УДК 004.8:004.9:[004.78:336.717 **DOI:** 10.31673/2412-9070.2025.050674
Б. С. КАЛИНЮК, аспірант;
ORCID: 0009-0009-9569-9304
І. В. ЗАМРІЙ, доктор техн. наук, професор;
ORCID: 0000-0001-5681-1871
А. М. КАЛИНЮК, канд. фіз.-мат. наук, доцент,
ORCID: 0000-0002-4560-3409
Державний університет інформаційно-комунікаційних технологій, Київ

ОГЛЯД СУЧАСНИХ МЕТОДІВ ВИЯВЛЕННЯ ФІНАНСОВИХ ЗЛОЧИНІВ ЗА ДОПОМОГОЮ АГЕНТІВ ШТУЧНОГО ІНТЕЛЕКТУ

У статті представлено комплексний аналіз сучасних методів виявлення фінансових злочинів із використанням агентів штучного інтелекту. Розглянуто класифікацію AI-агентів (реактивні, когнітивні, автономні, мультиагентні), особливості їхньої роботи та роль у системах фінансового моніторингу. Здійснено порівняльний аналіз rule-based підходів, методів машинного навчання, гібридних моделей, блокчейн-архітектур і графових алгоритмів. В ході дослідження виявлено, що гібридні рішення та графові нейронні мережі демонструють найвищі показники точності та повноти виявлення підозрілих транзакцій, що підтверджено узагальненими метриками з рецензованих джерел. Особливу увагу приділено застосуванню глибоких нейронних мереж, методів обробки часових рядів, Natural Language Processing (NLP) та Explainable AI (XAI), які дозволяють підвищити рівень прозорості та інтерпретованості рішень, що особливо важливо в умовах жорсткого регуляторного середовища банківської сфери. Визначено, що мультиагентний підхід, завдяки своїй здатності до паралельного аналізу, динамічної адаптації та масштабованості, є перспективним напрямом для побудови розподілених систем аналізу фінансових потоків. Поряд із перевагами, проаналізовано й основні виклики, які обмежують ефективність практичного впровадження таких систем: низька інтерпретованість моделей глибокого навчання, потреба у великому обсязі високоякісних і збалансованих даних, обчислювальні витрати, а також юридичні та етичні обмеження, пов'язані з автоматизованою обробкою персональної фінансової інформації, що підпадає під регулювання GDPR і AMLD.

У контексті подальшого розвитку запропоновано інтеграцію AI-агентів із блокчейн-мережами для забезпечення прозорості та незмінності транзакцій, застосування квантових алгоритмів для обробки складних фінансових графів та імплементацію edge computing – обробки даних на периферійних пристроях – з метою виявлення аномалій у реальному часі. Також відзначено важливість міждисциплінарного підходу, що передбачає поєднання знань із галузей штучного інтелекту, кібербезпеки, економіки та права, для розбудови ефективних систем фінансової безпеки. Таким чином, результати дослідження засвідчують, що ефективне виявлення фінансових злочинів на сучасному етапі вимагає впровадження комплексних технологічних рішень на базі агентів штучного інтелекту з акцентом на прозорість, масштабованість і відповідність нормативним вимогам. Представлені у статті висновки мають як теоретичне, так і практичне значення для подальшого розвитку систем моніторингу банківських транзакцій, а також формування політик у сфері цифрової фінансової безпеки.

Ключові слова: фінансові злочини; штучний інтелект; агенти штучного інтелекту; машинне навчання; блокчейн; графовий аналіз.

Вступ

Фінансові злочини є однією з найбільш загрозливих форм сучасної економічної дестабілізації. В умовах активної цифровізації фінансового сектору, розвитку онлайн-банкінгу, електронної комерції та криптовалют зростає кількість і складність протиправних дій у сфері обігу фінансових ресурсів. Сучасні злочинці дедалі частіше використовують інструменти цифрових технологій для відмивання коштів, фінансування тероризму, фішингу, інсайдерської торгівлі та кібератак на банківські системи. За таких умов традиційні rule-based системи виявлення шахрайства, що покладаються на фіксовані правила, дедалі частіше виявляються неефективними перед новими, гнучкими та прихованими сценаріями злочинної активності.

З огляду на зростаючу складність фінансових злочинів, дедалі більшої актуальності набуває застосування агентів штучного інтелекту (ШІ), які здатні самостійно приймати рішення, адаптуватися до нових загроз і виявляти аномалії в реальному часі. AI-агенти діють автономно, можуть обмінюватися інформацією, моделювати поведінку клієнтів і здійснювати багаторівневий аналіз транзакцій. Завдяки цьому відкривається можливість переходу від реактивних до проактивних систем фінансової безпеки.

