**УДК 004.85:657.6**

**А. С. Лановик**
**А. А. Яровий**

# ВИЯВЛЕННЯ ТА АНАЛІЗ АНОМАЛІЙ У ФІНАНСОВИХ ЗВІТАХ НА ОСНОВІ ГІБРИДНОГО АНСАМБЛЕВОГО ПІДХОДУ ISOLATION FOREST ТА LSTM AUTOENCODERS

Вінницький національний технічний університет

***Анотація***

*У роботі розглядається актуальна задача автоматизованого виявлення аномалій у фінансовій звітності підприємств. Запропоновано гібридну систему на основі ансамблю алгоритмів Isolation Forest та LSTM Autoencoder, що використовує набір із 17 ознак, який інтегрує розширений набір технічних індикаторів. Проведено порівняння запропонованого підходу порівняно з існуючими комерційними рішеннями.*

**Ключові слова:** аномалії, фінансова звітність, штучний інтелект, Isolation Forest, LSTM Autoencoder, виявлення шахрайства, фінансовий аналіз.

***Abstract***

*The paper considers the urgent task of automated detection of anomalies in financial reporting of enterprises. A hybrid system based on the Isolation Forest and LSTM Autoencoder ensemble of algorithms is proposed, which uses set of 17 features that integrates an extended set of technical indicators. A comparison of the proposed approach with existing commercial solutions is provided.*

**Keywords:** anomalies, financial reporting, artificial intelligence, Isolation Forest, LSTM Autoencoder, fraud detection, financial analysis.

## Вступ

Аналіз корпоративних фінансових звітів є складним аналітичним процесом, що потребує значних витрат часу та людських ресурсів. Збільшення обсягу фінансових даних та зростаючі вимоги до ефективності аудиту роблять необхідним автоматизацію цього процесу. Існуючі інструменти виявлення аномалій зазвичай аналізують ринкові дані або фінансові звіти окремо, що знижує точність виявлення та не враховує взаємозв'язок між поведінкою ринку та фактичним фінансовим станом компанії [1].

У попередніх дослідженнях було розроблено модуль виявлення аномалій на основі алгоритму Isolation Forest із 7-ма ознаками. Система була протестована на даних фондового ринку Apple Inc. (AAPL) за 2024 рік (327 торгових днів): було виявлено 17 аномалій, Silhouette Score = 0,68 [2].

Основним обмеженням розробленого модуля є використання єдиного детектора без урахування часових залежностей та розширених фінансових показників.

## Аналіз існуючих рішень

Огляд комерційних продуктів для виявлення аномалій у фінансових даних виявив суттєві обмеження існуючого інструментарію. Результати аналізу подано в табл. 1.

Зокрема, IBM Cognos Analytics надає розширені інструменти аналізу фінансових звітів, але виявлення аномалій у цій системі реалізовано за допомогою статичних порогових правил, без використання алгоритмів машинного навчання або механізмів адаптації [3].

Система SAP Financial Compliance дозволяє аналізувати квартальні звіти та виявляти аномальні транзакції за допомогою гнучких наборів правил, але вона не враховує часові залежності між спостереженнями і не має засобів для групування подібних відхилень [4].

Oracle Financial Services пропонує комплексний моніторинг транзакцій, але потребує складніших налаштувань і не має відкритого API для інтеграції власних алгоритмів [5].

Таким чином, жоден з розглянутих продуктів не пропонує гібридного виявлення аномалій, інтеграції розширеного набору фінансових показників, а також наявності адаптивної бази знань з механізмом перевірки аналітиком.

Таблиця 1 – Порівняльний аналіз функціональних можливостей систем виявлення аномалій у фінансових даних

| Критерії | IBM Cognos | SAP Financial | Oracle FS | Isolation Forest |
|---|---|---|---|---|
| *Тип аномалії* | | | | |
| Точкові аномалії | + | частково | – | + |
| Часові аномалії | частково | – | + | – |
| *Джерело даних* | | | | |
| Квартальна звітність | – | + | – | + |
| Біржові дані | + | + | + | частково |
| *Пояснюваність* | | | | |
| Пояснення аномалій | – | – | – | частково |
| Класифікація типу | – | – | – | – |
| *Доступність* | | | | |
| Відкритий/безкоштовний | – | – | – | + |

**Результати дослідження**

У попередніх дослідженнях було розроблено програмний модуль, який виявляє аномалії у фінансових звітах за допомогою алгоритму Isolation Forest. Програмний модуль було протестовано на даних фондового ринку Apple Inc. (AAPL) за 2024 рік (327 торгових днів). Були отримані наступні результати:

- виявлено 17 аномалій (5,20% спостережень) — система правильно ідентифікувала кластери турбулентності в травні та жовтні 2024 року;
- Silhouette Score = 0,68 — висока якість кластеризації (високе порогове значення 0,7);
- автоматичний вибір параметрів за допомогою Grid Search: contamination = 0,05, n_estimators = 100;
- аномальні дні демонструють волатильність до ±6%, нормальні дні ±3%;
- графічний інтерфейс, реалізований на Tkinter та автоматичне формування текстових звітів [2].

