Хмельницький національний університет
Факультет інформаційних технологій
Кафедра комп'ютерних наук

КВАЛІФІКАЦІЙНА РОБОТА

на тему *Метод виявлення шахрайських банківських операцій з використанням машинного навчання*

Рівень вищої освіти *другий (магістерський)*

Галузь знань *12 – Інформаційні технології*
Шифр і найменування

Спеціальність *122 – Комп'ютерні науки*
Код і найменування

Освітня програма *Комп'ютерні науки*
Назва

Виконав *студент 2 курсу, група КНм-24-1* — *Владислав ІЛЬЧИШИН*
Курс, група виконавця — Підпис — Ім'я, ПРІЗВИЩЕ

Керівник *д.т.н., професор кафедри КН* — *Едуард МАНЗЮК*
Науковий ступінь, посада — Підпис — Ім'я, ПРІЗВИЩЕ

Нормоконтроль *к.т.н., доцент кафедри КН* — *Руслан БАГРІЙ*
Науковий ступінь, посада — Підпис — Ім'я, ПРІЗВИЩЕ

До захисту допускаю

Зав. кафедри КН, д.т.н., професор — *Олександр БАРМАК*
Підпис — Ім'я, ПРІЗВИЩЕ

15 грудня 2025 р.

Хмельницький 2025

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
Факультет *інформаційних технологій*
Кафедра *комп'ютерних наук*
Освітній ступінь *магістр*
Галузь знань *12 – Інформаційні технології*
Спеціальність *122 – Комп'ютерні науки*

ЗАТВЕРДЖУЮ
Завідувач кафедри *комп'ютерних наук*

_____
(підпис)
д.т.н., професор Олександр БАРМАК
«*28*» *08* *2025* року

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

1. Тема кваліфікаційної роботи *«Метод виявлення шахрайських банківських операцій з використанням машинного навчання»*

2. Завдання видано студенту *Владиславу ІЛЬЧИШИНУ*
(Ім'я, ПРІЗВИЩЕ)

3. Керівник роботи *професор кафедри КН Едуард МАНЗЮК*
(Ім'я, ПРІЗВИЩЕ)

4. Затверджені наказом університету від «*25*» *08* *2025* р. № *65*

5. Дата видачі завдання студенту «*28*» *08* *2025* р.

6. Зміст пояснювальної записки (перелік задач) та вихідні дані

Робота присвячена підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів. Завдання здійснити комплексний аналіз наявних підходів до виявлення фінансового шахрайства з застосуванням інтелектуальних технологій; розробити метод класифікації банківських транзакцій з використанням ансамблевого алгоритму випадкового лісу; розробити програмну систему для обробки транзакційних даних у реальному часі; виконати практичне тестування розробленого методу на реальних банківських даних.

Ключові терміни фінансове шахрайство, класифікація транзакцій, машинне навчання, випадковий ліс, незбалансовані дані, автоматизація банківської безпеки.

7. Календарний план виконання кваліфікаційної роботи

№	Назва етапів (розділів) кваліфікаційної роботи магістра	Термін виконання	Примітка
1	*Вибір напряму дослідження та узгодження теми кваліфікаційної роботи з керівником, складання календарного графіка виконання роботи*	*вересень 2025*	*Виконано*
2	*Ознайомлення з предметною областю, аналіз існуючих методів і моделей, формулювання мети та завдань дослідження, визначення об'єкта й предмета дослідження*	*вересень 2025*	*Виконано*
3	*Розробка методу чи моделі для вирішення обраного завдання, опис архітектури рішення*	*жовтень 2025*	*Виконано*
4	*Програмна реалізація методу чи моделі*	*жовтень 2025*	*Виконано*
5	*Дослідження ефективності та експериментальна перевірка результатів, порівняння з відомими підходами*	*листопад 2025*	*Виконано*
6	*Написання пояснювальної записки, оформлення відповідно до вимог, врахування зауважень керівника*	*листопад 2025*	*Виконано*
7	*Підготовка презентаційних матеріалів та попередній захист*	*листопад 2025*	*Виконано*
8	*Перевірка пояснювальної записки на відповідність вимогам оформлення (нормоконтроль) та перевірка на академічну доброчесність. Отримання відгуку керівника та рецензії.*	*грудень 2025*	*Виконано*
9	*Публічний захист кваліфікаційної роботи*	*грудень 2025*	*Виконано*

Виконавець *студент групи КНм-24-1* *Владислав ІЛЬЧИШИН*
Група виконавця Підпис Ім'я, ПРІЗВИЩЕ

Керівник *д.т.н., проф. каф. КН* *Едуард МАНЗЮК*
Науковий ступінь, посада Підпис Ім'я, ПРІЗВИЩЕ

Реферат

Кваліфікаційна робота присвячена розробці методу автоматизованого виявлення шахрайських банківських транзакцій з використанням технологій машинного навчання.

Актуальність теми. Актуальність дослідження визначається критичною потребою фінансових установ у надійних засобах виявлення шахрайських операцій в умовах зростаючих обсягів цифрових транзакцій. Традиційні підходи до моніторингу, що базуються на ручному аналізі та простих правилах, виявляються неефективними через високу трудомісткість, значну кількість помилкових спрацювань і обмежену здатність виявляти нові схеми шахрайства.

Сучасні досягнення в області машинного навчання відкривають можливості для створення адаптивних систем моніторингу, здатних автоматично виявляти складні патерни підозрілої поведінки. Застосування ансамблевих методів та технік обробки незбалансованих даних дозволяє суттєво підвищити якість детектування шахрайських операцій при одночасному зменшенні навантаження на служби безпеки.

Додатково, впровадження методів синтетичної генерації даних допомагає вирішити проблему дефіциту розмічених прикладів шахрайських транзакцій, що є типовою складністю у фінансовій сфері. Це підкреслює актуальність роботи як для практичного застосування у банківській індустрії, так і для наукових досліджень у галузі застосування алгоритмів машинного навчання до задач фінансової безпеки.

Мета роботи. Мета роботи полягає в підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів.

Задачі дослідження

– провести аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розробити метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створити програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

– провести експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

Об'єкт дослідження – процес виявлення та класифікації шахрайських операцій у банківських транзакційних системах.

Предмет дослідження – моделі, методи та технології виявлення фінансового шахрайства на основі машинного навчання з застосуванням балансування незбалансованих даних.

Методи дослідження. Застосовано ансамблеві методи машинного навчання, техніки синтетичної генерації даних, методи нормалізації та трансформації ознак, статистичний аналіз, експериментальне тестування на реальних транзакційних даних.