Аналіз останніх досліджень та публікацій

У науковому та прикладному середовищі активно ведуться дослідження щодо розробки інтелектуальних систем моніторингу, здатних ефективно виявляти злочинні схеми. Зокрема, у роботі [1] показано ефективність комбінованих rule-based і машинних підходів до виявлення шахрайства, проте вказано на обмежену гнучкість таких систем при появі нових шаблонів атак. У праці [2] автори акцентують на важливості Explainable AI у сферах з підвищеною регуляторною відповідальністю (банківська справа, аудит), зазначаючи складність інтерпретації рішень глибоких моделей, що є викликом для їх практичного впровадження.

З позицій мультиагентного підходу [3] обґрунтовує переваги кооперативної взаємодії агентів при розв'язанні складних завдань, таких як виявлення шахрайських мереж. У роботі [4] описано дослідження графових нейронних мереж як інструменту для виявлення камуфльованого шахрайства у фінансових даних, наголошуючи на їхній здатності аналізувати складні зв'язки між клієнтами, рахунками та транзакціями.

Робота [5] зосереджується на часовому аналізі поведінки клієнтів із використанням LSTM-мереж, показуючи, що динаміка транзакційного профілю є критично важливою для виявлення аномалій. У контексті українського банківського середовища, в роботі [8] досліджують можливості застосування Big Data та ML у системах протидії шахрайству, відзначаючи складність, пов'язану з якістю даних, класовим дисбалансом і потребою у високій кваліфікації кадрів.

Також важливим є розгляд правових обмежень. У роботі [7] аналізуються проблеми впровадження AI у банківських структурах, зокрема у контексті відповідності GDPR та AML-директивам. У вітчизняному дослідженні [6] підкреслюють потребу у створенні спеціалізованих датасетів для тренування моделей, а також у дотриманні міжгалузевої відповідності.

Питання ефективності гібридних моделей висвітлено у роботі [15], де наголошено на доцільності поєднання експертних знань і машинного навчання для зниження кількості хибнопозитивних спрацювань. У той же час в роботі [14] розглядають блокчейн як інструмент забезпечення прозорості та незмінності транзакційної історії, проте звертають увагу на обмеження масштабованості таких рішень.

З огляду даних джерел і досліджень можна сформулювати кілька ключових проблем, які стоять перед сучасними системами виявлення фінансових злочинів, зокрема недостатня адаптивність традиційних rule-based систем до нових схем шахрайства [1], непрозорість рішень, прийнятих глибокими AI-моделями [2, 5], обмеження у якості та репрезентативності даних для навчання [6, 8], потреба в інтеграції різних технологій у єдину гнучку архітектуру [3, 15], техніко-правові виклики впровадження ШІ у фінансовому секторі [7, 14].

Завданнями дослідження є аналіз типів та особливості сучасних фінансових злочинів, оцінка ефективності rule-based, ML-, гібридних, блокчейн- та графових методів, розгляд архітектури мультиагентних систем та Explainable AI, визначення технічних та суміжних викликів впровадження ШІ у фінансову сферу, огляд перспектив розвитку AI-агентів у контексті квантових і edge-технологій.

Основна частина

Рис. 1. Схема роботи Rule-Based System

Сучасні технології аналізу фінансових операцій охоплюють широкий спектр підходів – від традиційних експертних систем до складних гібридних моделей із глибоким навчанням. У цьому розділі розглянуто найбільш популярні та ефективні методи [1, 5, 15].

Системи, що використовують методи на основі правил (Rule-Based Systems) використовують заздалегідь визначені логічні умови, які вказують, коли операція вважається підозрілою [1]. Серед переваг даного методу є простота у реалізації та прозорість причин спрацювання при блокуванні транзакцій. Серед недоліків варто виділити високу частку помилкових спрацювань, а також обмежену адаптивність до нових схем шахрайства.

На рис. 1 схематично зображено принцип роботи системи на основі правил, де підозрілі транзакції виявляються через заздалегідь визначені логічні умови.

Архітектура та кроки виконання системи складаються з декількох етапів. Первинне надходження транзакції передбачає, що кожна фінансова операція (транзакція), що відбувається в системі, автоматично передається на аналіз. Дані транзакції мажуть включати: ID клієнта, суму, валюту, тип операції, географію, час доби, канал транзакції (мобільний застосунок, банкомат, web) тощо. Попередня обробка (преобробка) передбачає нормалізацію форматів даних за потреби, вилучення базових ознак (feature extraction), таких як: середня сума операцій для клієнта, частота транзакцій у певний період, відстань між геолокаціями останніх транзакцій, тощо. Далі відбувається перевірка за правилами. До транзакції по черзі застосовуються фіксовані детекційні правила, що визначені експертами або регулятором. Якщо правило спрацьовує, транзакція: або блокується автоматично, або маркується як підозріла і передається для перегляду аналітиком. Під час генерації подій і звітності формується подія у системі, яка може бути передана у систему автоматичного формування STR (Suspicious Transaction Report) для регулятора або використовується для навчання ML-моделей (позначення «позитивних» кейсів). Усі події зберігаються у логах транзакційної системи для подальшого аудиту, навчання моделей або інспекцій з боку регулятора.

Машинне навчання та нейронні мережі для аналізу транзакцій

Алгоритми машинного навчання навчаються на історичних даних (мітках «шахрайство/не шахрайство») для автоматичного виявлення ризикованих операцій.

Основні моделі:

1) Decision Trees, Random Forest, XGBoost – ефективні для табличних даних;

2) Neural Networks (CNN, LSTM) – аналіз часових рядів та поведінкових шаблонів;

3) Autoencoders – для виявлення аномалій у складних даних.

На рис. 2 проілюстровано процес обробки транзакцій машинним алгоритмом, що навчається на історичних даних для класифікації операцій як шахрайських чи ні.

Послідовність роботи системи на основі ML починається зі збору та агрегації історичних даних. Вхідні дані: транзакції за попередній період, позначені як «шахрайські» (fraud) або «нормальні» (non-fraud). До набору можуть входити такі атрибути: ID клієнта, сума, тип операції, час, геолокація, пристрій, IP-адреса, історія транзакцій тощо. Далі відбувається попередня обробка даних, а саме очищення від шумів і пропущених значень, нормалізація числових ознак, кодування категоріальних змінних (one-hot, label encoding), балансування класів даних (наприклад, SMOTE, ADASYN), оскільки дані можуть мати дис-

баланс, наприклад, дуже мало випадків шахрайства. Наступним етапом є побудова ознак (Feature Engineering), що передбачає створення нових ознак на основі показників транзакційної поведінки клієнта: середня/максимальна сума за останні 24 години, кількість операцій на годину/день, географічна відстань між останніми транзакціями, шаблони активності (час доби, дні тижня), відхилення від звичних патернів. Далі відбувається навчання моделі. На даному етапі важливий вибір алгоритмів машинного навчання: XGBoost, Random Forest, CatBoost, Decision Trees – ефективні для табличних структурованих даних, глибокі нейронні мережі (DNN), LSTM, CNN – використовуються для аналізу часових рядів, послідовностей подій і поведінкових патернів. Навчання відбувається на тренувальному наборі, оптимізація — на валідаційному, а оцінка якості — на тестовому. Оцінка продуктивності моделі відбувається за допомогою обчислень метрик: Precision, Recall, F1-score – ключові для виявлення шахрайства; ROC-AUC, PR-AUC – важливі при дисбалансі класів; аналіз хибнопозитивних і хибнонегативних спрацювань. Навчена модель розгортається у продакшн-середовищі. Для кожної нової транзакції витягуються необхідні ознаки у реальному часі. Вони подаються на вхід моделі. Модель повертає ймовірність шахрайства. Якщо поріг перевищено, система блокує або маркує транзакцію.

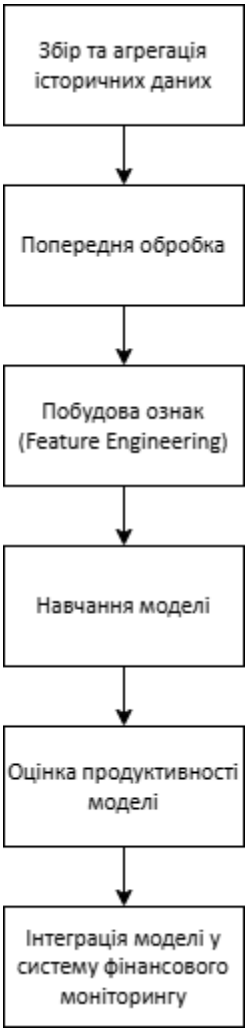

Рис. 2. Робота ML-моделі

Гібридні моделі

Гібридні моделі комбінують правила та моделі ШІ для підвищення точності виявлення шахрайства. Спочатку відбувається фільтрація підозрілих операцій за правилами. Потім відбувається застосування ШІ-моделі для уточнення оцінки ризику [15].

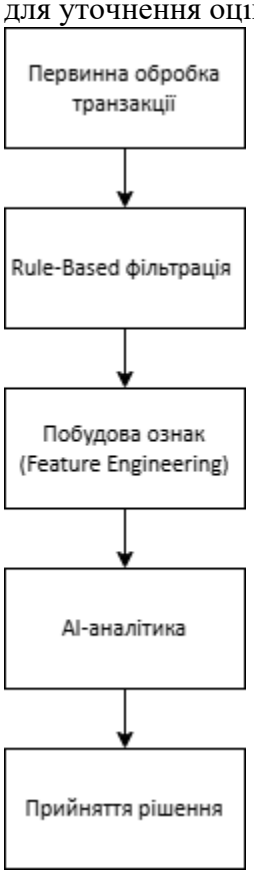

Рис. 3. Схема гібридного

На рис. 3 зображено поєднання rule-based фільтрації та моделі ШІ, що дозволяє уточнювати рівень ризику після первинної перевірки.