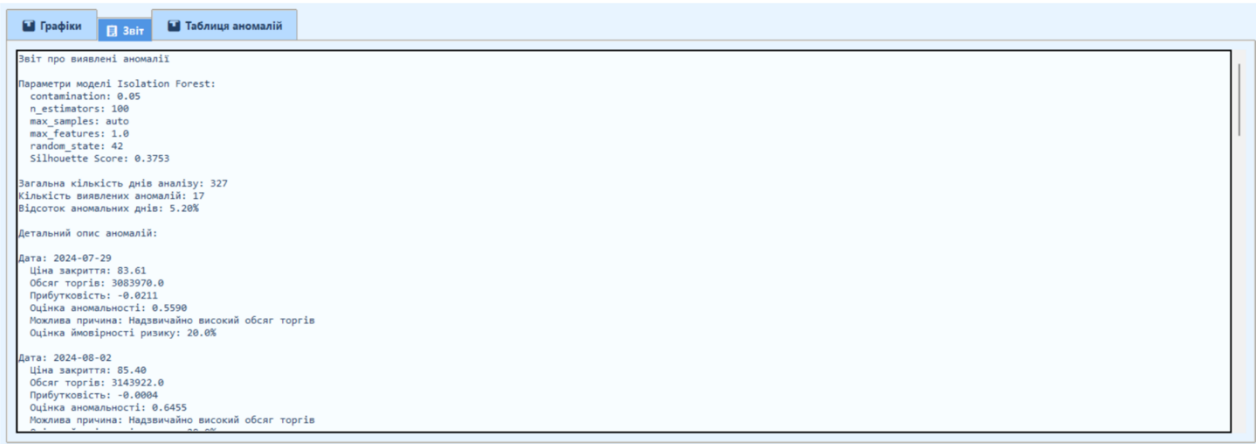


Рисунок 1 – Фрагмент звіту після завершення пошуку аномалій

Однак, було виявлено певні обмеження, які знижують практичну цінність системи:

- використання єдиного детектора (Isolation Forest) унеможливлює виявлення контекстуальних аномалій, спричинених тимчасовими залежностями між спостереженнями;
- набір із 7 характеристик не включає фундаментальні бізнес-показники, що унеможливлює розрізнення між поведінкою ринку та фактичним фінансовим станом;
- система не пояснює аналітику природу аномалії та не класифікує її тип;
- відсутній механізм накопичення знань: кожен запуск є незалежним, а досвід попередніх аналізів не фіксується.

На основі виявлених обмежень у подальших дослідженнях запропоновано три напрямки для вдосконалення.

1. Гібридний підхід пошуку аномалій. Ансамбль Isolation Forest та LSTM Autoencoder охоплює обидва класи аномалій. Кожній точці присвоюється score від 0 до 1, причому чим ближче значення до 1, тим більш точка аномальна. Остаточний score, за яким відбувається розподіл точок на нормальні та аномальні, формується за виразом (1):

$$combined\_score = IF\_score \times 0{,}5 + LSTM\_score \times 0{,}5 \quad (1)$$

2. Розширений набір ознак. Кількість ознак збільшено з 7 до 17: додано технічні індикатори та квартальні фінансові показники розповсюджені на кожен торговий день кварталу.

3. Розробка адаптивної експертної системи. База знань у форматі «умови» - «тип аномалії», де правила генеруються автоматично, а також впровадження механізму human-in-the-loop, який через діалог з аналітиком проводить верифікацію нових правил з поясненням сигналів.

## Висновки

Запропоновано напрямки удосконалення розробленого програмного продукту, що усувають основні обмеження існуючих комерційних рішень: поєднує гібридне виявлення точкових і часових аномалій, інтегрує квартальні дані фондового ринку та фінансові дані. Також такий підхід буде забезпечувати класифікацію типів аномалій та адаптивно накопичувати знання з досвіду аналітиків. На відміну від розглянутих аналогів, така система буде відкритою та розширюваною.

Майбутні дослідження будуть зосереджені на тестуванні гібридного підходу пошуку аномалій, створенні модуля експертної системи з автоматичним генеруванням правил і динамічним підходом до аналізу виявлених аномалій, розробці графічного інтерфейсу для аналітиків і тестуванні системи в цілому на реальних даних різних компаній.

## Список використаних джерел


1. AI Anomaly Detection Use Cases in Finance [Електронний ресурс]. — Режим доступу: https://www.lucid.now/blog/ai-anomaly-detection-use-cases-finance/.
2. Метод виявлення аномалій у фінансових звітах / А.С. Лановик, Я.В. Іванчук – Матеріали LIV Всеукраїнської науково-технічної конференції підрозділів ВНТУ (НТКП ВНТУ-2025). – [Електронний ресурс]. – Тип доступу: https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2025/paper/view/23674/19533
3. IBM Cognos Analytics. Forecasting and anomaly detection [Електронний ресурс]. — 2024. —Режим доступу: https://www.ibm.com/docs/en/cognos-analytics/11.2.x?topic=manuals
4. SAP. Finance NXT Anomaly Detection by Bosch Global Software Technologies [Електронний ресурс]. — Режим доступу: https://www.sap.com/ukraine/products/artificial-intelligence/partners/bosch-global-software-technologies-finance-nxt-anomaly-detection.html
5. Oracle. Anomaly Detection [Електронний ресурс]. — Режим доступу: https://docs.oracle.com/en/database/oracle/machine-learning/oml4sql/23/dmcon/anomaly-detection.html



***Лановик Анастасія Сергіївна*** — студентка групи 2КН-25м, факультет інформаційних інтелектуальних технологій та автоматизації, Вінницький технічний національних університет, м. Вінниця

***Яровий Андрій Анатолійович*** — д-р техн. наук, професор, завідувач кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця