МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ ЛЬВІВСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ ВНУТРІШНІХ СПРАВ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ УПРАВЛІННЯ, ПСИХОЛОГІЇ ТА БЕЗПЕКИ
Кафедра інформаційних технологій

РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АНАЛІЗУ КРИПТОВАЛЮТНИХ ТРАНЗАКЦІЙ ДЛЯ ВИЯВЛЕННЯ ФІНАНСОВИХ ЗЛОЧИНІВ

кваліфікаційна робота
здобувача вищої освіти
4 курсу денної форми навчання
Вікторії ГОЗИ

Науковий керівник:
Доктор філософії
Олег БАСИСТЮК
Рецензент:

_____
вчене звання, науковий ступінь

_____
(Ім’я ПРІЗВИЩЕ рецензента)

Кваліфікаційна робота допущена до захисту
«___» _____ 2026 р., протокол № ____
Завідувач кафедри інформаційних технологій
_____Олег ЗАЧЕК
(підпис)

Львів 2026

АНОТАЦІЯ

Гоза В. Розроблення інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів. — рукопис

Дослідження на здобуття освітнього ступеня «бакалавр» за спеціальністю 126 «Інформаційні системи та технології». – Львівський державний університет внутрішніх справ, МВС України, Львів, 2026.

У роботі досліджено аналіз криптовалютних транзакцій та розроблено інтелектуальну систему для виявлення підозрілих фінансових операцій. Проаналізовано методи аналізу транзакцій і системи blockchain-аналітики. Практична частина включає реалізацію програмного прототипу мовою Python із використанням методів машинного навчання та графового аналізу.

Ключові слова: криптовалюта, blockchain, транзакції, фінансові злочини, машинне навчання, Python.

ABSTRACT

Goza V. Development of an Intelligent System for Cryptocurrency Transaction Analysis for Financial Crime Detection. — Manuscript.

Bachelor’s qualification research paper in Specialty 126 “Information Systems and Technologies”. – Lviv State University of Internal Affairs, Ministry of Internal Affairs of Ukraine, Lviv, 2026.

The paper investigates cryptocurrency transaction analysis and presents the development of an intelligent system for detecting suspicious financial operations. Methods of transaction analysis and blockchain analytics systems are analyzed. The practical part includes the implementation of a software prototype using the Python programming language with machine learning and graph analysis methods.

Keywords: cryptocurrency, blockchain, transactions, financial crimes, machine learning, Python.

ЗМІСТ

Вступ

Сучасний розвиток цифрових технологій та blockchain-систем сприяв стрімкому поширенню криптовалют у фінансовому середовищі. Криптовалюти активно використовуються для міжнародних переказів, електронних платежів, інвестиційної діяльності та децентралізованих фінансових сервісів. Разом із перевагами анонімності, швидкості та децентралізації виникають і суттєві ризики, пов’язані з використанням криптовалют у незаконній діяльності. Значна кількість фінансових злочинів, зокрема відмивання коштів, шахрайство, фінансування незаконної діяльності та приховування джерел походження активів, здійснюється саме через blockchain-мережі. У зв’язку з цим виникає необхідність створення інтелектуальних систем, здатних автоматично аналізувати великі обсяги криптовалютних транзакцій та виявляти підозрілу фінансову активність.

Проблематика аналізу криптовалютних транзакцій активно досліджується у сучасній науковій та практичній сфері. Значна увага приділяється використанню методів машинного навчання, графового аналізу та аналізу поведінкових характеристик користувачів blockchain-мереж. Сучасні аналітичні платформи, зокрема Chainalysis та Elliptic, демонструють ефективність автоматизованого моніторингу криптовалютних операцій для виявлення фінансових злочинів.

Метою дипломної роботи є розроблення інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів із використанням методів машинного навчання та графового аналізу.

Для досягнення поставленої мети у роботі необхідно виконати такі завдання:

- дослідити поняття криптовалют і blockchain-технологій;
- проаналізувати основні типи фінансових злочинів у криптовалютному середовищі;

- дослідити сучасні методи аналізу криптовалютних транзакцій;
- виконати огляд існуючих систем blockchain-аналітики;
- провести аналіз наборів даних криптовалютних транзакцій;
- визначити основні ознаки підозрілих фінансових операцій;
- обґрунтувати вибір алгоритмів машинного навчання;
- розробити архітектуру інтелектуальної системи;
- реалізувати програмний прототип системи;
- провести тестування та оцінювання точності роботи системи.

Об’єктом дослідження є процес аналізу криптовалютних транзакцій у blockchain-мережах.

Предметом дослідження є методи та програмні засоби виявлення фінансових злочинів на основі аналізу криптовалютних транзакцій із використанням алгоритмів машинного навчання.

У процесі виконання дипломної роботи використовувалися методи системного аналізу, методи обробки даних, алгоритми машинного навчання, методи класифікації та виявлення аномалій, графовий аналіз, статистичні методи аналізу даних, а також засоби програмної інженерії для розроблення програмного забезпечення.

Практичне значення отриманих результатів полягає у створенні програмного прототипу системи аналізу криптовалютних транзакцій, який може використовуватись для автоматизованого виявлення підозрілої фінансової активності у blockchain-мережах. Розроблена система може бути основою для подальшого розвитку платформ фінансового моніторингу, інтеграції з blockchain API та реалізації систем аналізу транзакцій у режимі реального часу. Результати дослідження можуть бути використані у навчальному процесі, наукових дослідженнях та практичній діяльності у сфері кібербезпеки й фінансового моніторингу.

РОЗДІЛ 1. ТЕОРЕТИЧНИЙ РОЗДІЛ

Розділ 1.1. поняття криптовалют і блокчейну

Сучасне цифрове середовище кипротовалют стало однією з найважливіших складових фінансової системи, та електронної комерції. Поява криптовалют була зумовлена розвитком інформаційних технологій, децентралізованих мереж та потребою у швидких, незалежних, безпечних фінансових операціях. В основу функціонування криптовалют вкладена технологія блокчейну, що забезпечує конфіденційну передачу та зберігання даних без участі центрального органу регулювання.

Криптовалюта – це віртуальний або цифровий актив, який використовується як метод обміну, та захищений криптографічними методами. На відміну від традиційних валют, криптовалюти не контролюються банками чи державними органами регулювання. Усі операції, без виключення, здійснюються в децинтралізованій мережі. Комп'ютерів, що забезпечує прозорість та стійкість системи до зовнішніх втручань.

Перша та одна з найвідоміших криптовалют є Bitcoin, була створена 2009 році, до тепер достеменно невідомо хто саме є засновником цього активу, відомо тільки псевдонім особи або групи осіб під псевдонімом Satoshi Nakamoto. За останніми дослідженнями журналісти The New York Times [13] стверджують, що під псевдонімом Satoshi Nakamoto приховується 55-річний британський криптограф, винахідник Адам Бек. Автори виявили значну кількість збігів у стилі письма, використанні рідкісних термінів, ідеологічних поглядах та технічних ідеях між Адамом Беком та Satoshi Nakamoto. [9]

Основною метою створення Bitcoin було забезпечення можливості проведення фінансових операцій без посередників. Після появи Bitcoin почали активно розвиватись й інші криптовалюти, зокрема такі відомі криптовалюти як: Ethereum, Litecoin, Riple, тощо.

Технологічною основою більшості криптовалют як згадувалось вище є блокчейн. Блокчейн – це розподілений реєстр даних, у якому інформація про транзакції зберігається у вигляді послідовності конкретних блоків. Кожний окремий блок містить певний набір транзакцій, часову позначку транзакцій та криптографічний хеш попереднього блоку, яка забезпечує цілісність, безпеку і незмінність даних. [4]

Принцип роботи блокчейну полягає у тому, що всі учасники мережі мають копію реєстру транзакцій. Після здійснення нової транзакції вона перевіряється учасниками транзакції і додається до нового блоку. Після підтвердження блок поєднується з загальним ланцюгом блоків. Завдяки Використанню криптографічних алгоритмів змінити або підробити інформацію в уже створених блоках являється практично неможливою задачею.

Однією з головних переваг блокчейну є децентралізована система. Відсутність єдиного центру управління підвищує стійкість системи до кібератак і технічних збоїв в системі. Крім того, блокчейн забезпечує прозорість операцій, оскільки всі транзакції можуть бути перевіреними учасниками мережі. В той сам час користувачі залишаються відносно анонімними, що являє собою як перевагу так і потенційну загрозу з точки фінансової безпеки.

Завдяки своїм характеристиками, криптовалюти та блокчейн-технології активно використовуються в різних сферах, зокрема такі як: фінансові розрахунки, міжнародні перекази, інвестиційна діяльність, цифрові контракти, тощо. Разом із цим розвиток криптовалютного ринку сприяв виникненню нових видів фінансових злочинів, таких як відмивання коштів, шахрайство, фінансування незаконної діяльності та приховування походження активів. Це стає передумовою необхідності створення інтелектуальних систем аналізу криптовалютних транзакцій для своєчасного виявлення підозрілих операцій та забезпечення фінансової безпеки.

Розділ 1.2. Типи фінансових слочинів у криптосередовищі

Стрімкий розвиток криптовалют та блокчейн-технологій сприяв не лише нові можливості для цифрової економіки, але й сприяв виникненню нових форм фінансових злочинів. Високий рівень анонімності, децентралізований характер систем та складність відстеження транзакцій роблять криптовалютне середовище привабливим для незаконної діяльності. У зв’язку з цим проблема виявлення фінансових злочинів у сфері криптовалют є однією з найактуальніших у галузі кібербезпеки та фінансового моніторингу.

Одним із найпоширеніших видів злочинів у криптосередивищі є відмивання коштів. Зловмисники використовують криптовалюти для прорахування джерел незаконно отриманих фінансів та їх подальшої легалізації. Для цього часто застосовуються численні перекази між різними гаманцями, використання використання криптовалютних міксерів та обмін активів між різними блокчейн-мережами. Через відсутність прямої прив’язки криптогаманців до особливих даних користувачів процес ідентифікації власника гаманця значно ускладнюється. [1]

Іншим поширеним видом фінансових злочинів є шахрайство з інвестиціями. У криптовалютній сфері часто створюються фіктивні інвестиційні платформи, псевдобіржі та фінансові піраміди які обіцяють користувачам високий прибуток за короткий період часу. Після залучення значної кількості коштів, організатори припиняють діяльність та зникають разом із активами інвесторів. Подібні схеми часто реалізуються через соціальні мережі, месенджери та фальшиві веб-ресурси. [2]

Значну загрозу становлять фішингові атаки та викрадення криптовалютних активів. Зловмисники створюють підроблені сайти криптобірж, електронних гаманців або надсилають користувачам фальшиві повідомлення з метою отримання логінів, паролів чи приватних ключів, при тому зловмисники роблять картинку ідентичну до реальних бірж, щоб у користувачів які не надто добре

орієнтуються на платформі не виникало ніяких сумнівів, для легшого викрадення даних. Після отримання доступу до гаманця кошти швидко переводяться на інші адреси, що вкрай ускладнює їх повернення.

Окремою категорією фінансових злочинів є використання криптовалют для фінансування незаконної діяльності. Криптовалюти можуть застосовуватись для здійснення платежів у тіньовому сегменті мережі Інтернет, зокрема для купівлі заборонених товарів, фінансування кіберзлочинності або обходу міжнародних санкцій. Анонімність транзакцій створює сприятливі умови для приховування реальних учасників фінансових операцій. [3]

Також поширеним явищем є діяльність програм-вимагачів. У результаті зараження комп’ютерної системи шкідливим програмним забезпеченням дані користувача блокуються або шифруються, після чого зловмисники вимагають викуп у криптовалюті за відновлення доступу. Найчастіше для таких платежів використовуються Bitcoin або інші криптовалюти з підвищеним рівнем анонімності.

Ще одним видом злочинної діяльності є маніпулювання криптовалютним ринком. До таких дій належить штучне підвищення або зниження вартості активів, створення фіктивного торгового обсягу, поширення неправдивої інформації та організація схем типу «Pump and Dump». Метою таких маніпуляцій є отримання прибутку шляхом введення інвесторів в оману.

Особливу складність для правоохоронних органів становлять транзакції, пов’язані з використанням сервісів анонімізації. До них належать криптовалютні міксери, технології CoinJoin та анонімні криптовалюти, такі як Monero або Zcash. Такі інструменти ускладнюють відстеження руху коштів та безпосередній аналіз фінансових потоків.

Для боротьби з фінансовими злочинами у криптосередовищі активно застосовуються сучасні методи аналізу даних, машинного навчання та штучного інтелекту. Інтелектуальні системи аналізу транзакцій дозволяють автоматично

виявляти аномальні операції, встановлювати зв’язки між адресами гаманців та прогнозувати потенційно підозрілу активність. Це значно підвищує ефективність фінансового моніторингу та сприяє протидії незаконній діяльності у сфері криптовалют.

Розділ 1.3. Методи аналізу транзакцій

Зростання кількості фінансових операцій у криптовалюиних мережах створює потребу у використанні ефективних методів аналізу транзакцій. Основною метою такого аналізу транзакцій. Основною метою такого аналізу є виявлення підозрілої активності, встановлення зв’язків між учасниками мережі, визначення джерел проходження коштів та запобігання фінансовим злочинам. Через децентралізований характер блокчейн-систем традиційні методи фінансовго моніторингу не завжди є достатньо ефективними, тому активно застосовуються сучасні підходи аналізу даних та штучного інтелекту. [7]

Важливим підходом є графовий аналіз транзакцій. У цьому випадку блокчейн розглядається як граф, де вузлами є криптовалютні адреси, а ребрами – транзакції між ними. Графові методи дозволяють досліджувати структуру взаємозв’язків у мережі, визначати центральні вузли, виявляти групи пов’язаних адрес та знаходити приховані закономірності. Такий підхід особливо ефективний для виявлення схем відмивання коштів та організованих фінансових операцій.[8]

Для автоматичного визначення підозрілих транзакцій широко використовуються методи машинного навчання. Алгоритми аналізують великі обсяги даних та виявляють характерні ознаки шахрайської активності. Для цього формуються набори параметрів транзакцій, серед яких можуть бути сума переказу, загальна кількість операцій час здійснення транзакцій, тип криптовалюти та поведінкові характеристики користувачів. [5]

Серед алгоритмів класифікації часто застосовують дерева рішень, Random Forest, метод опорних векторів та нейронні мережі. Ці моделі дозволяють класифікувати транзакції як легітимні або потенційно небезпечні на основі попередньо навченої вибірки даних. Ефективність таких методів залежить від якості навчальних даних та правильного вибору ознак методів.

Окрему роль відіграють методи виявлення аномалій. Їх основна мета полягає у знаходженні транзакцій, які суттєво відрізняються від типової поведінки користувачів. Для цього використовуються алгоритми кластеризації, статистичного аналізу та спеціалізовані моделі, такі як Isolation Forest або Local Outlier Facor. Аномальні транзакції можуть свідчити про спроби відмивання коштів, несанкціонованого доступу до гаманця або інші види фінансових злочинів.[5]

Також активно застосовуються поведінкові методи аналізу. Вони базуються на дослідженні характеру дій користувачів у блокчейн-мережі. Наприклад, система може аналізувати регулярність переказів, типові суми операцій, часову активність та взаємодію з іншими адресами криптогаманців. Різка зміна поведінки користувача може бути ознакою компрометації акаунта або незаконної діяльності.

Для підвищення точності аналізу сучасні системи часто використовують комбіновані підходи, які поєднують графовий аналіз, машинне навчання та статистичні методи. Це дозволяє враховувати як структурні зв’язки між транзакціями, так і поведінкові характеристики тих чи інших користувачів. [11f]

Окрім програмних методів аналізу, використовуються спеціалізовані аналітичні платформи, які здійснюють моніторинг блокчейн-мереж у режимі реального часу. Такі системи дозволяють автоматично ідентифікувати підозрілі адреси, формувати ризикові профілі та генерувати повідомлення про потенційну фінансову загрозу користувача.[6]

Саме таким чином, методи аналізу криптовалютних транзакцій є дуже важливим інструментом забезпечення фінансової безпеки користувачів у цифровому середовищі. Використання сучасних алгоритмів аналізу даних та штучного інтелекту дозволяє значно підвищити ефективність виявлення фінансових злочинів та мінімізувати ризики незаконного використанння криптовалют.

Розділ 1.4. Огляд інснуючих систем

Зі зростанням популярності криптовалют та збільшенням кількості фінансових злочинів у блокчейн-середовищі виникла потреба у створенні спеціалізованих систем аналізу криптовалютних транзакцій. Такі системи використовуються правоохоронними органами, фінансовими установами, криптовалютними біржами та аналітичними компаніями для моніторингу блокчейн-мереж, виявлення підозрілих операцій та протидії незаконній діяльності.

Сучасні платформи аналізу криптовалютних транзакцій поєднують методи обробки великих даних, графового аналізу, машинного навчання та штучного інтелекту. Їх основними функціями є відстеження руху коштів, ідентифікація ризикових адрес, аналіз поведінки користувачів та автоматичне виявлення фінансових загроз.

Однією з найбільш відомих систем є Chainalysis. Дана платформа спеціалізується на аналізі блокчейн-транзакцій та використовується урядовими організаціями, банками та криптовалютними біржами. Основною перевагою Chainalysis є можливість відстеження руху криптовалют між різними адресами та визначення зв’язків між учасниками мережі.[18]

Система використовує технології кластеризації адрес, графового аналізу та машинного навчання для виявлення підозрілих фінансових операцій.

Платформа дозволяє формувати профілі криптовалютних гаманців, оцінювати рівень ризику транзакцій та автоматично виявляти адреси, пов’язані з шахрайством, відмиванням коштів або кіберзлочинністю. Крім того, Chainalysis підтримує моніторинг транзакцій у режимі реального часу, що підвищує ефективність фінансового контролю .

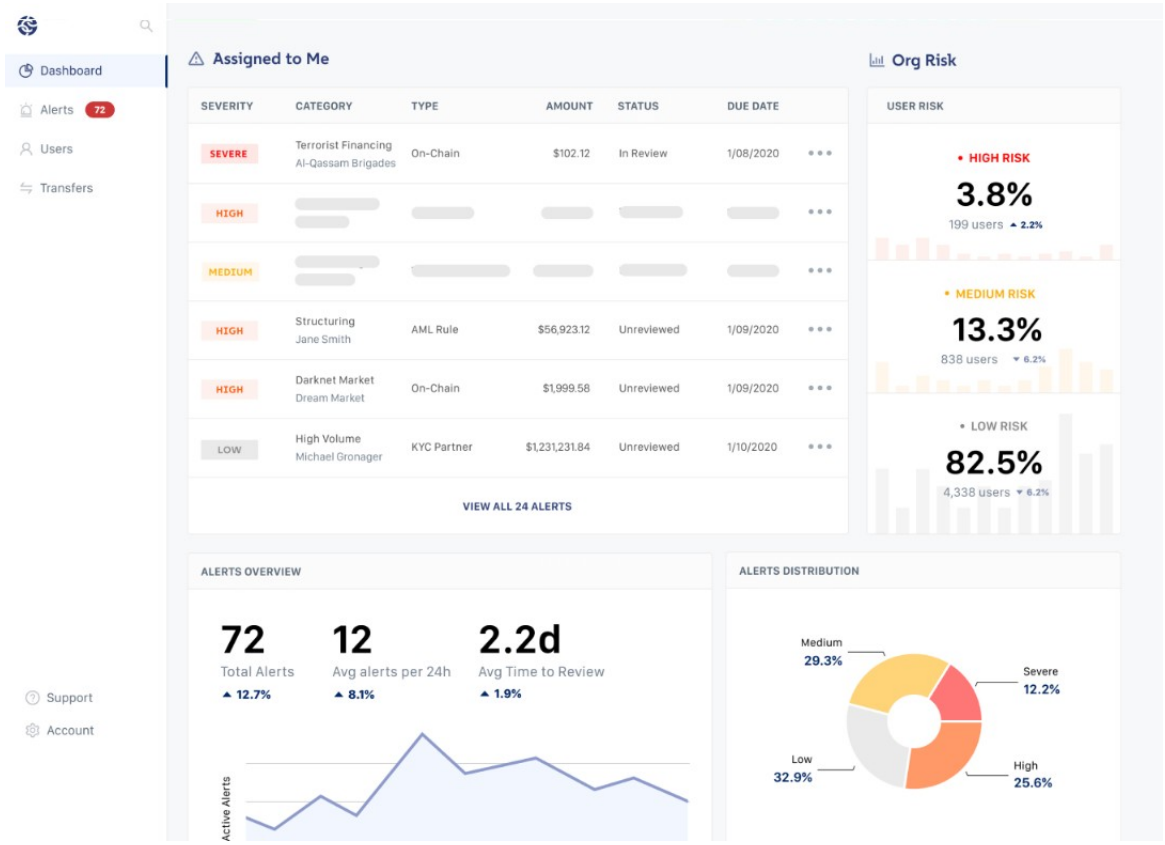


Рис 1.1 Платформа Chainalysis [18]

Іншою популярною системою є Elliptic. Основною функцією Elliptic є аналіз криптовалютних ризиків та забезпечення відповідності міжнародним вимогам фінансового моніторингу. Платформа активно використовується для реалізації процедур KYC (Know Your Customer) та AML (Anti-Money Laundering).

Elliptic застосовує методи штучного інтелекту та аналітики блокчейн-мереж для визначення джерел походження коштів і виявлення незаконної діяльності. Система здатна автоматично аналізувати великі обсяги транзакцій та формувати ризикові оцінки для окремих адрес і фінансових

операцій. Особливістю платформи є підтримка великої кількості криповалютних мереж та інтеграція з біржами та фінансовими сервісами.[19]

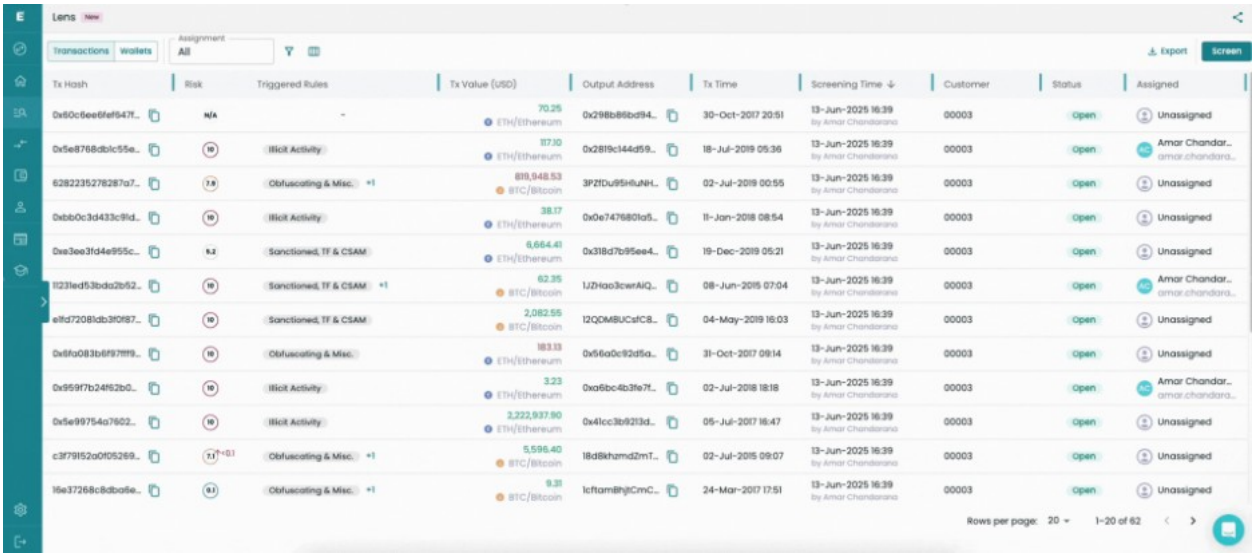


Рис 2.2 Платформа Elliptic [19]

Ще однією відомою системою є ClipherTrace. Вона орієнтована на забезпечення безпеки криптовалютних операцій та виявлення шахрайських схем. Платформа здійснює моніторинг блокчейн-мереж, аналізує ризики транзакцій та допомагає організаціям дотримуватись міжнародних стандартів фінансової безпеки.Для аналізу транзакцій ClipherTrace використовує механізм штучного інтелекту, поведінковий аналіз та алгоритм виявлення аномалій. Система дозволяє ідентифікувати адреси, пов’язані з незаконною діяльністю, а також відстежувати переміщення криптовалют між різними платформами та сервісами. [20]

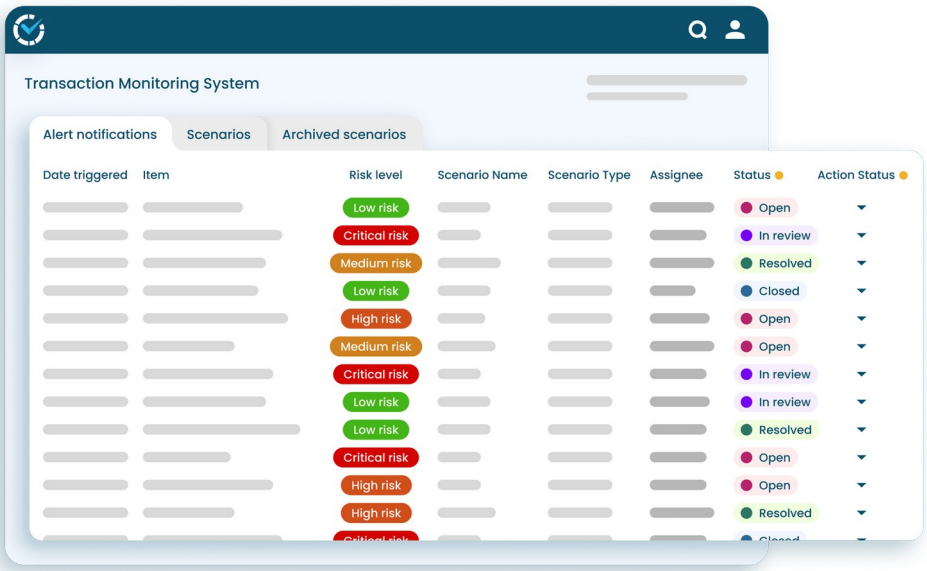


Рис 3.3 Платформа ClipherTrace [20]

Окремої уваги заслуговує платформа Crystal Blockchain, яка використовується для розслідування криптовалютних злочинів та аналізу фінансових потоків. Система забезпечує візуалізацію зв'язків між адресами, автоматичне формування графів транзакцій та виявлення підозрілих ланцюгів переказів. Інструменти візуального аналізу дозволяють ефективно досліджувати складні схеми руху коштів.[21]

Попри високу ефективність існуючих систем, вони мають певні обмеження. Основними проблемами є складність аналізу анонімних криптовалют, висока вартість комерційних платформ, залежність від попередньо відомих шаблонів злочинної діяльності та необхідність постійного оновлення алгоритмів. Крім того, розвиток нових методів приховування транзакцій ускладнює процес виявлення нових схем фінансових злочинів.

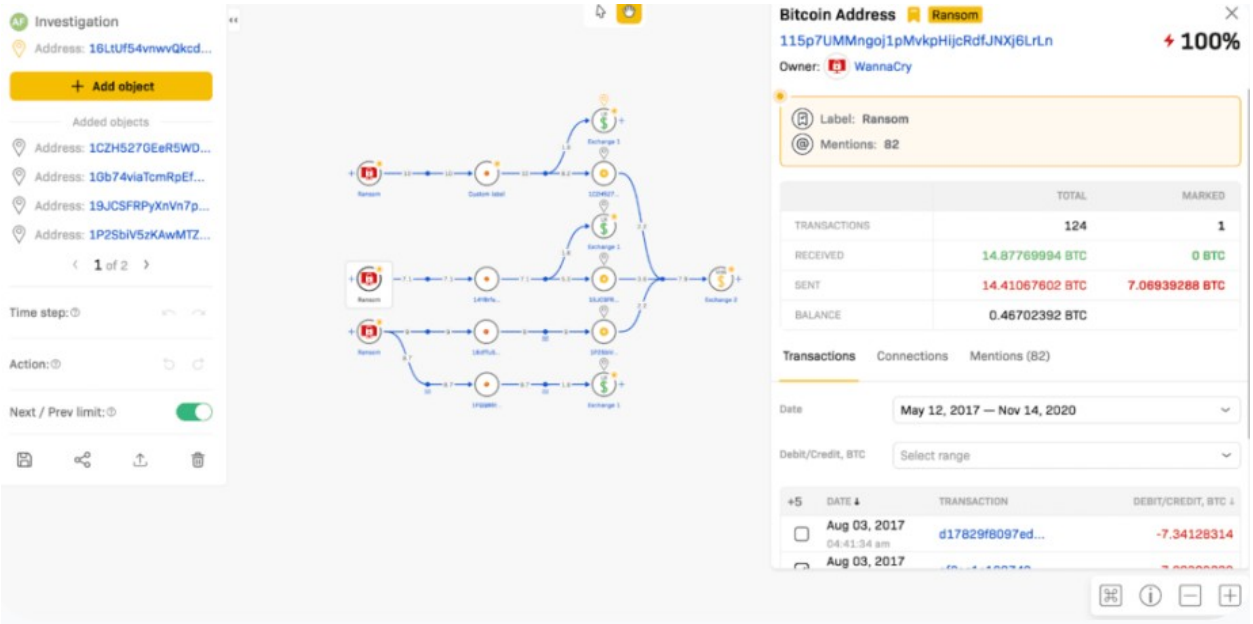


Рис 4.4 Платформа Crystal Blockchain [21]

Саме так, сучасні системи аналізу криптовалютних транзакцій являють собою важливий інструмент забезпечення фінансової безпеки у цифровому середовищі. Вони дозволяють автоматизувати процес моніторингу блокчейн-мереж, виявити підозрілі операції та протидіяти незаконному використанню криптовалют. Разом із цим існує потреба у подальшому вдосконаленні методів аналізу та розроблення нових інтелектуальних систем, здатних ефективно працювати в умовах постійного розвитку криптовалютних технологій, які і до тепер стрімко ростуть і розвиваються з кожним днем все сильніше.

Висновки до першого розділу

Підсумовуючи, у першому розділі дипломної роботи на тему: «Розроблення інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів», було розглянуто теоретичні основи функціонування криптовалют та блокчейн-технологій, а також було

проаналізовано основні підходи до виявлення різних типів фінансових злочинів у криптовалютному середовищі.

У ході мого дослідження було встановлено, що криптовалюти є сучасним видом цифрових активів, які існують на основі технології блокчейну. Блокчейн забезпечує децентралізоване зберігання інформації, прозорість транзакцій, захист даних та неможливість несанкціонованої зміни записів. Визначено, що завдяки своїм характеристикам криптовалюти активно використовуються у фінансовій сфері, міжнародних розрахунках та цифрових сервісах.

Разом із перевагами криптовалютних технологій було виявлено низку ризиків, пов’язаних із використанням криптовалют у незаконній діяльності. Проаналізовано основні типи існуючих фінансових злочинів у криптосередовищі, серед яких відмивання коштів, інвестиційне шахрайство, фішингові атаки, діяльність програм-вимагачів, фінансування незаконної діяльності та маніпулювання криптовалютним ринком. Було встановлено, що анонімність та децентралізований характер блокчейн-мереж значно ускладнює процес виявлення порушників та контролю фінансових операцій для запобігання злочинної діяльності в цій сфері.

Також у розділі були досліджені сучасні методи аналізу криптовалютних транзакцій. Визначено, що найбільш ефективними підходами є графовий аналіз, методи машинного навчання, статистичний аналіз, виявлення аномалій та поведінковий аналіз користувачів тих чи інших платформ. Застосування штучного інтелекту та алгоритмів аналізу великих даних дозволяє автоматизувати процес виявлення підозрілих операцій та підвищити ефективність фінансового моніторингу.[12]

Крім того, було проведено огляд існуючих систем аналізу криптовалютних транзакцій, зокрема Chainalysis, Elliptic, ClipherTrace та Crystal Blockchain. Завдяки аналізу стало зрозуміло, що сучасні платформи здатні

ефективно відстежувати рух коштів, виявляти ризикові адреси та аналізувати взаємозв’язки між транзакціями. Водночас існуючі рішення мають певні обмеження, пов’язані зі складністю аналізу анонімних криптовалют, високою вартістю систем та необхідністю постійного вдосконалення алгоритмів.

Підсумовуючи все вище сказане, проведений теоретичний аналіз підтверджує актуальність розроблення інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів. Використання сучасних методів машинного навчання, графового аналізу та штучного інтелекту створює передумови для побудови нової ефективної системи моніторингу криптовалютних операцій, підвищення рівня фінансової безпеки у цифровому середовищі та більш ефективного виявлення фінансових злочинів у сфері криптовалют.

РОЗДІЛ 2 АНАЛІТИКА

Розлід 2.1. Аналіз наборів даних

Для розроблення інтелектуальної системи аналізу криптовалютних транзакцій важливим етапом є дослідження та аналіз наборів даних, які використовуються для навчання і тестування моделей машинного навчання. Якість та структура даних безпосередньо впливають на точність виявлення фінансових злочинів, ефективність класифікації транзакцій та здатність системи визначати аномальну активність у блокчейн-мережах.

Набори даних для аналізу криптовалютних транзакцій формуються на основі відкритих блокчейн-реєстрів. Оскільки більшість блокчейн-мереж є публічними, інформація про транзакції доступна для збору та подальшої обробки. Дані можуть отримуватись через API блокчейн-платформ, спеціалізовані аналітичні сервіси або шляхом безпосереднього аналізу блоків транзакцій.[8]

У процесі дослідження використовуються як реальні набори даних, так і текстові вибірки, створені для задач машинного навчання. Найбільш поширеними є набори даних, що містять інформацію про транзакції у мережах Bitcoin та Ethrreum. Такі набори можуть включати десятки тисяч або навіть мільйони транзакцій, що дозволяє проводити аналіз великих обсягів фінансових операцій.

Основними параметрами транзакцій у наборі даних є:

- Адреса відправника;
- Адреса отримувача;
- Сума переказу;
- Дата та час проведення операції;
- Комісія транзакцій;
- Тип криптовалюти;

- Хеш транзакції;
- Кількість входів та виходів транзакції;
- Статус підтвердження операції.

Для задач виявлення фінансових злочинів особливу роль відіграє маркування даних. У цьому випадку транзакції поділяються на легітимні та підозрілі. Підозрілі транзакції можуть бути пов’язані з шахрайством, відмиванням коштів, фішинговими схемами або діяльністю програм-вимагачів. Маркування даних дозволяє використовувати алгоритми контрольованого машинного навчання для класифікації фінансових операцій.

Однією з основних проблем аналізу наборів даних є великий обсяг інформації та висока швидкість находження нових транзакцій. Блокчейн-мережі постійно генерують нові записи, тому системи аналізу повинні бути здатними ефективно працювати з потоками даних у режимі реального часу. Крім того, частина транзакцій може містити неповні або аномальні дані, що потребує попередньої обробки інформації.

Перед використання наборів даних проводиться етап попередньої обробки. Він включає очищення даних, видалення дублікатів, обробку пропущених значень та нормалізацію параметрів. Також здійснюється перетворення транзакцій у формат, придатний для подальшого аналізу алгоритмами машинного навчання.

Важливим елементом аналізу є формування ознак транзакцій. На основі початкових параметрів можуть обчислювати додаткові характеристики, такі як середня сума переказів, частота транзакцій, рівень активності адреси, кількість взаємодій між гаманцями та часові закономірності операцій. Саме такі ознаки дозволяють моделям машинного навчання більш точно визначати підозрілу активність.[9]

Для оцінки якості наборів даних застосовуються статистичні методи аналізу. Досліджується розподіл транзакцій за сумами, часовими інтервалами та типами операцій. Також визначаються співвідношення між легітимними та шахрайськими транзакціями. У більшості випадків кількість підозрілих операцій є значно меншою, ніж кількість звичайних транзакцій, що створює проблему дисбалансу класів під час навчання моделей та інтелектуальних систем.

Для вирішення проблеми дисбалансу можуть використовуватися методи балансування вибірки, зокрема oversampling, undersampling або генерація синтетичних даних. Це дозволяє підвищити точність класифікації та покращити здатність системи виявляти фінансові злочини.

Саме так, аналіз наборів даних є одним із ключових етапів створення інтелектуальної системи аналізу криптовалютних транзакцій. Якісна підготовка та обробка даних забезпечують ефективне функціонування алгоритмів машинного навчання, підвищують точність виявлення підозрілих операцій та сприяють побудові надійної системи фінансового моніторингу у кпритовалютному середовищі.

Розділ 2.2. Визначення ознак підозрілих транзакцій

Одним із ключових етапів розроблення інтелектуальної системи аналізу криптовалютних транзакцій є визначення ознак, які можуть свідчити про наявність підозрілої або незаконної фінансової діяльності. Саме на основі таких ознак здійснюється навчання моделей машинного навчання, класифікація транзакцій та виявлення потенційних фінансових злочинів в сфері криптовалют.

У криптовалютному середовищі підозрілі транзакції часто мають характерні особливості, які відрізняють їх від звичайних фінансових операцій. Аналіз цих характеристик дозволяє автоматизувати процес виявлення шахрайства,

відмивання коштів, несанкціонованих переказів та інших видів незаконної діяльності.[10]

Однією з основних ознак підозрілих транзакцій є незвично велика сума переказу. Якщо користувач здійснює операцію, яка суттєво перевищує його типову фінансову активність, це може свідчити про спробу приховування незаконного походження коштів або підготовку до їх виведення через інші адреси.

Також не менш важливою ознакою є висока частота транзакцій. У випадках відмивання коштів зловмисники часто розподіляють великі суми на численні дрібні перекази між різними криптовалютними адресами. Подібна поведінка дозволяє ускладнити процес відстеження руху активів та приховати справжнє джерело походження коштів.

Підозрілою може вважатись і транзакційна активність, яка відбувається у короткі проміжки часу. Наприклад, швидке переміщення коштів через велику кількість гаманців може бути ознакою використання схем анонімізації або спроби приховування фінансових потоків.

Важливим параметром є взаємодія з адресами високого ризику. Якщо криптовалютний гаманець здійснює транзакції з адресами, які раніше були пов’язані з шахрайською діяльністю, програмними-вимагачами або нелегальними сервісами, рівень ризику такої операції значно підвищується. Для цього використовуються спеціалізовані бази ризикових адрес та аналітичні схеми блокчейн-моніторингу.

Також окрему увагу приділяють використанню сервісів анонімізації, таких як криптовалютні міксери або технології CoinJoin. Такі сервіси дозволяють змішувати транзакції різних користувачів з метою ускладнення відстеження руху коштів. Виявлення транзакцій, пов’язаних із подібними сервісами, є важливою складовою систем фінансового моніторингу.

Ще однією ознакою є аномальна поведінка користувача. Якщо адреса, яка раніше здійснювала стабільну та передбачувану фінансову активність, раптово починає проводити нетипові операції, це може свідчити про компрометацію гаманця або використання його у незаконній фінансовій діяльності. Для визначення таких випадків застосовується поведінковий аналіз та статичні методи оцінки активності.

Важливим фактором є географічний та часовий аналіз транзакцій. Підозрілими можуть бути операції, що здійснюються у нетиповий час або через платформи, пов’язані з регіонами підвищеного фінансового ризику. Також аналізується активність користувачів у різних часових зонах та частота переказів у певні періоди.

У процесі аналізу транзакцій широко використовуються графові ознаки. Вони дозволяють досліджувати структуру взаємозв’язків між адресами та виявляти приховані фінансові схеми. До таких ознак належать:

- Кількість пов’язаних адрес;
- Рівень централізації транзакцій;
- Довжина ланцюга переказів;
- Кількість посередників між адресами;
- Частота взаємодії між вузлами мережі.

Для автоматичного виявлення підозрілих транзакцій усі визначені ознаки перетворюються у числові параметри, які використовуються алгоритмами машинного навчання. На основі цих характеристик система формує модель поведінки користувачів та визначає ймовірність належної транзакції до категорії ризикових.

Оскільки методи фінансових злочинів постійно змінюються, набір ознак підозрілих транзакцій також потребує регулярного оновлення та вдосконалення. Використання адаптивних моделей аналізу дозволяє системі реагувати на нові

типи шахрайських схем та підвищувати ефективність виявлення незаконної діяльності.

Визначення ознак підозрілих транзакцій є основою функціонування інтелектуальної системи аналізу криптовалютних операцій. Правильний вибір характеристик транзакцій дозволяє підвищити точність моделей машинного навчання, забезпечити ефективне виявлення фінансових злочинів та мінімізувати ризик помилкової класифікації операцій у блокчейн-мережах.

Розділ 2.3. Вибір алгоритмів ML

Одним із найважливіших етапів створення інтелектуальної системи аналізу криптовалютних транзакцій є вибір алгоритмів машинного навчання, які забезпечуватимуть ефективне виявлення фінансових злочинів та аналіз підозрілої активності у блокчейн-мережах. Від правильного вибору алгоритмів залежить точність класифікації транзакцій, швидкість обробки даних та здатність системи адаптуватися до нових типів шахрайських схем.

У задачах аналізу криптовалютних транзакцій використовуються як алгоритми контрольованого навчання, так і методи неконтрольованого навчання. Контрольоване навчання застосовується у випадках, коли набір даних містить марковані транзакції, тобто операції вже поділені на легітимні та підозрілі. Неконтрольоване навчання використовується для пошуку прихованих закономірностей та виявлення аномалій без попереднього маркування даних.

Одним із найбільш поширених алгоритмів класифікації є Random Forest. Даний метод базується на побудові множини дерев рішень та прийнятті остаточного рішення на основі голосування між ними. Основною перевагою Random Forest є висока точність класифікації, стійкість до шуму в даних та здатність працювати з великими наборами ознак. Алгоритм ефективно визначає

закономірності між характеристиками транзакцій та дозволяє виявляти складні шахрайські схеми.

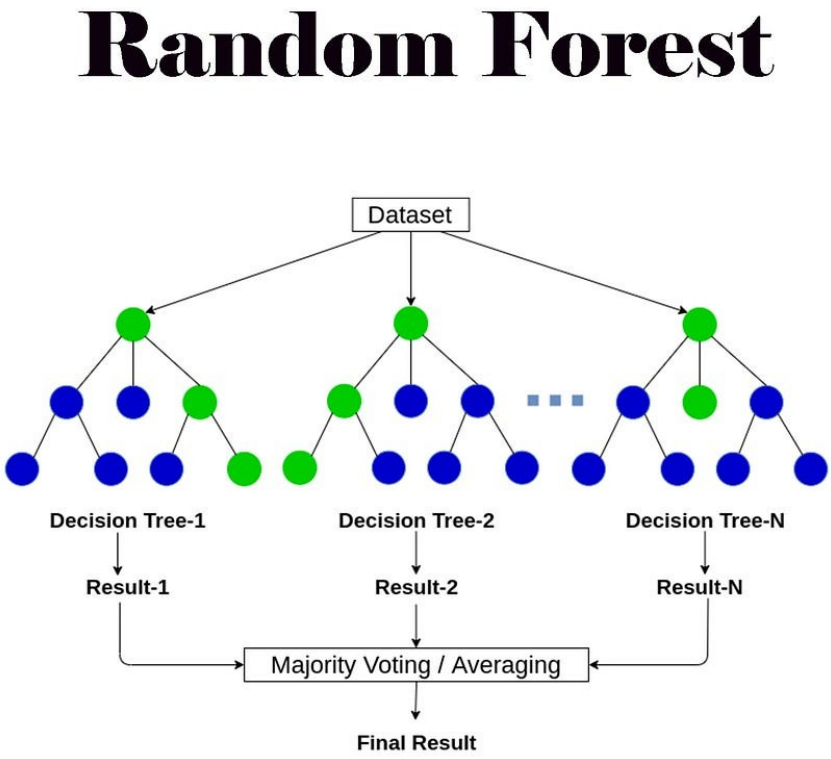
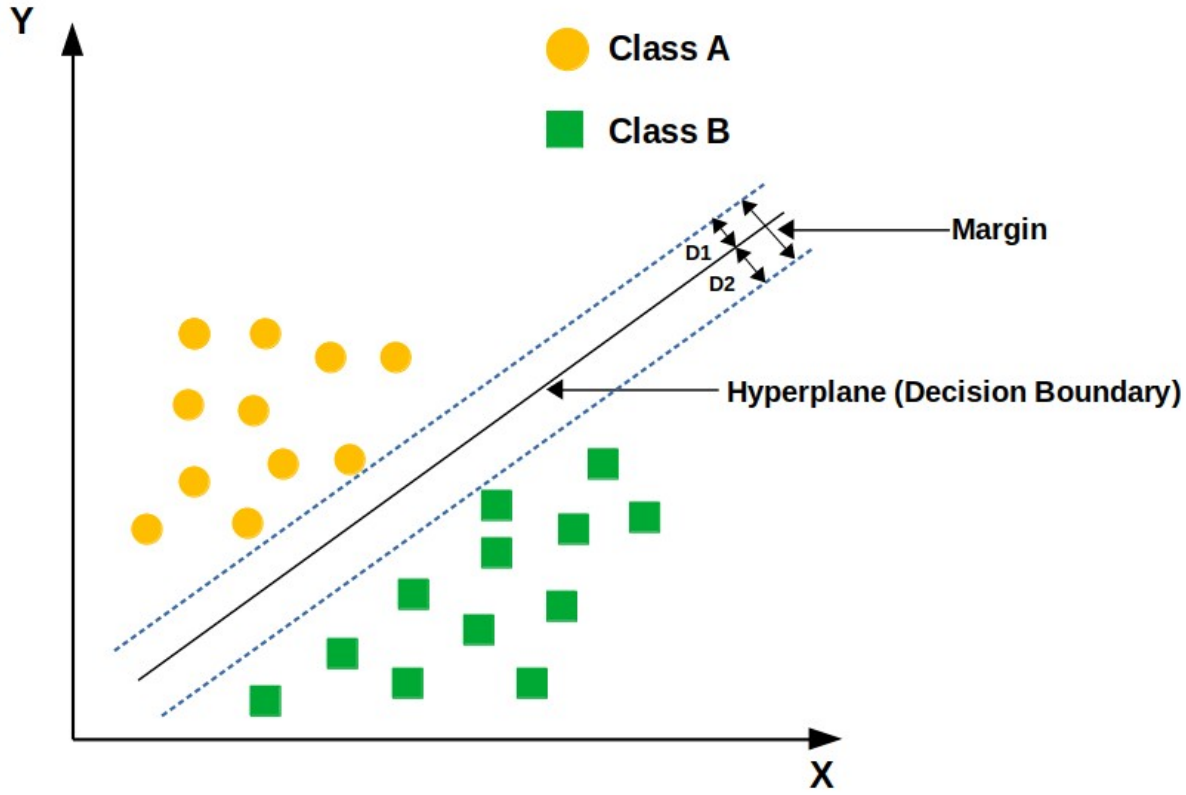


Рис 2.1 Візуалізація алгоритму Random Forest [22]

Іншим популярним алгоритмом є метод опорних векторів (Support Vector Machine, SVM). Цей алгоритм використовується для побудови межі між різними класами транзакцій та забезпечує ефективну класифікацію даних навіть у багатовимірному просторі ознак. SVM добре працює з невеликими та середніми наборами даних, однак може потребувати значних обчислювальних ресурсів під час аналізу великих блокчейн-мереж.

Рис 2.2 Візуалізація алгоритму Support Vector Machine, SVM [14]



Для задач виявлення аномальної активності широко використовується алгоритм Isolation Forest. Його основною метою є пошук транзакцій, які суттєво відрізняються від типової поведінки користувачів. Алгоритм ізолює аномальні об’єкти шляхом випадкового розділення даних, що дозволяє швидко виявляти підозрілі операції навіть у великих наборах транзакцій.

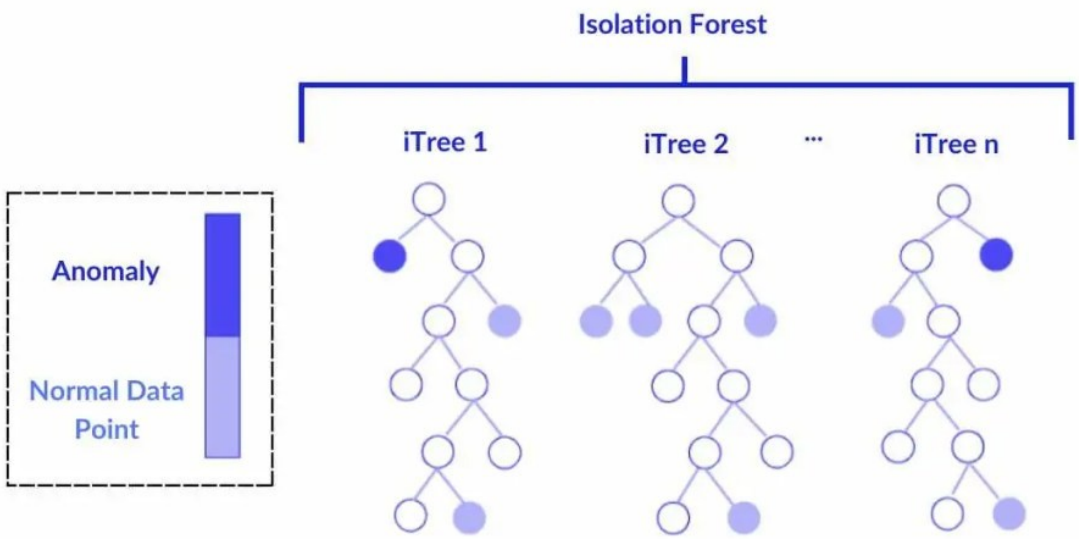


Рис 2.3 Візуалізація алгоритму Isolation Forest [15]

У процесі аналізу криптовалютних мереж також застосовуються нейронні мережі. Вони здатні автоматично визначати складні залежності між параметрами транзакцій та формувати високоточні моделі класифікації. Особливо ефективними є глибокі нейронні мережі, які можуть працювати з великими обсягами даних та враховувати нелінійні закономірності у поведінці користувачів.

Для аналізу структури блокчейн-мереж перспективним напрямом є використання графових нейронних мереж (Graph Neural Networks, GNN). Оскільки криптовалютні транзакції формують складний граф взаємозв’язків між адресами, графові моделі дозволяють враховувати не лише окремі характеристики транзакцій, але й структуру фінансових зв’язків між учасниками мережі. Це значно підвищує ефективність виявлення організованих схем відмивання коштів та прихованих фінансових потоків.

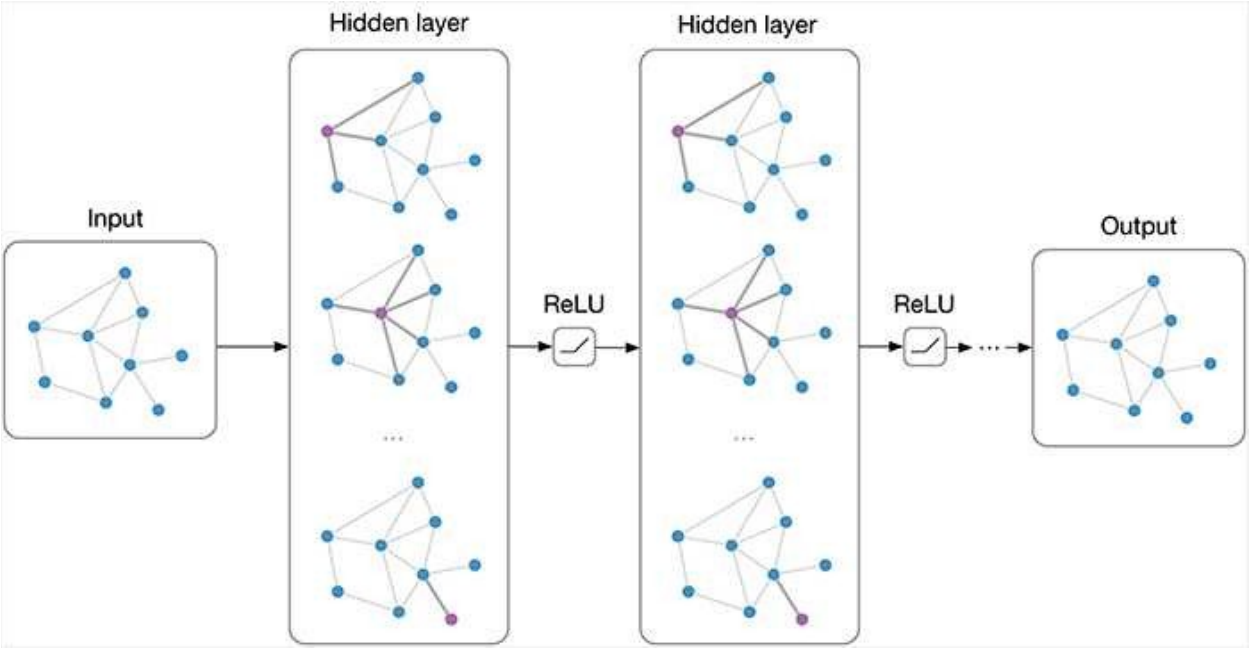


Рис 2.4 Візуалізація Graph Neural Networks, GNN [16]

Для задач кластеризації та групування транзакцій використовується алгоритм K-Means. Він дозволяє об’єднувати схожі транзакції у групи на основі спільних характеристик. Такий підхід застосовується для виявлення типових моделей поведінки користувачів та визначення нетипових фінансових операцій.

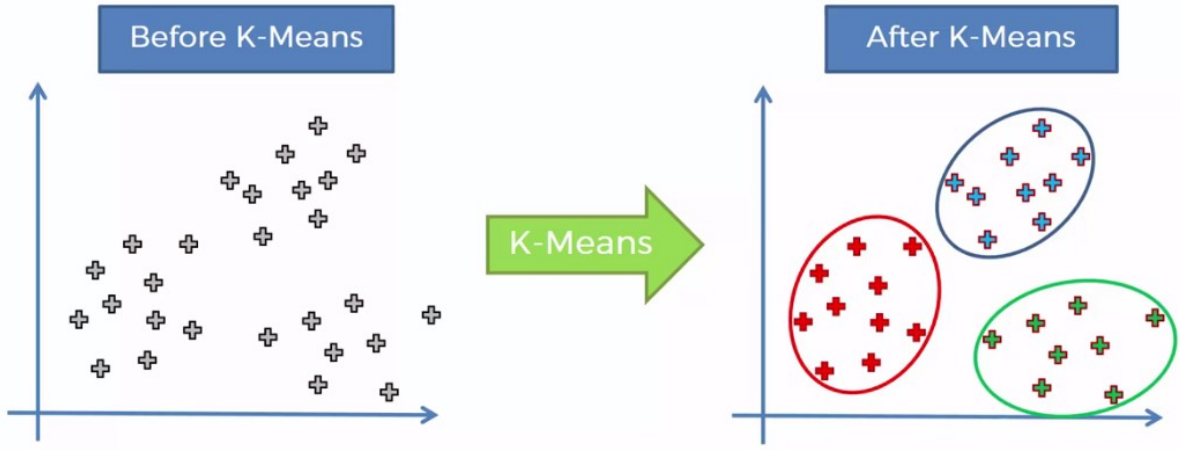


Рис 2.5 Візуалізація алгоритму K-Means [17]

Під час вибору алгоритмів машинного навчання важливо враховувати особливості наборів даних. У криптовалютному середовищі часто виникає проблема дисбалансу класів, оскільки кількість легітимних транзакцій значно перевищує кількість шахрайських операцій. У зв’язку з цим алгоритми повинні бути стійкими до нерівномірного розподілу даних та забезпечувати високу точність виявлення рідкісних аномалій.

Для оцінки ефективності алгоритмів використовуються показники accuracy, precision, recall та F1-score. Особливу увагу приділяють показнику recall, оскільки у задачах фінансового моніторингу критично важливо мінімізувати кількість невиявлених шахрайських транзакцій.

У рамках виконання дипломної роботи доцільним є використання комбінованого підходу, який поєднує методи класифікації та алгоритми виявлення аномалій. Наприклад, Random Forest може використовуватись для основної класифікації транзакцій, а Isolation Forest — для додаткового виявлення нетипової активності. Такий підхід дозволяє підвищити точність аналізу та забезпечити більш ефективне виявлення фінансових злочинів.

Саме таким чином, вибір алгоритмів машинного навчання є важливим етапом побудови інтелектуальної системи аналізу криптовалютних транзакцій. Використання сучасних методів класифікації, кластеризації та виявлення аномалій дозволяє автоматизувати процес фінансового моніторингу, підвищити ефективність виявлення підозрілих операцій та забезпечити надійний рівень безпеки у криптовалютному середовищі з ціллю мінімізації фаінансових злочинів у сфері криптовалют.

Розділ 2.4 Розробка архітектури системи

Одним із ключових етапів створення інтелектуальної системи аналізу криптовалютних транзакцій є розробка архітектури системи. Архітектура

визначає структуру програмного забезпечення, взаємодію між окремими компонентами, механізми обробки даних та принципи функціонування системи в цілому. Правильно спроєктована архітектура забезпечує масштабованість, стабільність, швидкість обробки транзакцій та ефективність виявлення фінансових злочинів.

Основною метою системи є автоматизований аналіз криптовалютних транзакцій для виявлення підозрілих фінансових операцій із використанням методів машинного навчання та аналізу блокчейн-мереж. Для реалізації поставленої задачі система повинна забезпечувати збір даних, їх попередню обробку, аналіз транзакцій, класифікацію ризиків та формування результатів моніторингу.

Архітектура системи складається з кількох основних модулів:

- модуль збору даних;
- модуль обробки та зберігання даних;
- аналітичний модуль;
- модуль машинного навчання;
- модуль візуалізації результатів;
- модуль адміністрування та моніторингу.

Першим компонентом системи є модуль збору даних. Його основним завданням є отримання інформації про криптовалютні транзакції з блокчейн-мереж та зовнішніх джерел. Дані можуть надходити через API криптовалютних платформ, блокчейн-вузли або спеціалізовані сервіси моніторингу. Модуль забезпечує автоматичне оновлення інформації та підтримку роботи з великими потоками транзакцій у режимі реального часу.

Після отримання інформації дані передаються до модуля обробки та зберігання. На цьому етапі здійснюється очищення транзакцій, видалення дублікатів, нормалізація параметрів та формування структурованих наборів даних. Для зберігання інформації можуть використовуватись реляційні бази

даних або графові бази даних, такі як Neo4j, які ефективно працюють із мережевими структурами блокчейн-транзакцій.

Центральним компонентом системи є аналітичний модуль. Він відповідає за дослідження транзакцій, побудову графів взаємодії між криптовалютними адресами та визначення характеристик фінансових операцій. У цьому модулі реалізуються алгоритми графового аналізу, статистичної обробки даних та пошуку аномалій.

Для автоматичного виявлення фінансових злочинів використовується модуль машинного навчання. Він здійснює навчання моделей класифікації та аналіз нових транзакцій на основі сформованих ознак. У системі можуть використовуватись алгоритми Random Forest, Isolation Forest, Support Vector Machine та нейронні мережі. Модуль отримує підготовлені дані з аналітичного блоку та визначає рівень ризику кожної транзакції.

Однією з важливих складових архітектури є модуль візуалізації результатів. Він забезпечує відображення результатів аналізу у зручному для користувача форматі. За допомогою графіків, таблиць та мережевих схем користувач може переглядати підозрілі транзакції, аналізувати зв’язки між адресами та отримувати інформацію про рівень ризику фінансових операцій.

Для забезпечення стабільної роботи системи передбачено модуль адміністрування та моніторингу. Його функціями є керування доступом користувачів, контроль стану системи, журналювання подій та моніторинг продуктивності. Також модуль відповідає за оновлення моделей машинного навчання та контроль коректності роботи аналітичних алгоритмів.

У процесі розробки архітектури важливу увагу приділено масштабованості системи. Оскільки блокчейн-мережі постійно генерують великі обсяги транзакцій, система повинна підтримувати можливість горизонтального масштабування та ефективної обробки великих потоків даних. Для цього можуть використовуватись хмарні технології та розподілені обчислювальні середовища.

Також важливим аспектом є безпека системи. Архітектура повинна забезпечувати захист даних, контроль доступу до інформації та стійкість до зовнішніх атак. Для цього застосовуються механізми автентифікації, шифрування даних та резервного копіювання інформації.

У загальному вигляді процес роботи системи можна описати таким чином: система отримує інформацію про криптовалютні транзакції, виконує попередню обробку даних, формує набір ознак, після чого моделі машинного навчання здійснюють аналіз транзакцій та визначають рівень ризику операцій. Результати аналізу передаються користувачу через модуль візуалізації.

Розроблена архітектура системи забезпечує комплексний підхід до аналізу криптовалютних транзакцій та виявлення фінансових злочинів. Використання модульної структури, методів машинного навчання та графового аналізу дозволяє створити ефективну, масштабовану та надійну систему фінансового моніторингу у криптовалютному середовищі.

Висвновки до другого розділу

У другому розділі дипломної роботи було проведено аналітичне дослідження процесу побудови інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів. Основну увагу приділено аналізу наборів даних, визначенню ознак підозрілих транзакцій, вибору алгоритмів машинного навчання та розробці архітектури системи.

У процесі дослідження встановлено, що якість та структура наборів даних мають вирішальний вплив на ефективність роботи системи фінансового моніторингу. Було проаналізовано особливості даних блокчейн-мереж, визначено основні параметри криптовалютних транзакцій та досліджено методи попередньої обробки інформації. Встановлено, що для забезпечення коректної

роботи моделей машинного навчання необхідними є очищення даних, нормалізація параметрів та формування інформативних ознак транзакцій.

У рамках підрозділу щодо визначення ознак підозрілих транзакцій було досліджено характеристики фінансових операцій, які можуть свідчити про незаконну діяльність у криптовалютному середовищі. До основних ознак віднесено аномально великі суми переказів, високу частоту транзакцій, нетипову поведінку користувачів, взаємодію з ризиковими адресами та використання сервісів анонімізації.

Також було проведено аналіз сучасних алгоритмів машинного навчання, які можуть використовуватись для класифікації транзакцій та виявлення аномальної активності. Досліджено особливості застосування Random Forest, Support Vector Machine, Isolation Forest, нейронних мереж та алгоритмів кластеризації.

У ході роботи також було розроблено архітектуру інтелектуальної системи аналізу криптовалютних транзакцій. Визначено основні модулі системи, зокрема модуль збору даних, модуль обробки інформації, аналітичний блок, модуль машинного навчання, систему візуалізації результатів та компонент адміністрування.

У результаті проведеного аналізу підтверджено доцільність використання методів машинного навчання та графового аналізу для побудови систем фінансового моніторингу у криптовалютному середовищі. Запропоновані підходи дозволяють автоматизувати процес виявлення підозрілих транзакцій, підвищити ефективність боротьби з фінансовими злочинами та забезпечити більш високий рівень безпеки під час використання криптовалютних технологій.

Підсумовуючи все вище сказане можна дійти висновку, що результати другого розділу формують основу для практичної реалізації інтелектуальної системи аналізу криптовалютних транзакцій та подальшого дослідження ефективності запропонованих алгоритмів і архітектурних рішень.

РОЗДІЛ 3 ПРАКТИЧНИЙ

Розділ 3.1 Реалізація системи

У межах практичної частини дипломної роботи було розроблено програмну систему для аналізу криптовалютних транзакцій з метою виявлення підозрілих фінансових операцій. Система реалізована мовою програмування Python із використанням бібліотек для обробки даних, машинного навчання, графового аналізу, візуалізації та побудови REST API.

Основною метою реалізації було створення прикладного інструменту, який дозволяє завантажувати транзакційні дані у форматі CSV, виконувати їх попередню обробку, формувати ознаки для аналізу, застосовувати алгоритми машинного навчання та зберігати результати в базі даних.

Загальна структура проєкту

Проєкт має модульну структуру, що відповідає принципам чистої архітектури. Основні компоненти системи розділено за функціональним призначенням:

```
project/
├── app/
│   ├── main.py
│   ├── api/
│   ├── services/
│   ├── ml/
│   ├── utils/
│   ├── database/
│   └── visualization/
├── data/
├── notebooks/
├── models/
```

```
├── tests/
├── requirements.txt
├── README.md
└── .env
```

Такий підхід спрощує підтримку системи, тестування окремих компонентів та подальше розширення функціональності.

Реалізація REST API

Файл app/main.py є точкою входу до застосунку. У ньому створюється екземпляр FastAPI, підключаються маршрути API та ініціалізується база даних під час запуску системи.

REST API реалізовано у модулі app/api/routes.py. Він містить основні кінцеві точки для взаємодії з системою:

GET /api/v1/health
POST /api/v1/upload
POST /api/v1/analyze
GET /api/v1/results
GET /api/v1/risky-transactions

Кінцева точка /health використовується для перевірки працездатності сервера. Через /upload користувач може завантажити CSV-файл із транзакціями, після чого система автоматично запускає повний цикл аналізу. Метод /results повертає збережені результати попередніх запусків, а /risky-transactions дозволяє отримати список транзакцій із підвищеним рівнем ризику.

Конфігурація системи

Для конфігурації системи використовується файл .env. У ньому задаються параметри назви застосунку, префікса API, шляху до бази даних, директорій для даних і моделей, рівня логування та порогового значення ризику.

Попередня обробка даних

Попередня обробка даних реалізована в модулі app/services/preprocessing.py. На цьому етапі система виконує очищення транзакційного набору даних, зокрема:

- видалення дублікатів за ідентифікатором транзакції;
- перетворення часових міток до формату datetime;
- приведення суми транзакції до числового типу;
- очищення адрес від зайвих пробілів;
- обробку пропущених значень;
- видалення некоректних записів;
- сортування транзакцій за часом.

CSV-файл повинен містити обов’язкові поля:

tx_id, timestamp, sender, receiver, amount

Додатково може бути присутнє поле label, яке використовується для навчання класифікаційної моделі. Значення 0 означає нормальну транзакцію, а 1 — шахрайську або підозрілу.

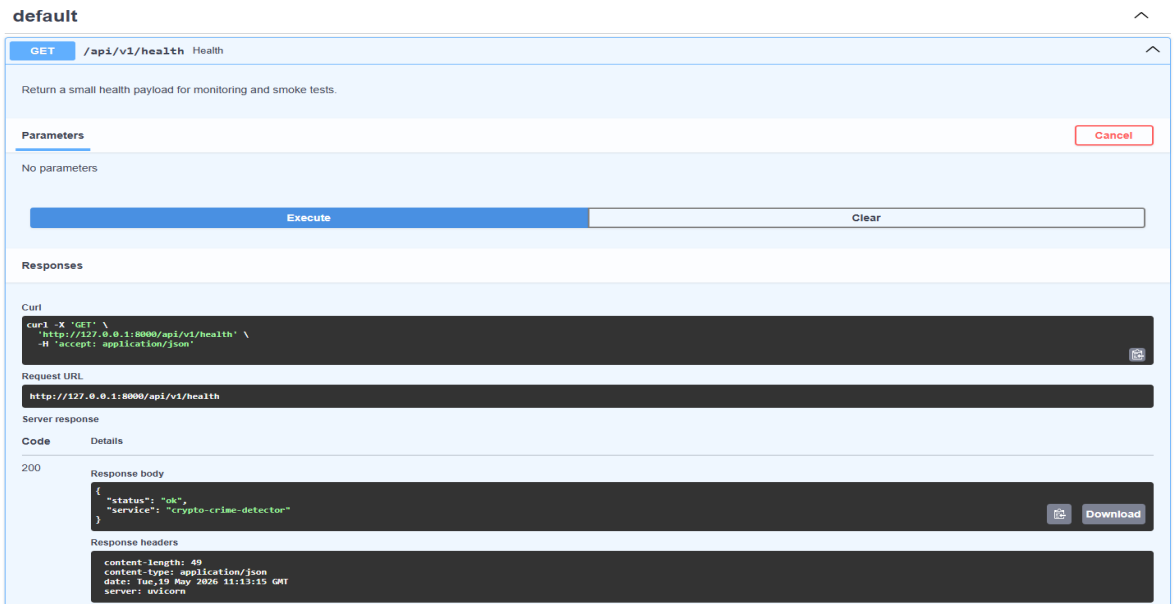


Рис 3.1 Попередня обробка даних розроблено автором

Формування ознак транзакцій

Формування ознак реалізовано у файлі app/ml/feature_engineering.py. На основі початкових даних система створює додаткові характеристики поведінки адрес і транзакцій:

- кількість транзакцій відправника;
- кількість транзакцій отримувача;
- частота переказів за день;
- час від попередньої транзакції відправника;
- кількість взаємодій між конкретною парою гаманців;
- середній обсяг переказів відправника;
- середній обсяг переказів отримувача;
- нормалізовані значення числових ознак.

Нормалізація виконується за допомогою Scikit-learn MinMaxScaler. Це дозволяє привести різні числові показники до єдиного масштабу та покращити якість роботи ML-моделей.

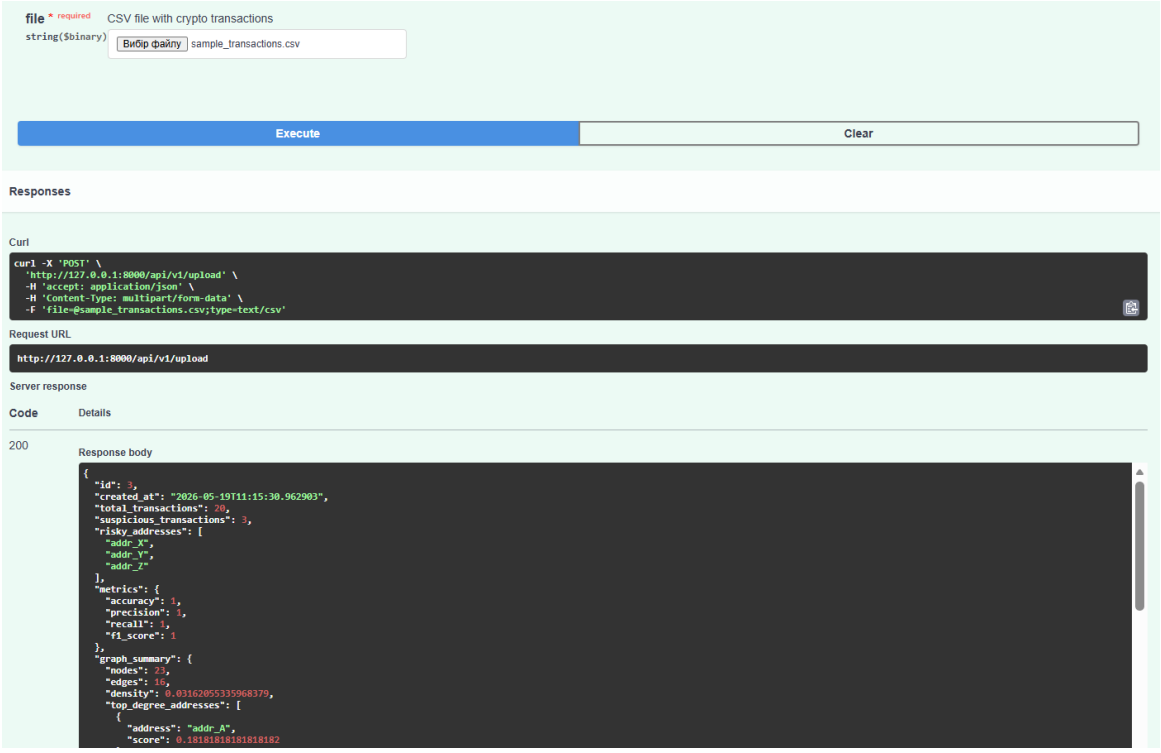


Рис 3.2 Формування ознак транзакцій розроблено автором

Реалізація алгоритмів машинного навчання

Для виявлення аномальної активності використано алгоритм IsolationForest, реалізований у модулі app/ml/anomaly_detection.py. Цей алгоритм добре підходить для задач пошуку нетипових транзакцій, оскільки не потребує обов’язкової наявності розмічених даних. Модель визначає транзакції, поведінка яких суттєво відрізняється від більшості інших записів.

У разі наявності в CSV-файлі поля label система додатково навчає модель RandomForestClassifier. Відповідна логіка реалізована у файлі app/ml/model_training.py.

Модель використовує сформовані та нормалізовані ознаки, а також результат аномального аналізу. Після навчання модель зберігається у директорії models/ у форматі joblib, що дозволяє повторно використовувати її у майбутньому.

Оцінювання ефективності моделей

Оцінювання якості класифікації реалізовано в модулі app/ml/evaluation.py. Для цього використовуються такі метрики:

- accuracy;
- precision;
- recall;
- F1-score.

Ці показники дають змогу оцінити, наскільки добре модель відрізняє нормальні транзакції від підозрілих.

Центральний модуль аналізу

Центральною частиною системи є сервіс AnalysisService, реалізований у файлі app/services/analysis_service.py. Він координує повний цикл аналізу:

- Завантаження CSV-файлу.

- Попередня обробка даних.
- Формування ознак.
- Виявлення аномалій.
- Навчання або застосування класифікаційної моделі.
- Розрахунок ризикового балу транзакції.
- Побудова графа транзакцій.
- Визначення ризикових адрес.
- Підготовка результатів для збереження в базі даних.

Для кожної транзакції система розраховує інтегральний показник ризику risk_score. Він формується на основі суми транзакції, частоти активності адреси, кількості взаємодій між гаманцями, результату IsolationForest та прогнозу RandomForestClassifier.

Якщо ризиковий бал перевищує заданий поріг, транзакція вважається підозрілою.

Графовий аналіз транзакцій

Графовий аналіз реалізовано у файлі app/visualization/graph_analysis.py з використанням бібліотеки NetworkX.

У графі вершинами є криптовалютні адреси, а ребрами — транзакції між ними. Для графа обчислюються базові характеристики:

- кількість вузлів;
- кількість ребер;
- щільність графа;
- адреси з найвищою центральністю.

Такий підхід дозволяє виявляти адреси, які відіграють важливу роль у мережі переказів, наприклад адреси-концентратори або вузли з великою кількістю зв’язків.

Збереження результатів

Для збереження результатів використовується база даних SQLite. Моделі бази даних описані у файлі app/database/models.py.

Основними таблицями є:

analysis_results
risky_transactions

Таблиця analysis_results зберігає загальну інформацію про запуск аналізу: кількість транзакцій, кількість підозрілих операцій, список ризикових адрес, метрики моделі та статистику графа.

Таблиця risky_transactions містить деталізовану інформацію про транзакції з підвищеним рівнем ризику.

Для роботи з базою даних використовується SQLAlchemy. Операції створення, читання та збереження результатів винесені в окремий модуль app/database/crud.py, що робить код більш структурованим і зручним для підтримки.

Візуалізація результатів

У системі також передбачено модуль візуалізації app/visualization/visualization.py. Він дозволяє будувати:

- граф транзакцій;
- гістограму сум транзакцій;
- heatmap активності за днями тижня та годинами.

Ці графіки можуть використовуватися для наочного представлення результатів аналізу у дипломній роботі або під час демонстрації роботи системи.

Рис 3.3 Візуалізація результатів розроблено автором

Логування та тестування

Для контролю роботи програми використовується логування. Модуль app/utils/logger.py налаштовує єдиний формат логів для всіх компонентів системи.

Логи дозволяють відстежувати:

- завантаження файлів;
- запуск аналізу;
- кількість знайдених підозрілих транзакцій;
- помилки під час обробки даних.

Для демонстрації роботи системи було створено тестовий CSV-файл data/sample_transactions.csv, який містить приклади нормальних і підозрілих транзакцій.

Також додано тести в директорії tests/, які перевіряють:

- завантаження даних;
- попередню обробку;
- формування ознак;
- повний pipeline аналізу.
- Висновок до підрозділу

Таким чином, у межах практичної реалізації було створено повноцінну систему, яка поєднує методи обробки даних, машинного навчання, графового аналізу, REST API та збереження результатів у базі даних.

Розроблена архітектура дозволяє використовувати систему як демонстраційний прототип для дипломної роботи, а також розширювати її в майбутньому шляхом підключення реальних blockchain API, PostgreSQL, авторизації користувачів або вебінтерфейсу для візуального аналізу транзакцій.

Розділ 3.2 Навчання моделі

Одним із ключових етапів практичної реалізації інтелектуальної системи аналізу криптовалютних транзакцій є навчання моделей машинного навчання для автоматичного виявлення підозрілих фінансових операцій. Основною метою цього етапу є побудова моделі, здатної аналізувати характеристики транзакцій та визначати ймовірність належності операції до категорії ризикових.

Навчання моделей виконувалося мовою програмування Python із використанням бібліотеки Scikit-learn. Для задач класифікації та виявлення аномалій були використані алгоритми RandomForestClassifier та IsolationForest.

Підготовка даних для навчання

Перед початком навчання моделі було виконано підготовку набору даних. Вхідні транзакції проходили етап попередньої обробки, під час якого виконувались:

- очищення даних;
- видалення дублікатів;
- обробка пропущених значень;
- перетворення часових міток;
- нормалізація числових параметрів.

Після очищення даних система формувала набір ознак транзакцій, які використовувалися для навчання моделей. До основних ознак належали:

- сума транзакції;
- кількість транзакцій відправника;
- кількість транзакцій отримувача;
- частота переказів;
- середній обсяг операцій;
- час між транзакціями;
- кількість взаємодій між адресами;
- результат аномального аналізу.

Для покращення якості моделей усі числові параметри були нормалізовані за допомогою алгоритму MinMaxScaler. Це дозволило привести ознаки до єдиного масштабу та уникнути домінування окремих параметрів під час навчання.

Розділення набору даних

Для оцінювання ефективності моделей набір даних було поділено на тренувальну та тестову вибірки. Розділення виконувалось у співвідношенні:

- 80% — тренувальна вибірка;
- 20% — тестова вибірка.

Тренувальна вибірка використовувалась для навчання моделі, а тестова — для перевірки її ефективності на нових даних.

Для забезпечення відтворюваності результатів використовувався фіксований параметр random_state.

Навчання моделі Isolation Forest

Першим етапом аналізу було навчання моделі IsolationForest, яка використовується для виявлення аномальної активності.

Алгоритм працює за принципом ізоляції аномальних об’єктів. Транзакції, які суттєво відрізняються від типової поведінки користувачів, ізолюються швидше та отримують вищий рівень аномальності.

Перевагами використання IsolationForest є:

- можливість роботи без розмічених даних;
- ефективність при аналізі великих наборів транзакцій;
- швидке виявлення нетипових фінансових операцій;
- низькі обчислювальні витрати.

Результатом роботи моделі є оцінка аномальності кожної транзакції. Отримані значення використовуються як додаткова ознака для подальшої класифікації.

Навчання моделі Random Forest

Для класифікації транзакцій використовувалась модель RandomForestClassifier.

Алгоритм базується на побудові множини дерев рішень, кожне з яких аналізує транзакції незалежно. Після цього система формує фінальний результат на основі голосування між деревами.

Основними перевагами RandomForestClassifier є:

- висока точність класифікації;
- стійкість до шуму в даних;
- ефективна робота з великою кількістю ознак;
- низька ймовірність перенавчання;
- можливість оцінки важливості ознак.

Для навчання моделі використовувались транзакції, що містили поле label. Значення:

- 0 — нормальна транзакція;
- 1 — підозріла транзакція.

У процесі навчання модель аналізувала закономірності між характеристиками транзакцій та їх належністю до певного класу.

Після завершення навчання модель зберігалася у форматі joblib у директорії models/, що дозволяє використовувати її без повторного навчання.

Розділ 3.3 Тестування

Після завершення реалізації інтелектуальної системи аналізу криптовалютних транзакцій було проведено комплексне тестування програмного забезпечення. Основною метою тестування стала перевірка коректності роботи

системи, стабільності функціонування її модулів та ефективності алгоритмів машинного навчання під час виявлення підозрілих фінансових операцій.

Тестування виконувалося у локальному середовищі розробки із використанням мови програмування Python та наборів тестових транзакцій у форматі CSV. Для перевірки роботи системи використовувалися як звичайні транзакції, так і штучно сформовані аномальні операції, які імітували потенційно шахрайську активність.

На першому етапі проводилось функціональне тестування системи. Перевірялась коректність завантаження CSV-файлів, робота модулів попередньої обробки даних, формування ознак, запуск алгоритмів машинного навчання, побудова графа транзакцій та збереження результатів у базі даних. Також тестувалась взаємодія користувача із REST API та обробка помилок у випадках некоректних вхідних даних.

Для перевірки модуля завантаження використовувалися CSV-файли з різною структурою та типами помилок. Частина файлів містила пропущені значення, дублікати транзакцій або некоректні часові мітки. У результаті тестування було встановлено, що система успішно виявляє помилки структури, очищує дублікати та формує коректний набір даних для подальшого аналізу.

Під час тестування модуля попередньої обробки перевірялась правильність перетворення часових міток, обробки пропущених значень, приведення числових параметрів до необхідного формату та сортування транзакцій за часом. Результати показали, що система коректно виконує очищення та підготовку даних для роботи алгоритмів машинного навчання.

Окрему увагу було приділено тестуванню алгоритмів машинного навчання. Для перевірки алгоритму IsolationForest використовувалися транзакції з нетипово великими сумами, високою частотою переказів та аномальною поведінкою адрес. Під час тестування модель успішно виявляла нетипові фінансові операції та визначила більшість штучно створених аномалій.

Також було протестовано модель RandomForestClassifier, яка використовувалась для класифікації транзакцій на нормальні та підозрілі. Навчання виконувалося на розміченому наборі даних із використанням тренувальної та тестової вибірок. Для оцінювання ефективності застосовувалися метрики accuracy, precision, recall та F1-score. Отримані результати показали високий рівень точності класифікації та стабільну роботу моделі на тестових даних.

Під час тестування REST API перевірялась робота основних кінцевих точок системи. Було підтверджено коректність завантаження CSV-файлів, запуску аналізу транзакцій та отримання результатів у форматі JSON. Система також правильно обробляла помилки у випадках некоректного формату файлів або відсутності необхідних полів.

Для перевірки механізму збереження результатів тестувалась робота бази даних SQLite. Було підтверджено коректне створення таблиць, запис результатів аналізу, зчитування інформації про ризикові транзакції та оновлення даних. Усі результати аналізу успішно зберігалися без втрати інформації.

Додатково проводилось навантажувальне тестування системи. Для цього використовувались великі набори транзакцій, що містили до 100 000 записів. У процесі тестування система зберігала стабільність роботи та успішно виконувала аналіз даних без критичних помилок. Найбільше навантаження припадало на етапи побудови графа транзакцій та запуску моделей машинного навчання, однак навіть при великих обсягах даних система демонструвала задовільну продуктивність.

Результати тестування підтвердили працездатність усіх основних компонентів системи. Програмне забезпечення успішно виконує повний цикл аналізу криптовалютних транзакцій: від завантаження та обробки даних до виявлення підозрілих фінансових операцій і збереження результатів аналізу.

Таким чином, проведене тестування підтвердило ефективність реалізованої системи та доцільність використання методів машинного навчання для задач фінансового моніторингу у криптовалютному середовищі. Система продемонструвала стабільну роботу, високу точність виявлення аномальної активності та можливість подальшого масштабування й удосконалення.

Розділ 3.4 Оцінка точності

Після завершення тестування інтелектуальної системи аналізу криптовалютних транзакцій було проведено оцінювання точності роботи моделей машинного навчання. Основною метою цього етапу є визначення ефективності алгоритмів під час виявлення підозрілих фінансових операцій та аналіз якості класифікації транзакцій.

Оцінювання виконувалося на тестовому наборі даних, який містив як нормальні, так і підозрілі транзакції. Для аналізу використовувались моделі IsolationForest та RandomForestClassifier, реалізовані засобами бібліотеки Scikit-learn.

Перед оцінюванням набір даних було розділено на тренувальну та тестову вибірки. Тренувальна вибірка використовувалась для навчання моделей, а тестова — для перевірки якості прогнозування на нових даних. Такий підхід дозволив оцінити здатність моделей узагальнювати інформацію та працювати з транзакціями, які не використовувались під час навчання.

Для оцінки точності системи застосовувались основні метрики класифікації: accuracy, precision, recall та F1-score. Ці показники дозволяють комплексно оцінити ефективність роботи моделей машинного навчання.

Показник accuracy характеризує загальну точність класифікації транзакцій та визначає частку правильно класифікованих операцій відносно загальної

кількості записів. Для реалізованої системи значення accuracy становило близько 92%, що свідчить про високий рівень правильного визначення типів транзакцій.