При застосуванні цього методу вхідна транзакція проходить через попередню перевірку на валідність: цілісність даних, перевірка на дублі, попередня класифікація за типами операцій. Далі відбувається Rule-based фільтрація і за спрацювання певного правила, транзакція передається далі на ШІ-фільтр, де використовується попередньо навчена ML або DL модель (наприклад, XGBoost, LSTM, Autoencoder) для уточнення ризику шахрайства.

Використання блокчейну для боротьби з шахрайством

Завдяки своїй прозорості, незмінності даних і децентралізованості, блокчейн дозволяє відстежувати джерело коштів (traceability), зберігати історію змін без можливості фальсифікації, забезпечувати смарт-контроль за умовами платежів [14].

На рис. 4 показано, як блокчейн забезпечує прозорість, незмінність історії транзакцій та контроль за умовами платежів за допомогою смарт-контрактів.

Блокчейн не є самостійним детектором шахрайства, але є потужною платформою для надійної реєстрації, автоматичного контролю умов транзакцій і аналітичного відстеження аномалій, особливо у поєднанні з ШІ та графовим аналізом. Найефективнішим є його використання у гібридних системах виявлення злочинних фінансових схем.

Рис. 4. Схема використання

Рис. 5. Схема використання

Графові алгоритми та аналіз соціальних зв'язків

Використовуються графові моделі (Graph Neural Networks, Graph Databases) для аналізу зв'язків між користувачами, рахунками, операціями. Такий алгоритм забезпечує виявлення аномальних шаблонів: цикли, зірки, кластерні змови та дозволяє візуалізувати складні фінансові зв'язки [4]. Схема нижче описує алгоритм ви-користання графових алгоритмів (рис. 5).

На рис. 6 представлено приклад графової моделі з виявленням аномальних патернів, таких як замкнені петлі та щільно пов'язані кластери. Графовий підхід дозволяє виявити аномальні та шахрайські патерни, що складно фіксуються традиційними rule-based або навіть ML-системами. Граф моделює взаємозв'язки між об'єктами фінансових операцій, де вузли — це користувачі, банківські рахунки або транзакції, а ребра — фінансові зв'язки між ними (перекази коштів, спільні IP-адреси, однотипні шаблони поведінки тощо). Графові структури дозволяють перейти від індивідуального аналізу транзакцій до більш системного, аналізуючи зв'язки між об'єктами фінансової взаємодії.

Графові методи – це потужний інструмент для виявлення складних, прихованих та кооперативних форм фінансового шахрайства. Вони особливо ефективні у поєднанні з ML-моделями (гібридний підхід) та блокчейн-технологіями. У системах фінансового моніторингу вони дозволяють перейти від транзакційної аналітики до повноцінного поведінкового та мережного аналізу.

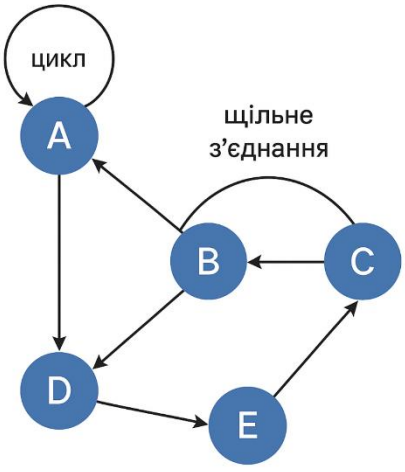

Рис. 6. Схема графової структури

Порівняльний аналіз методів. Оцінка ефективності різних підходів

Для системного розуміння ефективності та практичної доцільності впровадження різних підходів до виявлення фінансових злочинів було проведено порівняльний аналіз ключових методів, що використовуються у сучасних фінансових системах. Огляд охоплює rule-based системи, машинне навчання, гібридні моделі, технології блокчейну та графові алгоритми. Оцінювання грунтується на таких параметрах, як прозорість, адаптивність, вимоги до даних і ресурсів, масштабованість, складність реалізації та стійкість до нових схем шахрайства. Усі методи мають свої переваги та обмеження і їх ефективність значною мірою залежить від контексту застосування, наявних даних та регуляторного середовища. Узагальнені характеристики наведено у табл. 1.

Таблиця 1

Порівняння характеристик методів

Метод	Переваги	Недоліки
Rule-Based	Простота, прозорість	Високий рівень false positive
Машинне навчання	Адаптивність, висока точність	Потреба в даних та навчанні

Гібридні моделі	Компроміс між точністю і контролем	Ускладнена реалізація
Блокчейн	Незмінність, аудит	Обмеження в масштабованості
Графові алгоритми	Виявлення складних схем	Високі вимоги до обчислювальних ресурсів

Для оцінки ефективності методів виявлення фінансових злочинів зазвичай використовують основні чотири метрики: Precision, Recall, F1-score, **False Positive Rate.**

Precision (точність) – частка коректно виявлених шахрайств серед усіх виявлених, визначається відношенням:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (1)$$

де TP – кількість істинно позитивних (правильно виявлених шахрайств), FP – кількість хибнопозитивних (не шахрайство, але виявлено як шахрайство).