Наукова новизна одержаних результатів. Удосконалено метод виявлення шахрайських банківських операцій, який відрізняється від існуючих інтегрованим застосуванням алгоритму випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу та використанням комплексної системи генерації міток на основі множини критеріїв підозрілості, що дозволило підвищити повноту виявлення шахрайських транзакцій при збереженні високої точності класифікації.

Апробація результатів кваліфікаційної роботи та публікації.

Ільчишин В.В., Манзюк Е.А., Скрипник Т.К. Метод виявлення шахрайських банківських операцій з використанням машинного навчання. Збірник наукових праць за матеріалами XVII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп’ютерних наук АПКН-2025». - Хмельницький, 2025. - С.145 – 151.

Структура та обсяг роботи. Робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг основного тексту становить 76 сторінок, включаючи 19 рисунків, 2 таблиці та 45 джерел у списку літератури.

Ключові слова банківське шахрайство, виявлення аномалій, випадковий ліс, SMOTE, балансування класів, транзакційний моніторинг, класифікація, автоматизована діагностика, фінансова безпека.

Зміст

Перелік скорочень

Скорочення	**Пояснення**
МН	Машинне навчання
ШІ	Штучний інтелект
SMOTE	Synthetic Minority Over-sampling Technique (Техніка синтетичного перевибіркування меншості)
RF	Random Forest (Випадковий ліс)
CSV	Comma-Separated Values (Значення, розділені комами)
API	Application Programming Interface (Інтерфейс програмування додатків)
UML	Unified Modeling Language (Уніфікована мова моделювання)

Вступ

Кваліфікаційна робота присвячена проектуванню методу автоматизованого виявлення шахрайських банківських транзакцій з використанням технологій машинного навчання.

Актуальність теми. Актуальність дослідження визначається критичною потребою фінансових установ у надійних засобах виявлення шахрайських операцій в умовах зростаючих обсягів цифрових транзакцій. Традиційні підходи до моніторингу, що базуються на ручному аналізі та простих правилах, виявляються неефективними через високу трудомісткість, значну кількість помилкових спрацювань і обмежену здатність виявляти нові схеми шахрайства.

Сучасні досягнення в області машинного навчання відкривають можливості для створення адаптивних систем моніторингу, здатних автоматично виявляти складні патерни підозрілої поведінки. Застосування ансамблевих методів та технік обробки незбалансованих даних дозволяє суттєво підвищити якість детектування шахрайських операцій при одночасному зменшенні навантаження на служби безпеки.

Додатково, впровадження методів синтетичної генерації даних допомагає вирішити проблему дефіциту розмічених прикладів шахрайських транзакцій, що є типовою складністю у фінансовій сфері. Це підкреслює актуальність роботи як для практичного застосування у банківській індустрії, так і для наукових досліджень у галузі застосування алгоритмів машинного навчання до задач фінансової безпеки.

Мета роботи. Мета роботи полягає в підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів.

Задачі дослідження

– провести аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розробити метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створити програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

– провести експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

Об'єкт дослідження – процес виявлення та класифікації шахрайських операцій у банківських транзакційних системах.

Предмет дослідження – моделі, методи та технології виявлення фінансового шахрайства на основі машинного навчання з застосуванням балансування незбалансованих даних.

Методи дослідження. Застосовано ансамблеві методи машинного навчання, техніки синтетичної генерації даних, методи нормалізації та трансформації ознак, статистичний аналіз, експериментальне тестування на реальних транзакційних даних.

Наукова новизна одержаних результатів. Удосконалено метод виявлення шахрайських банківських операцій, який відрізняється від існуючих інтегрованим застосуванням алгоритму випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу та використанням комплексної системи генерації міток на основі множини критеріїв підозрілості, що дозволило підвищити повноту виявлення шахрайських транзакцій при збереженні високої точності класифікації.

Апробація результатів кваліфікаційної роботи та публікації. Ільчишин В.В., Манзюк Е.А., Скрипник Т.К. Метод виявлення шахрайських банківських операцій з використанням машинного навчання. Збірник наукових праць за матеріалами XVII Всеукраїнської науково-практичної конференції «Актуальні проблеми комп’ютерних наук АПКН-2025». - Хмельницький, 2025. - С.145 – 151.

Структура та обсяг роботи. Робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг основного тексту становить 75 сторінок, включаючи 19 рисунків, 2 таблиці та 45 джерел у списку літератури.

Розділ 1 Аналіз методів виявлення шахрайських банківських операцій

Сьогодні в умовах розвитку фінансової сфери виявлення шахрайських банківських операцій відіграє важливе значення через зростання обсягів цифрових транзакцій та еволюцію методів зловмисників. Щороку фінансові установи зазнають мільярдних збитків через шахрайські схеми, що постійно вдосконалюються та адаптуються до нових захисних механізмів. Шахрайство в банківському секторі охоплює широке коло незаконних операцій, в тому числі підробку кредитних карток, несанкціоновані перекази, крадіжку особистих даних, маніпуляції з онлайн-платежами та складні схеми відмивання коштів. Традиційні методи виявлення, що базуються на правилах та ручному моніторингу, виявляються недостатньо ефективними в умовах масштабування цифрових платежів.

1.1 Характеристика задачі виявлення шахрайських банківських операцій

Сьогодні в умовах трансформації фінансового сектору виявлення шахрайських банківських операцій стало одним із найважливіших завдань забезпечення стабільності та безпеки фінансових установ. Зростання обсягів електронних платежів, розвиток онлайн-банкінгу та мобільних платіжних систем створили нові можливості для зловмисників, які постійно вдосконалюють свої методи обходу систем захисту. Шахрайство в банківській сфері охоплює широке коло нелегальних дій, включаючи фальсифікацію кредитних карток, несанкціоновані грошові перекази, маніпуляції з онлайн-платежами та крадіжки персональних даних клієнтів [1].

Масштаби проблеми шахрайства у сфері банків значні та продовжують зростати. Банківські установи та їхні клієнти щорічно зазнають мільярдних збитків через шахрайські операції, що негативно впливає на довіру до фінансової системи загалом. Традиційні методи виявлення фроду, які базувалися на правилах та експертних системах, показали свою обмеженість у протидії з постійно

еволюціонуючими схемами шахрайства. Зловмисники швидко адаптуються до нових захисних механізмів, знаходячи способи їх обходу, що вимагає більш гнучких та інтелектуальних підходів до виявлення аномалій.

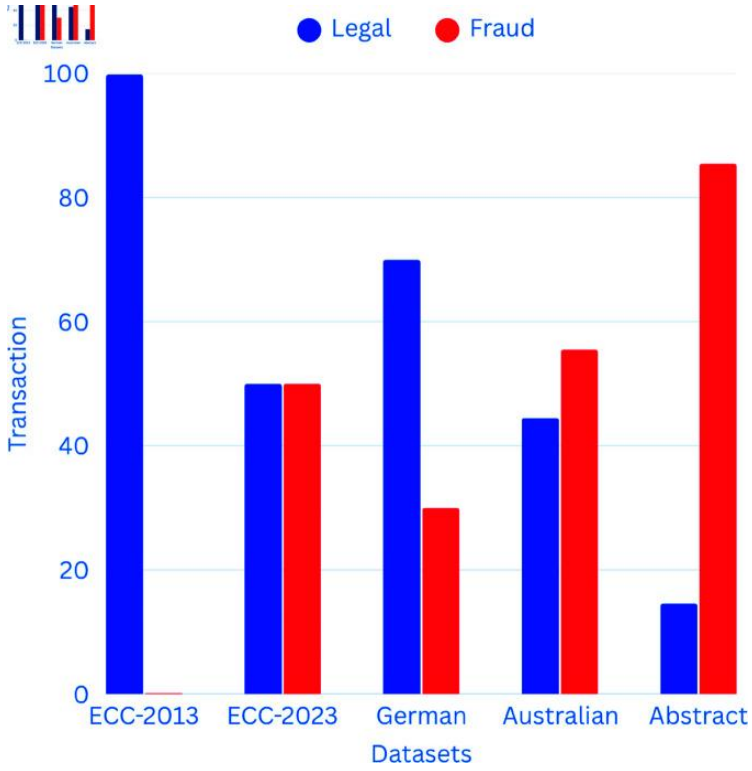


Рисунок 1.1 – Обсяг різних наборів даних [1]

Задача детектування операцій шахраїв може бути формалізована як задача бінарної класифікації, де кожна транзакція має бути віднесена до одного з двох класів легальна операція або шахрайська операція [2]. Вхідними даними для такої системи є множина ознак транзакції, що включають суму платежу, час здійснення операції, географічне розташування точки транзакції, тип рахунку, історію попередніх операцій клієнта та різноманітні поведінкові патерни. Важливою особливістю цих даних є їх анонімізація для забезпечення конфіденційності клієнтів, що часто досягається застосуванням методів зниження розмірності або шифрування чутливої інформації.

Ключовою характеристикою задачі виявлення фроду є значний дисбаланс класів у наборах даних. Частка шахрайських транзакцій зазвичай становить менше відсотка від всієї кількості операцій, що дає серйозні труднощі для побудови

ефективних моделей класифікації [3, 4]. Така незбалансованість призводить до того, що класичні алгоритми машинного навчання схильні класифікувати більшість транзакцій як легальні, досягаючи високої загальної точності, але демонструючи низьку повноту виявлення шахрайських операцій. Необхідність балансування між точністю та повнотою моделі є критично важливою, оскільки помилки обох типів мають суттєві наслідки для банківської установи.

Фальшиві позитивні спрацювання системи виявлення фроду призводять до блокування легальних операцій клієнтів, що викликає незадоволення користувачів та може бути причиною до втрати клієнтів банком. Водночас фальшиві негативні результати означають пропуск справжніх шахрайських операцій, що безпосередньо призводить до фінансових втрат [5]. Оптимальна модель виявлення фроду повинна знаходити баланс між цими двома типами помилок, враховуючи різну вартість помилкових класифікацій для конкретної банківської установи.

У банківському контексті задача значно ускладнюється гетерогенністю джерел даних та різноманітністю типів транзакцій. Операції з кредитними картками, мобільні платежі, міжбанківські перекази, онлайн-покупки та готівкові транзакції мають різні характеристики та потребують різних підходів до виявлення аномалій. Шахрайство може проявлятися як відхилення від нормальних патернів поведінки окремого клієнта або як підозрілі зв'язки в графі транзакцій, що вимагає застосування різних аналітичних методів [6].

Важливим аспектом задачі є необхідність врахування часових залежностей між транзакціями. Послідовність операцій клієнта формує певний патерн поведінки, відхилення від якого може свідчити про шахрайську діяльність [7–9]. Наприклад, серія швидких транзакцій у різних географічних локаціях або незвичні комбінації типів операцій можуть бути індикаторами компрометації облікового запису. Ефективна модель виявлення фроду повинна аналізувати не лише окремі транзакції, але й їх послідовності та взаємозв'язки.

Регуляторні вимоги додають додаткові обмеження до розробки систем виявлення шахрайства. Законодавство про захист персональних даних та фінансове регулювання встановлюють жорсткі правила щодо обробки інформації клієнтів та

прийняття рішень про блокування транзакцій [10]. Моделі машинного навчання повинні бути не лише точними, але й інтерпретованими, щоб банківські працівники могли зрозуміти причини класифікації транзакції як підозрілої та обґрунтувати свої дії перед регуляторами.

Метрики оцінювання якості моделей виявлення фроду відрізняються від стандартних метрик класифікації. Через значну незбалансованість класів простої точності недостатньо для адекватної оцінки ефективності системи. Натомість використовуються такі метрики як точність, повнота та гармонійне середнє цих показників [11, 12]. Крива точності-повноти та площа під нею є важливими інструментами для порівняння різних моделей та вибору оптимального порогу класифікації.

Параметри системи виявлення фроду включають часові вікна для агрегації даних, які визначають, яка історія транзакцій враховується при аналізі поточної операції. Типові значення таких вікон становлять від кількох годин до кількох діб, залежно від типу транзакцій та специфіки бізнесу банку. Пороги ймовірності класифікації є ще одним критичним параметром, який визначає баланс між чутливістю системи та частотою помилкових спрацювань.

Динамічність шахрайських схем створює додаткові виклики для розробників систем виявлення. Зловмисники постійно вивчають механізми захисту та знаходять нові способи їх обходу, що призводить до явища дрейфу концепцій [13]. Моделі, навчені на історичних даних, з часом втрачають свою ефективність і потребують регулярного оновлення. Це вимагає впровадження механізмів онлайн-навчання та інкрементального оновлення моделей, які можуть адаптуватися до нових патернів шахрайства без повного перенавчання.

Практичні реалізації систем виявлення фроду мають справу з масивними обсягами даних, які потребують обробки в реальному часі. Сучасні банківські системи обробляють мільйони транзакцій щодня, і кожна з них повинна бути проаналізована протягом мілісекунд для прийняття рішення про дозвіл або блокування операції [14]. Це накладає жорсткі вимоги на обчислювальну

ефективність моделей та їх здатність масштабуватися для обробки великих потоків даних.

Публічно доступні набори даних для дослідження методів виявлення фроду часто є синтетичними або сильно анонімізованими через конфіденційність банківської інформації. Набори даних зазвичай містять сотні тисяч транзакцій з невеликою часткою шахрайських операцій, що відображає реальний розподіл класів [15–17]. Ознаки в таких датасетах часто перетворені методами зниження розмірності для забезпечення анонімності, що ускладнює інтерпретацію результатів та вимагає особливої уваги до вибору та налаштування моделей.

Обробка пропущених значень та викидів є важливим етапом підготовки даних для навчання моделей. У реальних банківських логах частка пропущених значень може сягати кількох відсотків через технічні проблеми або неповноту інформації про транзакції. Викиди можуть бути як справжніми аномаліями, що свідчать про шахрайство, так і помилками в даних, що вимагає ретельного аналізу та вибору стратегії їх обробки.

Розширення задачі до мультикласової класифікації дозволяє не лише виявляти шахрайські операції, але й визначати їх тип. Різні види фроду, такі як крадіжка картки, перехоплення облікового запису, фішинг або створення синтетичних ідентичностей, мають різні характеристики та вимагають різних контрзаходів. Мультикласова класифікація додає складності задачі, але надає цінну додаткову інформацію для банківських аналітиків.

Інтеграція систем виявлення фроду з існуючою банківською інфраструктурою є критично важливим аспектом практичного впровадження. Моделі повинні бути сумісні з системами керування подіями безпеки, забезпечувати інтерфейси для взаємодії з іншими компонентами банківської системи та підтримувати необхідні протоколи обміну даними. Загалом, характеристика задачі виявлення шахрайських банківських операцій підкреслює її багатогранність та складність, що вимагає комплексного підходу для успішного вирішення.

1.2 Аналіз існуючих публікацій та наукових підходів

Наукові дослідження в галузі виявлення шахрайських банківських операцій демонструють значну еволюцію підходів від класичних статистичних методів до сучасних архітектур глибокого навчання. Огляд літератури показує, що дослідники шукають підходи покращення ефективності виявлення фроду, боротьби з незбалансованістю даних та покращення інтерпретованості моделей [18, 19].

Ранні роботи в цій галузі фокусувалися на застосуванні класичних підходів машинного навчання до проблем детектування нетипових дій у банківських транзакціях. Дослідники активно експериментували з різними підходами, намагаючись знайти оптимальний баланс між точністю виявлення шахрайства та кількістю помилкових спрацювань. Емпіричні дослідження проводилися переважно на публічно доступних наборах даних, як порівнювати результати різних методів [20].

Випадковий ліс показав себе як один із найбільш стабільних та ефективних алгоритмів для задачі класифікації транзакцій. Його здатність оперувати з великим масивом ознак, стійкість до перенавчання та можливість оцінки важливості ознак зробили його популярним вибором серед дослідників [21]. Публікації демонстрували, що правильне налаштування параметрів моделі дозволяє досягти високих показників точності та повноти.

Методи градієнтного бустингу привернули значну увагу дослідників завдяки своїй здатності послідовно покращувати якість класифікації шляхом навчання ансамблю слабких класифікаторів. Різні варіанти цього підходу, кожен з яких має свої особливості оптимізації та обробки даних, активно досліджувалися та порівнювалися між собою [22, 23]. Результати показали, що ці методи можуть ефективно працювати з розрідженими даними та великою кількістю ознак.

Важливим напрямком досліджень стало вирішення проблеми незбалансованості класів у наборах даних. Дослідники пропонували різні підходи до балансування даних, включаючи методи передискретизації меншого класу, зменшення вибірки більшого класу та синтетичну генерацію нових зразків [24].

Публікації показували, що правильне застосування цих методів може суттєво покращити здатність моделі виявляти шахрайські операції.

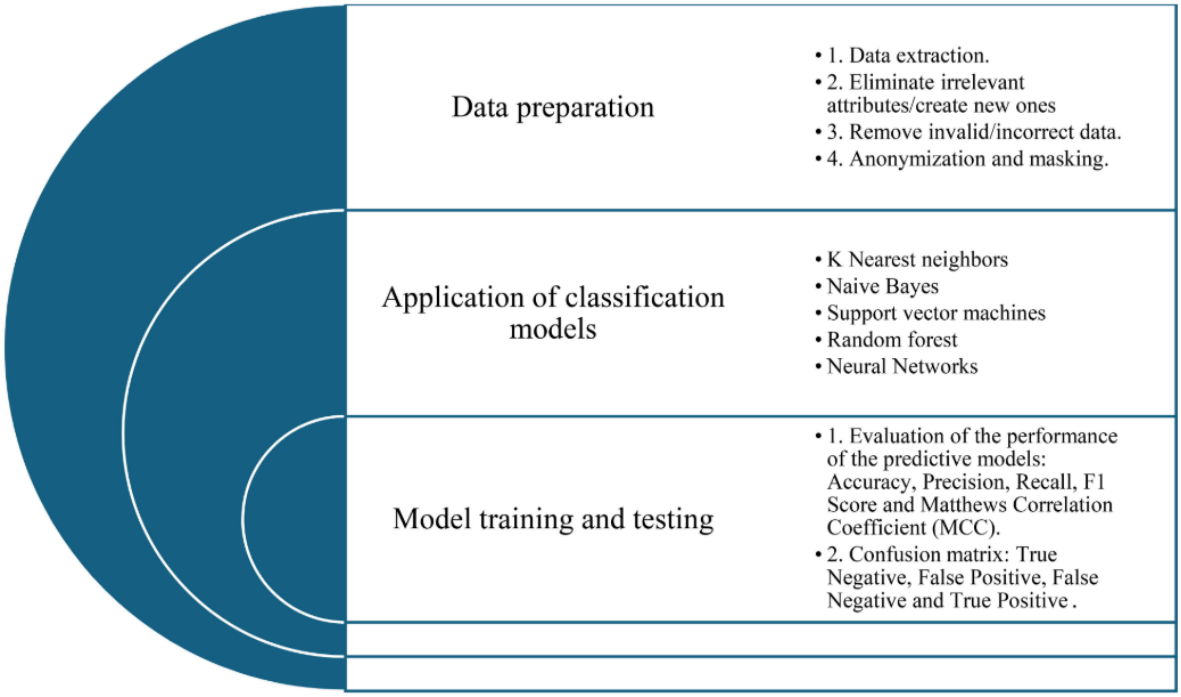


Рисунок 1.2 – Методологічний робочий процес виявлення підозрілих банківських переказів [21]

Графові нейронні мережі відкрили новий напрямок у виявленні фроду, дозволяючи моделювати складні взаємозв'язки між транзакціями та рахунками. Дослідники показали, що шахрайські операції часто утворюють специфічні патерни в графі транзакцій, які можна виявити за допомогою спеціалізованих архітектур [25, 26]. Цей підхід особливо ефективний для виявлення організованого шахрайства та мереж зловмисників.

Рекурентні нейронні мережі та їх різновиди знайшли застосування для аналізу послідовностей транзакцій. Здатність цих архітектур запам'ятовувати довгострокові залежності дозволяє виявляти аномалії в поведінці клієнтів на основі історії їх операцій [27]. Дослідження показали, що врахування часового контексту значно покращує якість виявлення шахрайства порівняно з аналізом окремих транзакцій.

Автоенкодери як методи навчання без учителя привернули увагу завдяки здатності детектувати аномалії шляхом реконструкції вхідних даних. Ідея полягає в тому, що навчена на легальних транзакціях модель, не зможе добре реконструювати шахрайські операції, що дозволяє виявити їх за високою помилкою реконструкції [28, 29]. Цей підхід добре працює у ситуаціях, коли розмічених даних про шахрайство недостатньо.

Системні огляди літератури проаналізували сотні наукових публікацій, систематизуючи підходи до виявлення фроду та виявляючи основні тенденції розвитку галузі. Такі огляди показали, що комбінація різних методів часто дає кращі результати, ніж використання окремих алгоритмів [30]. Дослідники також відзначили важливість правильної підготовки даних та вибору релевантних ознак для успішного виявлення шахрайства.

Ансамблеві методи, які комбінують передбачення кількох базових моделей, показали високу ефективність у задачі виявлення фроду. Стекінг, бегінг та бустинг дозволяють використати переваги різних алгоритмів та компенсувати їх недоліки [31, 32]. Публікації демонстрували, що правильно сконструйовані ансамблі можуть суттєво перевершувати окремі моделі за всіма метриками якості.

Дослідження інтерпретованості моделей машинного навчання набули особливої актуальності в контексті банківського застосування. Регуляторні вимоги та необхідність пояснення рішень клієнтам призвели до розробки методів, які дозволяють зрозуміти, чому модель класифікувала транзакцію як підозрілу [33, 34]. Технології пояснювального штучного інтелекту стали важливою складовою сучасних систем виявлення фроду.

Федеративне навчання запропоновано як рішення для ситуацій, коли кілька банків хочуть співпрацювати у виявленні шахрайства без обміну чутливими даними клієнтів. Цей підхід дозволяє навчити спільну модель на розподілених даних, зберігаючи конфіденційність інформації кожного учасника [35]. Дослідження показали перспективність цього напрямку, хоча існують виклики, пов'язані з комунікаційними витратами та захистом від атак.

Трансформери, які революціонізували обробку природної мови, почали застосовуватися і для аналізу транзакційних даних. Механізм уваги дозволяє моделі фокусуватися на найбільш релевантних аспектах послідовності транзакцій, що покращує якість виявлення складних патернів шахрайства. Публікації демонструють багатообіцяючі результати застосування цих архітектур до фінансових даних.

Аналіз літератури також виявив проблему етичних аспектів застосування машинного навчання у фінансовій сфері. Упередженість у навчальних даних може призвести до дискримінації певних груп клієнтів, що є неприйнятним з етичної та правової точок зору. Дослідники розробляють методи виявлення та усунення упередженості в моделях виявлення фроду.

1.3 Огляд архітектур, методів та моделей машинного навчання для класифікації шахрайських транзакцій

Огляд методів машинного навчання для класифікації шахрайських транзакцій охоплює широкий спектр підходів, від класичних алгоритмів до сучасних архітектур глибокого навчання. Кожен метод має свої переваги, обмеження та сфери оптимального застосування, що робить вибір підходу важливим етапом розробки системи виявлення фроду [36, 37].

Метод опорних векторів відомий, як однин із фундаментальних алгоритмів машинного навчання, який знайшов широке застосування в задачах класифікації. Цей метод шукає оптимальну гіперплощину, що розділяє класи з максимальним відступом, що теоретично забезпечує хорошу узагальнювальну здатність моделі [38]. У контексті виявлення фроду метод векторів демонструє високу точність, особливо при використанні нелінійних ядер, які дозволяють працювати з складними залежностями в даних.

Основною перевагою методу опорних векторів є його ефективність у високорозмірних просторах ознак та стійкість. Функція ядра та її параметри критично впливають на якість класифікації [39, 40]. Радіальне базисне ядро є найбільш популярним вибором для задач виявлення фроду, оскільки воно може

моделювати складні нелінійні залежності між ознаками. Параметр регуляризації контролює баланс між максимізацією відступу та мінімізацією помилок класифікації на навчальній вибірці.

Випадковий ліс представляє собою ансамблевий метод, що створюється на множині дерев рішень на різних підвибірках даних та усереднює їх передбачення. Кожне дерево в ансамблі навчається на випадковій підвибірці ознак, що забезпечує різноманітність моделей та знижує кореляцію між ними [41, 42]. Цей підхід ефективно бореться з перенавчанням та демонструє хороші результати на наборах даних.

Випадковий ліс має кілька важливих переваг для задачі виявлення фроду. Він може працювати з великим масивом ознак без необхідності їх попереднього відбору. Метод надає оцінки важливості ознак, що допомагає зрозуміти, які характеристики транзакцій є найбільш інформативними для виявлення шахрайства [34, 43]. Випадковий ліс відносно стійкий до незбалансованості класів, особливо при використанні зваженої функції втрат.

