

9. *Марченко О. О., Россада Т. В.* Актуальні проблеми Data Mining: навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. Київ: Київський національний університет імені Тараса Шевченка, 2017. 150 с.

#### References:

1. Top 33 Data Mining software, available at: <https://www.predictiveanalyticstoday.com/top-data-mining-software>

2. 50 top free data mining software, available at: <https://www.predictiveanalyticstoday.com/top-free-data-mining-software/#>

3. Tierney Brendan (2014), Predictive Analytics Using Oracle Data Miner: Develop & Use Data Mining Models in Oracle Data Miner, SQL & PL/SQL: Oracle Press, 429 p.

4. Oracle Advanced Analytics Customer Success Stories, available at: <https://www.oracle.com/technetwork/database/options/advanced-analytics/odm/odm-customers-086483.html>

5. Geourjon A.-M., Laporte B., Coundoul O. and Gadiaga M. (2012), Inspecting less to inspect better : the use of data mining for risk management by customs administrations, in Reform by numbers : Measurement Applied to Customs and Tax Administrations in Developing Countries / ed. by T. Cantens, R. Ireland and G. Raballand. Washington, DC: World Bank, November.

6. *Habibollah Arasteh Rad\**, *Arash Saeed*, *Rahbar Farhad*, *Ruhollah Rahmani et al.* A Novel Unsupervised Classification Method for Customs Fraud Detection // Indian Journal of science and technology. 2015. №8(35). DOI: 10.17485/ijst/2015/v8i35/87306

7. V DFS funkcionuye suchasna avtomatyzovana systema mytnoho kontrolyu ta oformlennya [The State Customs Service has a modern automated customs control and clearance system], available at: <http://zp.sfs.gov.ua/media-ark/news-ark/print-249074.html>

8. Elektronna mytnytsya [Electronic Customs], available at: [https://www.mdoffice.com.ua/pdf/\\_jeais\\_dfs\\_\\_1602281713.pdf](https://www.mdoffice.com.ua/pdf/_jeais_dfs__1602281713.pdf)

9. Marchenko O. O. and Rossada T. V. (2017), Aktual'ni problemy Data Mining: navchal'nyy posibnyk dlya studentiv fakul'tetu komp'yuternykh nauk ta kibernetiky [Current issues of Data Mining], Tutorial for students of the faculty of computer science and cybernetics, Kyiv, 150 p.