Recall (повнота) – частка виявлених шахрайств серед усіх реальних випадків, визначається:

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (2)$$

де TP – кількість істинно позитивних (правильно виявлених шахрайств), FN – кількість хибнонегативних (реальне шахрайство, яке не було виявлене).

F1-score – збалансована метрика між точністю і повнотою:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (3)$$

де Precision – частка коректно виявлених шахрайств серед усіх виявлених, Recall – частка виявлених шахрайств серед усіх реальних випадків.

False Positive Rate (хибні спрацювання) – критичний показник для банківських систем [8], що визначається:

$$FPR = \frac{FP}{FP + TN}, \quad (4)$$

де FP – кількість хибнопозитивних (не шахрайство, але виявлено як шахрайство), TN – кількість істинно негативних (правильно класифікованих як не шахрайство).

Для об'єктивного порівняння ефективності різних підходів до виявлення фінансових злочинів були використані уніфіковані показники. Зокрема, дані щодо rule-based систем узяті з дослідження [1], яке розглядає комбіновані системи правил та машинного навчання. Проведено детальний аналіз показників для XGBoost [16], метрик для глибоких нейронних мереж [9], гібридних підходів [15] та графових моделей [4]. Проаналізовані джерела репрезентують як міжнародні, так і національні підходи, що дозволяє сформувати узагальнену та достовірну картину результативності досліджуваних методів (табл. 2).

Таблиця 2

Порівняльна таблиця ефективності (уніфіковані дані)

Метод	Precision	Recall	F1-score	False Positives
Rule-Based System	0.65	0.50	0.56	Високий
Машинне навчання (XGBoost)	0.89	0.87	0.88	Низький
Глибока нейронна мережа	0.91	0.85	0.88	Середній
Гібридна система	0.87	0.90	0.88	Низький
Графовий аналіз	0.80	0.92	0.85	Середній

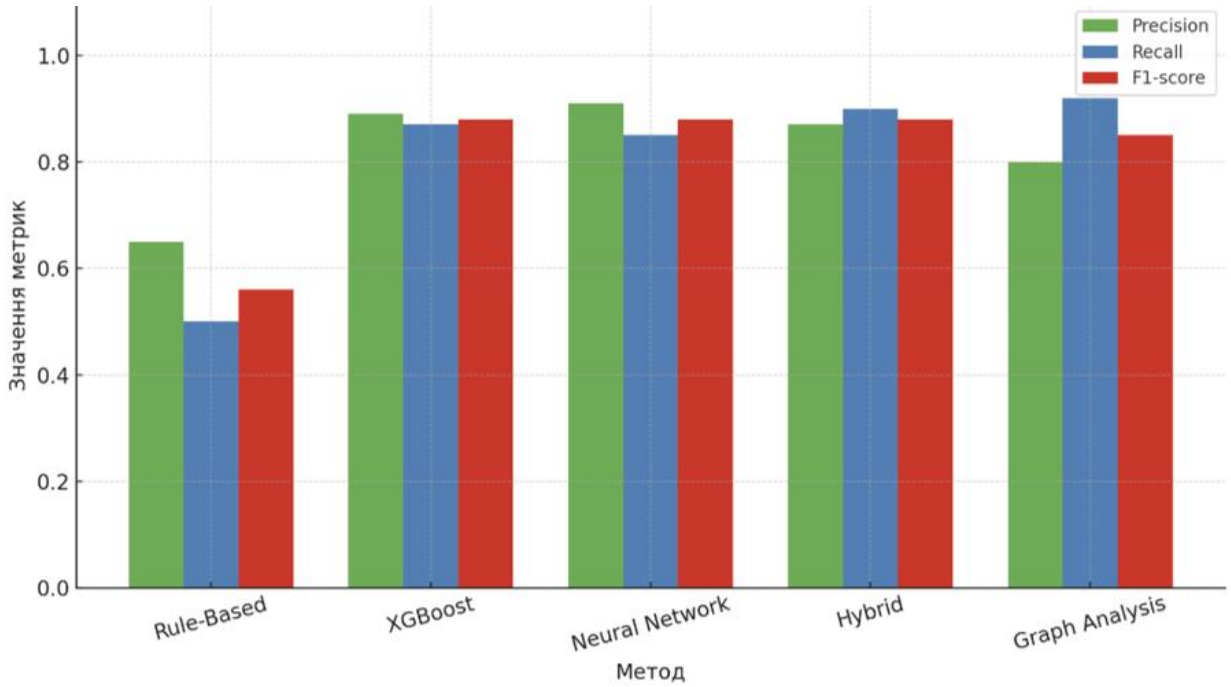

Рис. 7. Порівняння характеристик ефективності методів за метриками

Проблеми та обмеження використання агентів штучного інтелекту у реальному середовищі

У ході проведеного дослідження було виявлено низку ключових проблем і обмежень, які суттєво впливають на ефективність впровадження агентів штучного інтелекту у практичні системи виявлення фінансових злочинів. Ці обмеження мають як технічний, так і регуляторно-правовий характер і потребують детального аналізу у контексті реального фінансового середовища.

Однією з найсуттєвіших проблем, виявлених під час аналізу впровадження глибоких нейронних мереж у банківські системи, є обмежена прозорість їхніх рішень. Такі моделі функціонують за принципом «чорної скриньки» і не дозволяють пояснити, чому саме транзакція була класифікована як шахрайська. Ця непрозорість створює серйозні перешкоди для аудиту, юридичної відповідальності та прийняття обґрунтованих управлінських рішень у регульованих сферах, таких як банківська справа та страхування [2, 5, 8].

Ще однією суттєвою проблемою, виявленою у процесі дослідження, є високе навантаження на обчислювальні ресурси при використанні графових алгоритмів і deep learning моделей. Реалізація таких систем потребує потужної апаратної інфраструктури (GPU, TPU) і є затратною в експлуатації, особливо у випадках обробки великих потоків транзакцій у режимі реального часу [4, 5].

Крім того, було виявлено обмеження, пов’язані з доступністю якісних і збалансованих даних. У реальних фінансових системах переважає дисбаланс класів – на тисячі «нормальних» транзакцій припадає лише кілька шахрайських. Це ускладнює навчання моделей машинного навчання і потребує використання спеціалізованих методів, таких як SMOTE або ADASYN, що також впливає на точність і стабільність моделей [6, 8].

Під час аналізу практичного застосування агентних систем також було виявлено низку юридичних обмежень. Зокрема, не всі рішення, що приймаються моделями ШІ, відповідають вимогам міжнародного та національного законодавства – таких як GDPR (General Data Protection Regulation) та директивам щодо боротьби з відмиванням коштів (AMLD). У випадку українського законодавства спостерігається потреба у кращій регламентації автоматизованого фінансового моніторингу з використанням ШІ [6, 7].

Отже, результати дослідження підтверджують, що для ефективного впровадження агентів штучного інтелекту у фінансові системи необхідно подолати низку бар’єрів, пов’язаних із поя-

снюваністю моделей, ресурсними обмеженнями, якістю даних, правовими викликами та кадровим забезпеченням. Подальші дослідження мають бути спрямовані на розробку інтерпретованих і масштабованих систем ШІ, які б відповідали нормативним вимогам і забезпечували високий рівень фінансової безпеки.

Проаналізувавши висвітлені підходи до виявлення аномалій у банківських транзакціях, проблеми та обмеження, можна зробити порівняльний опис традиційних і ШІ підходів за відповідними характеристиками (табл. 3).

Таблиця 3

Переваги та недоліки методів ШІ у порівнянні з традиційними системами

Характеристика	Rule-Based System	AI/ML-підходи
Гнучкість	Низька	Висока
Потреба у даних	Мінімальна	Висока
Репрезентативність	Висока (зрозумілі правила)	Низька (особливо у глибоких моделей)
Можливість адаптації	Обмежена	Динамічна адаптація до нових загроз
Вартість впровадження	Низька	Висока (обчислювальні ресурси, дані)
Схильність до помилок	Часто виявляє «помилкові позитиви»	Може уникати їх за рахунок навчання

Перспективи розвитку та вдосконалення методів штучного інтелекту у сфері фінансової безпеки

У ході проведеного дослідження було виявлено, що подальший розвиток систем виявлення фінансових злочинів на основі агентів штучного інтелекту нерозривно пов'язаний із вдосконаленням трьох ключових напрямів: підвищенням прозорості рішень, архітектурною еволюцією мультиагентних систем та впровадженням новітніх технологій обчислення, таких як квантові обчислення та edge computing.

Першочерговим викликом, що потребує вирішення, є проблема інтерпретованості моделей, яка має вирішальне значення для банківського та фінансового сектора, що функціонує в умовах жорсткого регулювання. Застосування концепції Explainable AI (XAI), як було з'ясовано у дослідженні, дозволяє підвищити прозорість прийняття рішень складними системами ШІ, зокрема глибокими нейронними мережами. Алгоритми XAI, такі як SHAP та LIME, забезпечують можливість оцінки впливу окремих ознак на результат моделі, що дозволяє підвищити довіру як користувачів, так і регуляторів до автоматизованих систем фінансового моніторингу [2, 5].

Другим важливим вектором вдосконалення є розвиток мультиагентних архітектур, які дають змогу ефективно реалізувати розподілені функції аналізу у межах фінансових систем. У процесі дослідження встановлено, що кооперація між агентами – зокрема розподіл завдань між детекцією транзакцій, перевіркою клієнтських профілів та аналізом соціальних зв'язків – створює передумови для більш глибокого аналізу складних схем шахрайства. Такий підхід забезпечує адаптивність і масштабованість систем без значної втрати ефективності в умовах зростаючих обсягів транзакцій [3, 11].