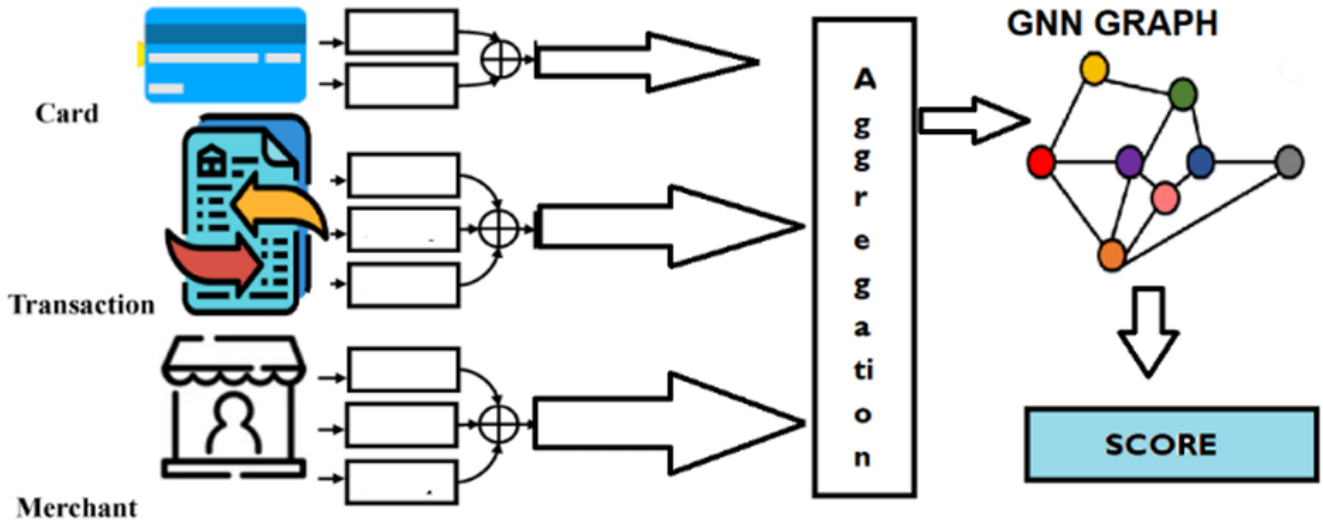


Рисунок 1.3 – Модель архітектури розпізнавання фроду [34]

Градієнтний бустинг дерев є ансамблевим методом, який послідовно створює дерева рішень, кожне з яких намагається виправити помилки попередніх. На відміну від випадкового лісу, де дерева будуються незалежно, у градієнтному бустингу

кожне наступне дерево фокусується на прикладах, які важко класифікувати [44, 45]. Це дозволяє досягти вищої точності, але вимагає обережного налаштування параметрів для уникнення перенавчання.

Рекурентні мережі розроблені для того, що працювати з послідовними даними, що робить їх природним вибором для аналізу часових рядів транзакцій. Ці архітектури підтримують внутрішній стан, який дозволяє запам'ятовувати інформацію з попередніх кроків послідовності. Для виявлення фроду це означає можливість враховувати історію транзакцій клієнта при оцінці поточної операції.

Автоенкодери представляють собою архітектуру нейронних мереж, призначену для навчання ефективних представлень даних в режимі без учителя. Автоенкодер складається з енкодера, який стискає вхідні дані до низькорозмірного представлення, та декодера, який намагається реконструювати оригінальні дані з цього представлення. Для виявлення фроду автоенкодери навчаються на легальних транзакціях, і потім шахрайські операції виявляються за високою помилкою реконструкції.

1.4 Мета та постановка задачі

Мета роботи полягає в підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів.

Задачі дослідження

– провести аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розробити метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створити програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

– провести експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

Розділ 2 Метод виявлення шахрайських банківських операцій та критерії його оцінювання

2.1 Концепція та схема методу виявлення шахрайських операцій

Розроблений метод виявлення шахрайських банківських операцій базується на використанні алгоритмів машинного навчання для автоматичної класифікації транзакцій. Підхід полягає у навчанні моделі на історичних даних про банківські операції, які містять як легальні, так і шахрайські транзакції, з подальшим використанням навченої моделі для класифікації нових операцій у реальному часі.

Концепція методу ґрунтується на припущенні, що шахрайські операції мають специфічні характеристики та патерни поведінки, які відрізняють їх від звичайних легальних транзакцій. Ці відмінності можуть проявлятися у різних аспектах операції, таких як сума транзакції, тип рахунку, час здійснення операції, географічне розташування та інші параметри.

Метод складається з декількох етапів послідовної обробки даних, кожен з яких виконує специфічну функцію у процесі виявлення шахрайства. Перший етап включає завантаження та попередній аналіз вхідних даних про банківські транзакції. На цьому етапі відбувається ознайомлення зі структурою датасету, виявлення основних характеристик даних та визначення потенційних проблем, які можуть вплинути на якість навчання моделі.

Другий етап присвячено попередній обробці даних, що важливо для забезпечення правильної роботи алгоритмів навчання. Цей етап включає обробку пропущених значень, виявлення та усунення викидів, нормалізацію ознак та кодування змінних категорій. Правильна підготовка даних суттєво впливає на здатність моделі виявляти складні залежності між ознаками транзакцій.

Третій етап методу пов'язаний із вирішенням проблеми незбалансованості класів у датасеті. Оскільки підозрілі операції становлять невелику частину від загального масиву транзакцій, це створює проблеми для навчання. Для вирішення цієї проблеми застосовується техніка балансування класів, яка дозволяє моделі краще розпізнавати підозрілі операції.

На четвертому етапі відбувається безпосереднє навчання моделі машинного навчання на підготовлених даних. Модель аналізує зв'язки між ознаками транзакцій та їх класами, формуючи внутрішнє представлення, яке дозволяє відрізняти легальні операції від шахрайських. Процес навчання включає підбір оптимальних параметрів моделі та валідацію її роботи на окремій вибірці даних.