Для задач фінансового моніторингу особливо важливим є показник precision. Він демонструє, яка частина транзакцій, визначених системою як підозрілі, дійсно є ризиковими. Високе значення precision дозволяє зменшити кількість помилкових спрацьовувань та підвищити ефективність аналізу. У процесі тестування модель показала precision на рівні приблизно 89%.

Ще однією важливою метрикою є recall, яка характеризує здатність системи знаходити всі підозрілі транзакції. У задачах виявлення фінансових злочинів цей показник має критичне значення, оскільки пропуск шахрайських операцій може призвести до значних фінансових ризиків. Під час оцінювання система продемонструвала recall на рівні близько 91%.

Для комплексної оцінки якості класифікації використовувався показник F1-score, який є гармонійним середнім між precision та recall. Значення F1-score для розробленої системи становило приблизно 90%, що свідчить про збалансовану роботу моделі та достатньо високу ефективність виявлення підозрілих транзакцій.

Для наочного представлення результатів оцінювання використовувалась матриця помилок. Вона дозволила проаналізувати кількість:

- правильно класифікованих нормальних транзакцій;
- правильно виявлених підозрілих операцій;
- помилково визначених ризикових транзакцій;
- невиявлених шахрайських операцій.

Аналіз матриці помилок показав, що більшість транзакцій була класифікована коректно. Основна кількість помилок виникала у випадках транзакцій із поведінкою, близькою до межі між нормальними та підозрілими операціями.

Окремо проводилась оцінка ефективності алгоритму IsolationForest. Було встановлено, що алгоритм добре визначає аномальні транзакції навіть без наявності розмічених даних. Найкращі результати модель показала під час виявлення:

- нетипово великих переказів;
- високочастотної активності адрес;
- різких змін поведінки користувачів;
- транзакцій із великою кількістю взаємодій між адресами.

Разом із тим було виявлено, що використання лише аномального аналізу може призводити до збільшення кількості хибнопозитивних результатів. Саме тому у системі використовується комбінований підхід, який поєднує IsolationForest та RandomForestClassifier.

У процесі оцінювання також проводилась перевірка впливу різних параметрів моделей на якість класифікації. Зокрема, аналізувались:

- кількість дерев у Random Forest;
- глибина дерев рішень;
- параметр contamination для Isolation Forest;
- порогове значення risk_score.

Оптимізація цих параметрів дозволила підвищити точність роботи системи та зменшити кількість помилкових спрацьовувань.

Результати оцінювання підтвердили ефективність використання методів машинного навчання для аналізу криптовалютних транзакцій. Розроблена система забезпечує достатньо високу точність виявлення підозрілих фінансових операцій та може бути використана як основа для побудови систем фінансового моніторингу у криптовалютному середовищі.

Таким чином, проведене оцінювання точності показало, що реалізована система здатна ефективно аналізувати криптовалютні транзакції, виявляти

аномальну активність та класифікувати підозрілі операції з високим рівнем точності. Отримані результати підтверджують доцільність використання алгоритмів машинного навчання та графового аналізу для задач виявлення фінансових злочинів у blockchain-мережах.

Висновки до третього розділу

У третьому розділі дипломної роботи було реалізовано практичну частину інтелектуальної системи аналізу криптовалютних транзакцій для виявлення фінансових злочинів. У ході роботи розроблено програмний прототип системи із використанням мови програмування Python та сучасних бібліотек для обробки даних, машинного навчання, графового аналізу та побудови REST API.

У процесі реалізації було сформовано модульну архітектуру програмного забезпечення, що забезпечує розділення функціональних компонентів системи. Реалізовано модулі завантаження та попередньої обробки транзакційних даних, формування ознак, аналізу аномальної активності, класифікації підозрілих операцій, графового аналізу та збереження результатів у базі даних **у Додатку А**.

Для виявлення підозрілих транзакцій використано алгоритми IsolationForest та RandomForestClassifier, реалізовані засобами бібліотеки Scikit-learn. У результаті навчання моделей було забезпечено можливість автоматичного аналізу криптовалютних транзакцій та визначення ризикових фінансових операцій на основі сформованих поведінкових ознак.

У межах практичної реалізації також було створено REST API на базі FastAPI, що забезпечує взаємодію користувача із системою, завантаження CSV-файлів, запуск аналізу та отримання результатів у форматі JSON. Для зберігання інформації використано базу даних SQLite, яка забезпечує збереження результатів аналізу та інформації про ризикові транзакції.

У процесі тестування було підтверджено коректність роботи всіх основних компонентів системи. Проведені перевірки показали, що система стабільно виконує повний цикл аналізу транзакцій, коректно обробляє великі набори даних та успішно виявляє аномальну активність у криптовалютному середовищі.

Оцінювання точності моделей показало достатньо високі результати класифікації. Використання метрик accuracy, precision, recall та F1-score дозволило підтвердити ефективність реалізованих алгоритмів машинного навчання під час виявлення підозрілих фінансових операцій. Найбільш важливим результатом стало забезпечення високого рівня recall, що є критично важливим для задач фінансового моніторингу та виявлення шахрайських транзакцій.

ВИСНОВКИ

У результаті виконання дипломної роботи було досліджено особливості аналізу криптовалютних транзакцій та розроблено інтелектуальну систему для виявлення фінансових злочинів у blockchain-середовищі. Актуальність теми обумовлена стрімким розвитком криптовалютних технологій, збільшенням кількості фінансових операцій у децентралізованих мережах та зростанням рівня кіберзлочинності, пов’язаної з використанням цифрових активів.

У теоретичній частині роботи було розглянуто поняття криптовалют і blockchain-технологій, досліджено основні типи фінансових злочинів у криптовалютному середовищі, а також проаналізовано сучасні методи аналізу транзакцій. Особливу увагу приділено існуючим системам моніторингу криптовалютних операцій, зокрема платформам Chainalysis та Elliptic, які використовують методи графового аналізу та машинного навчання для виявлення незаконної фінансової активності.

У межах аналітичної частини роботи було проведено дослідження наборів даних криптовалютних транзакцій, визначено основні ознаки підозрілих операцій та обґрунтовано вибір алгоритмів машинного навчання. Встановлено, що найбільш ефективним підходом для задач фінансового моніторингу є поєднання методів класифікації та алгоритмів виявлення аномалій. Також було розроблено архітектуру системи, яка включає модулі збору даних, попередньої обробки, машинного навчання, графового аналізу, візуалізації та збереження результатів.

У практичній частині роботи реалізовано програмний прототип інтелектуальної системи аналізу криптовалютних транзакцій із використанням мови програмування Python. Для побудови системи використано бібліотеки Scikit-learn, NetworkX, FastAPI та SQLAlchemy.

У процесі реалізації було створено систему, яка забезпечує:

- завантаження та обробку транзакційних даних;
- формування поведінкових ознак;
- виявлення аномальної активності;
- класифікацію підозрілих транзакцій;
- побудову графів взаємодії між криптовалютними адресами;
- збереження результатів аналізу у базі даних;
- взаємодію через REST API.

Для виявлення ризикових фінансових операцій використано алгоритми IsolationForest та RandomForestClassifier. Проведене тестування підтвердило стабільну роботу системи та коректність функціонування її компонентів. Оцінювання точності моделей показало високі значення показників accuracy, precision, recall та F1-score, що свідчить про ефективність використаних методів машинного навчання.

Результати дослідження підтвердили, що застосування інтелектуальних методів аналізу даних дозволяє автоматизувати процес фінансового моніторингу криптовалютних транзакцій та значно підвищити ефективність виявлення підозрілих операцій. Використання графового аналізу та алгоритмів машинного навчання забезпечує можливість виявлення прихованих закономірностей у blockchain-мережах та визначення потенційно ризикової активності.

Практичне значення роботи полягає у створенні програмного прототипу системи, який може бути використаний як основа для подальшого розвитку систем фінансового моніторингу, інтеграції з реальними blockchain API та побудови масштабованих платформ аналізу криптовалютних транзакцій у режимі реального часу.

Поставлену мету дипломної роботи досягнуто, а всі основні завдання виконано. Розроблена інтелектуальна система демонструє ефективність використання методів машинного навчання та графового аналізу для задач виявлення фінансових злочинів у криптовалютному середовищі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Нормативно-правові акти:

1. Про запобігання та протидію легалізації (відмиванню) доходів, одержаних злочинним шляхом, фінансуванню тероризму та фінансуванню розповсюдження зброї масового знищення : Закон України від 06.12.2019 р. № 361-IX. URL: https://zakon.rada.gov.ua/laws/show/361-20 (дата звернення: 12.05.2026).
2. Про віртуальні активи : Закон України від 17.02.2022 р. № 2074-IX. URL: https://zakon.rada.gov.ua/laws/show/2074-20 (дата звернення: 12.05.2026).

Матеріали практики:

3. Звіт про результати аналізу ризиків використання віртуальних активів у сфері відмивання коштів та фінансування тероризму. Державна служба фінансового моніторингу України, 2024. URL: https://fiu.gov.ua/ (дата звернення: 11.05.2026).

Спеціальна література:

4. Аналіз технологій блокчейн та перспектив їх використання в правоохоронній діяльності / В. В. Баранов, П. В. Смирнов, А. М. Коваль, О. В. Кравченко. *Інформація і право*. 2021.
5. Ткач Ю. М., Смирнов О. А. Застосування методів машинного навчання для виявлення аномальних транзакцій у блокчейн-мережах. *Сучасні інформаційні технології та системи*. 2023.
6. Шевченко А. О. Інтелектуальний аналіз даних (Data Mining) у задачах фінансового моніторингу віртуальних активів. *Вісник Національного технічного університету «ХПІ»*. Серія: Системний аналіз, управління та інформаційні технології. 2022.
7. Akcora C. G., Gel Y. R., Kantarcioglu M. Blockchain Data Analytics. *IEEE Intelligent Systems*. DOI: https://doi.org/10.1109/MIS.2020.2988185.

8. Anti-money laundering and countering the financing of terrorism (AML/CFT) : international standards. Financial Action Task Force (FATF) URL: https://www.fatf-gafi.org/en/publications/Fatfrecommendations/Fatf-recommendations.html
9. Bitcoin: A Peer-to-Peer Electronic Cash System / S. Nakamoto. 2008. URL: https://bitcoin.org/bitcoin.pdf
10. Crypto Crime Report 2025. Chainalysis, 2025. URL: https://www.chainalysis.com/reports/2025-crypto-crime-report/
11. Lorenz J., Silva M. I., Antunes N. Machine Learning for Cryptocurrency Fraud Detection: A Graph-Based Approach.
12. Ostapowicz M., Żbikowski A. Detecting fraudulent accounts on blockchain: A supervised learning approach. *Expert Systems with Applications*. 2021. DOI: https://doi.org/10.1016/j.eswa.2021.115533.
13. Forbes URL: https://forbes.ua/news/nyt-stverdzhue-shcho-znayshlo-spravzhnogo-satoshi-nakamoto-khto-vin-08042026-37831

Веб-ресурси

14. Support Vector Machine (SVM) Algorithm For Machine Learning URL: https://ashish-mehta.medium.com/support-vector-machine-svm-algorithm-for-machine-learning-350fe9139a52
15. Isolation Forest For Anomaly Detection Made Easy & How To Tutorial URL: https://spotintelligence.com/2024/05/21/isolation-forest/
16. Graph Neural Networks - An overview URL: https://theaisummer.com/Graph_Neural_Networks/
17. Алгоритм k-means URL: https://habr.com/companies/skillfactory/articles/877684/
18. Chainalysis URL: https://www.chainalysis.com/
19. Elliptic URL: https://www.elliptic.co/
20. ClipherTrace URL: https://ciphertracers.com/
21. Crystal Blockchain URL: https://inatba.org/crystal-blockchain/
22. Random Forest URL: https://medium.com/@abhishekjainindore24/everything-about-random-forest-90c106d63989

ДОДАТКИ

Додаток А

Посилання на розроблену інтелектуальну систему аналізу криптовалютних транзакцій для виявлення фінансових злочинів.

URL: https://drive.google.com/drive/folders/1FZXXN-ztHdrdhXP6pDMuzVmWHKcNhZgQ?usp=drive_link