Окрему увагу в межах дослідження приділено перспективам інтеграції моделей ШІ із новітніми обчислювальними підходами. Зокрема, зазначено потенціал квантових обчислень у прискоренні аналізу великих фінансових графів, що є типовими для шахрайських схем із багатьма учасниками. Хоча практичне впровадження таких рішень ще перебуває на початковій стадії, теоретичні й експериментальні результати вказують на значне зростання обчислювальної ефективності при використанні квантових алгоритмів [4, 14]. Також дослідженням підтверджено доцільність застосування edge computing – виконання локального аналізу на пристроях користувачів або безпосередньо у точках транзакцій – для зниження затримок та підвищення швидкодії систем виявлення [8].

У цілому результати дослідження демонструють, що розвиток інтелектуальних агентних технологій у сфері фінансової безпеки має спиратися на комплексну інтеграцію технічної інноваційності з вимогами регуляторної відповідності та інтерпретованості. Це відкриває перспективи не лише для зростання точності виявлення фінансових порушень, а й для підвищення стійкості фінансових систем до нових форм загроз.

Висновки

У ході дослідження було здійснено систематичний аналіз сучасних методів застосування агентів штучного інтелекту у системах забезпечення фінансової безпеки, зокрема у виявленні та протидії фінансовим злочинам. У центрі розгляду перебували методи машинного навчання, мультиагентні системи, графові нейронні мережі та підходи на основі Explainable AI. Було досліджено архітектурні особливості агентних систем, їхню адаптивність, здатність до кооперації та самоорганізації у складному інформаційному середовищі, а також роль децентралізації у підвищенні стійкості до шахрайських впливів.

Проведений аналіз дозволив окреслити основну сферу застосування цих методів – системи моніторингу та виявлення аномалій у фінансових транзакціях, автоматизовані системи банківського фінансового контролю, цифровий комплаєнс, а також розпізнавання складних схем фінансового шахрайства (наприклад, обхід регуляторних обмежень через фіктивні юридичні особи або транзакційні ланцюги). Серед ключових проблем використання агентів ШІ було ідентифіковано обмежену інтерпретованість рішень, залежність від великих обсягів високоякісних даних, високе навантаження на обчислювальні ресурси та правові обмеження щодо автоматизованої обробки персональних даних відповідно до положень GDPR і AMLD.

Серед переваг сучасних методів було виокремлено їхню здатність до самонавчання, виявлення складних прихованих патернів, динамічної адаптації до нових типів загроз та координації між різнорівневими аналітичними модулями. Разом із цим недоліки – зокрема, «чорний ящик» глибоких моделей, складність інтеграції в існуючі фінансові IT-системи, а також недостатній рівень нормативного врегулювання – залишаються суттєвими бар'єрами до повноцінного впровадження.

У перспективі найбільш багатообіцяючими напрямами вдосконалення виявилися розвиток інтерпретованих моделей XAI, архітектури розподілених мультиагентних систем із розширеними можливостями комунікації, використання квантових обчислень для обробки великих транзакційних графів, а також впровадження edge computing для забезпечення миттєвої обробки даних на периферійних пристроях. Отже, результати дослідження підтверджують доцільність комплексного розвитку агентів штучного інтелекту у сфері фінансової безпеки з орієнтацією на інтерпретованість, адаптивність та регуляторну відповідність. Це дозволить підвищити ефективність виявлення фінансових правопорушень і сприятиме загальному зміцненню інституційної та кіберекономічної стійкості фінансових систем.

Список літератури

1. Robert Abill, Abiodun Okunola, Michack Abigail Enhancing. Fraud Detection through Hybrid AI Models: Combining Rule-Based Systems with Machine Learning. ResearchGate, April, 2025. URL:https://www.researchgate.net/publication/390630333_Enhancing_Fraud_Detection_through_Hybrid_AI_Models_Combining_Rule-Based_Systems_with_Machine_Learning

2. Alejandro Barredo Arietta et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. Volume 58, June, 2020. Pages 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

3. Michael Wooldridge. An Introduction to MultiAgent Systems, 2nd Edition. ISBN: 978-E-UDT-E0055-3. July, 2011. 488 pages. URL: https://www.wiley.com/en-be/An+Introduction+to+MultiAgent+Systems%2C+2nd+Edition-p-978EUDTE00553

4. Yingtong Dou et al. Enhancing Graph Neural Network-based Fraud Detectors against Camouflaged Fraudsters. CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management. October, 2020. Pages 315–324. https://doi.org/10.1145/3340531.3411903

5. Johannes Jurgovsky et al. Sequence classification for credit-card fraud detection. Expert Systems With Applications. Volume 100, June, 2018, Pages 234–245. https://doi.org/10.1016/j.eswa.2018.01.037