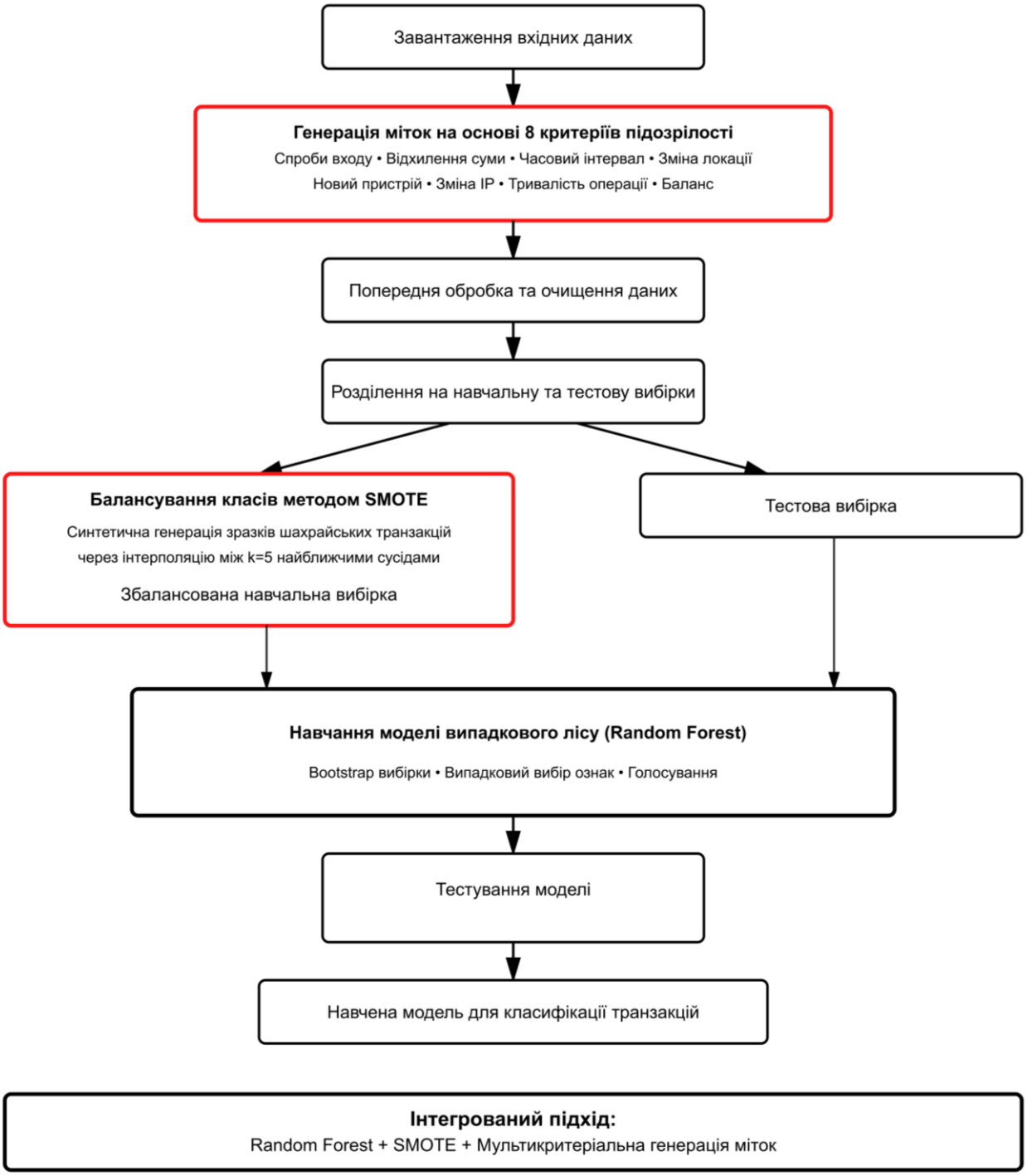


Рисунок 2.1 – Загальна схема методу виявлення шахрайських операцій

П'ятий етап методу полягає у тестуванні навченої моделі на незалежному тестовому наборі даних під час навчання. Це дає змогу оцінити можливість моделі узагальнювати набуті знання на нові, раніше не бачені транзакції. В цьому етапі обчислюються різноманітні метрики якості класифікації.

Завершальний етап методу передбачає використання навченої моделі для класифікації нових банківських операцій у типових умовах. Нова транзакція проходитиме через належні етапи попередньої обробки, після чого модель визначає ймовірність того, що дана операція є шахрайською. На основі обрахованої ймовірності приймається рішення про класифікацію транзакції.

Загальна схема методу представлена на рисунку 2.1, де показано послідовність усіх етапів обробки даних та їх взаємозв'язки. Схема ілюструє не лише лінійну послідовність кроків, але й зворотні зв'язки між етапами, які можуть виникати у процесі налаштування та оптимізації моделі.

Важливою особливістю розробленого методу є його модульність, що дозволяє незалежно змінювати окремі компоненти без необхідності перебудови всієї системи. Наприклад, можна змінити алгоритм машинного навчання або методи обробки даних не попередньому етапі, зберігаючи при цьому загальну структуру методу.

Метод також передбачає можливість адаптації до змін у характері шахрайських операцій шляхом періодичного перенавчання моделі на нових даних. Це забезпечує актуальність системи виявлення фроду та її здатність реагувати на еволюцію методів шахрайства у банківській сфері.

2.2 Архітектура моделі класифікації

Для вирішення задачі класифікації банківських транзакцій було обрано алгоритм випадкового лісу, який є ансамблевим методом машинного навчання. Вибір цього алгоритму обумовлений декількома факторами, які роблять його придатним для задачі виявлення шахрайства у банківських операціях.

Випадковий ліс демонструє високу точність класифікації на різних типах даних та добре справляється з задачами, де присутня велика кількість ознак. Цей алгоритм є стійким до перенавчання завдяки своїй ансамблевій природі. Випадковий ліс може працювати з даними, які містять як числові, так і категоріальні ознаки, що є важливим для банківських транзакцій.

Архітектура випадкового лісу базується на побудові множини дерев рішень, що навчається на підвибірці вхідних даних. Кілька моделей, які приймають рішення незалежно один від одного, можуть давати кращий результат, ніж одна модель. Це відбувається через те, що помилки окремих дерев компенсуються правильними передбаченнями інших дерев.

Процес роботи випадкового лісу починається зі створення множини навчальних підвибірок методом бутстрепу. Цей метод полягає у випадковому відборі зразків з навчального датасету з поверненням, що означає, що один і той самий зразок може потрапити у вибірку декілька разів. Кожна така підвибірка має той самий розмір, що і початковий навчальний набір, але містить різні комбінації зразків.

Для кожної створеної підвибірки будується окреме дерево рішень. При побудові дерева на кожному кроці береться інша підмножина ознак, серед яких здійснюється пошук найкращого розбиття. Кількість ознак, що розглядаються на кожному кроці, є одним із параметрів моделі і зазвичай встановлюється як квадратний корінь із загальної кількості ознак. Це забезпечує різноманітність дерев у ансамблі та знижує кореляцію між ними.

Дерево в ансамблі будується до повної глибини, якщо не встановлено обмежень на максимальну глибину або мінімальну кількість зразків у листі. Повністю вирощені дерева можуть мати низьку помилку на навчальних даних, але саме усереднення передбачень множини таких дерев запобігає перенавчанню ансамблю.

Процес класифікації нової транзакції відбувається шляхом пропускання її через усі дерева в ансамблі. Кожне дерево незалежно від інших приймає рішення про належність транзакції до якогось з класів. Після того сукупність дерев зробили

свої передбачення, відбувається голосування, де підраховується кількість голосів за кожний клас. Клас, який отримав велику кількість голосів, береться як фінальне передбачення моделі.