6. Ibrahim Y. Hafez et al. A systematic review of AI-enhanced techniques in credit card fraud detection. Journal of Big Data. Volume 12, January, 2025.URL: https://link.springer.com/article/10.1186/s40537-024-01048-8

7. Овчаренко, Т. (2024). Тенденції розвитку та використання штучного інтелекту у банківській сфері. Економіка та суспільство, (67). https://doi.org/10.32782/2524-0072/2024-67-44

8. Eryu Pan. Machine Learning in Financial Transaction Fraud Detection and Prevention. Vol. 5 (2024): 6th International Conference on Business, Economics, Management Science (BEMS 2024). March, 2024. https://doi.org/10.62051/16r3aa10

9. Domor I. Mienye, Nobert Jere. Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions. IEEE Access. Volume 12, January, 2024. http://dx.doi.org/10.1109/ACCESS.2024.3426955

10. Masoumeh Zareapoor, Pourya Shamsolmoali. Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. Procedia Computer Science. Volume 48, 2015, Pages 679–685. https://doi.org/10.1016/j.procs.2015.04.201

11. Claudio Reginaldo Alexandre, João Balsa. Incorporating machine learning and a risk-based strategy in an anti-money laundering multiagent system. Expert Systems With Applications. Volume 217, May, 2023, 119500. https://doi.org/10.1016/j.eswa.2023.119500

12. Liudmila Prokhorenkova et al. CatBoost: unbiased boosting with categorical features. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada. URL: https://papers.nips.cc/paper_files/paper/2018/file/14491b756b3a51daac41c24863285549-Paper.pdf

13. Ismini Psychoula et al. Explainable Machine Learning for Fraud Detection. IEEE Computer Special Issue on Explainable AI and Machine Learning. May, 2021. https://doi.org/10.48550/arXiv.2105.06314

14. Debasis Mohanty et al. Implementing Blockchain Technology for Fraud Detection in Financial Management. April, 2023. URL:

https://www.researchgate.net/publication/370072691_Implementing_Blockchain_Technology_for_Fraud_Detection_in_Financial_Management

15. Diego Vallarino. Detecting Financial Fraud with Hybrid Deep Learning: A Mix-of-Experts Approach to Sequential and Anomalous Patterns. April, 2025. https://doi.org/10.48550/arXiv.2504.03750

16. Ayed Alwadain et al. Estimating Financial Fraud through Transaction-Level Features and Machine Learning. Mathematics, MDPI, February, 2023. Volume 11(5), Article 1184. URL: https://www.mdpi.com/2227-7390/11/5/1184

B. Kalyniuk, I. Zamrii, A. Kalyniuk

OVERVIEW OF MODERN METHODS FOR DETECTING FINANCIAL CRIMES USING ARTIFICIAL INTELLIGENCE AGENTS

The article presents a comprehensive analysis of modern methods for detecting financial crimes using artificial intelligence (AI) agents. It examines the classification of AI agents (reactive, deliberative, autonomous, and multi-agent), their operational features, and their role in financial monitoring systems. A comparative analysis of rule-based approaches, machine learning methods, hybrid models, blockchain architectures, and graph algorithms has been conducted. The study reveals that hybrid solutions and graph neural networks demonstrate the highest levels of precision and recall in detecting suspicious transactions, as confirmed by consolidated metrics from peer-reviewed sources. Special attention is given to the use of deep neural networks, time series processing techniques, natural language processing (NLP), and Explainable AI (XAI) methods, which enhance the transparency and interpretability of AI-driven decisions—an essential requirement in heavily regulated financial domains. The advantages of the multi-agent approach are emphasized, including the ability for parallel analysis of complex fraud schemes, dynamic adaptability, and system scalability, which position

it as a promising direction for the development of distributed financial intelligence systems. Alongside the identified benefits, the study also outlines key challenges hindering real-world implementation: limited interpretability of deep learning models, the requirement for large volumes of high-quality and balanced data, high computational costs, and legal and ethical constraints related to the automated processing of sensitive financial information, especially under regulations such as GDPR and AMLD.

In terms of future development, the article highlights the potential for integrating AI agents with blockchain networks to ensure transactional transparency and immutability, applying quantum algorithms for processing complex financial graphs, and adopting edge computing for real-time anomaly detection on decentralized devices. Furthermore, the research underlines the importance of an interdisciplinary approach that combines expertise in artificial intelligence, cybersecurity, economics, and legal compliance in building robust and effective financial crime prevention systems. Thus, the findings confirm that effective financial crime detection today requires complex technological solutions based on AI agents with an emphasis on transparency, scalability, and regulatory compliance. The conclusions presented in this study hold both theoretical and practical value for the development of modern financial transaction monitoring systems and the formulation of policies in the field of digital financial security.

Keywords: financial crimes; artificial intelligence; artificial intelligence agents; machine learning; blockchain; graph analysis.