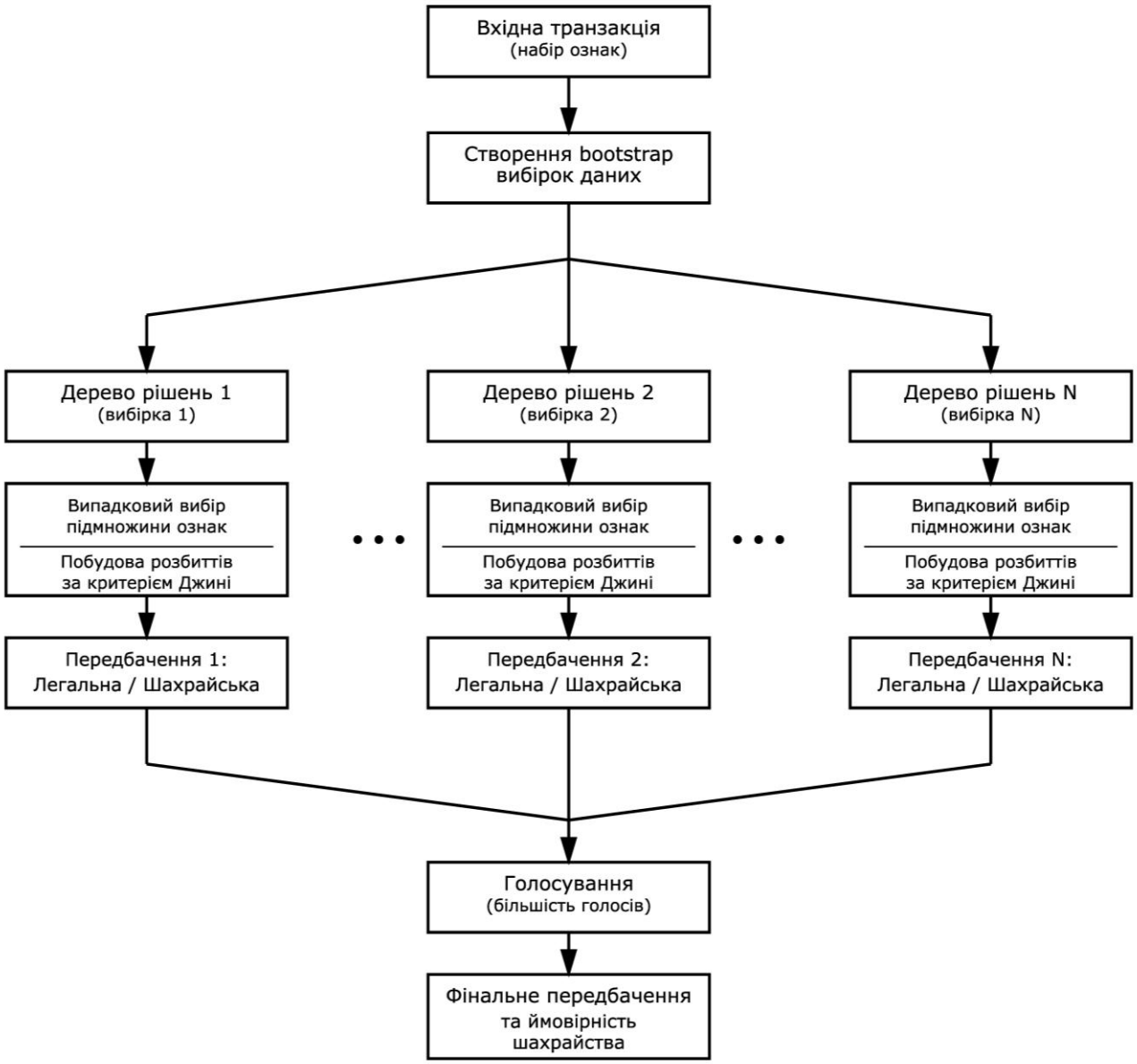


Рисунок 2.2 – Архітектура моделі класифікації транзакцій

У контексті виявлення шахрайства випадковий ліс також може надавати оцінку ймовірності належності транзакції до класу шахрайських операцій. Ця ймовірність обчислюється як частка дерев, які класифікували транзакцію як шахрайську, від загальної сукупності дерев в ансамблі. Ймовірнісна інтерпретація

дозволяє встановлювати порогові значення для прийняття рішення та гнучко налаштовувати баланс між виявленням шахрайства та кількістю помилкових спрацювань.

Важливою характеристикою випадкового лісу є можливість оцінки важливості ознак. Це досягається шляхом вимірювання того, наскільки зменшується помилка класифікації при використанні конкретної ознаки для розбиття вузлів дерев. Ознаки, які часто використовуються для розбиття і призводять до значного покращення класифікації, отримують вищі оцінки важливості. Ця інформація може бути корисною для розуміння того, які характеристики транзакцій найбільш важливі для детектування шахрайства.

Архітектура включає кілька ключових аспектів, які впливають на її роботу. Сукупність дерев в ансамблі визначає, скільки незалежних моделей буде побудовано. Збільшення кількості дерев зазвичай покращує якість класифікації, але також збільшує обчислювальні витрати. Максимальна глибина дерев контролює вагомість моделей у ансамблі. Мінімальна сукупність зразків для вузла та мінімальна кількість зразків у листі запобігають створенню надто специфічних розбиттів.

Для задачі виявлення шахрайських операцій було встановлено такі параметри моделі кількість дерев дорівнює 100, що забезпечує достатню стабільність передбачень без надмірних обчислювальних витрат. Максимальна глибина дерев не обмежується, що дозволяє моделі вивчати складнощі в даних. Мінімальна сукупніс зразків для розбиття встановлена як 2, а мінімальна кількість зразків у листі як 1, що відповідає стандартним налаштуванням для випадкового лісу.

Кількість ознак, що розглядаються при кожному розбитті, визначається автоматично як квадратний корінь від загальної кількості ознак у датасеті. Це забезпечує оптимальний баланс між різноманітністю дерев та їх індивідуальною точністю. При навчанні використовується критерій Джині для оцінки якості розбиття, який вимірює нечистоту вузла на основі розподілу класів.

2.3 Модифікація моделі та покращення класифікації

Основною проблемою при роботі з даними про банківські транзакції є значний дисбаланс класів, коли кількість легальних операцій значно перевищує кількість шахрайських. У типовому датасеті шахрайські транзакції можуть становити менше 1% від загальної кількості операцій. Така незбалансованість призводить до того, що модель машинного навчання схильна класифікувати більшість транзакцій як легальні, оскільки це забезпечує високу загальну точність.

Для розв’язання проблеми небалансу класів було застосовано техніку передискретизації меншого класу, яка відома під назвою SMOTE. Ця техніка дозволяє збільшити кількість зразків шахрайських транзакцій у навчальній вибірці шляхом генерації синтетичних прикладів. Основна ідея методу полягає в створенні нових зразків не шляхом простого копіювання існуючих шахрайських транзакцій, а через інтерполяцію між сусідніми зразками в просторі ознак.

Процес генерації синтетичних зразків починається з вибору випадкового зразка шахрайської транзакції з навчальної вибірки. Для цього зразка визначаються його найближчі сусіди серед інших шахрайських транзакцій за допомогою евклідової метрики у просторі ознак. Зазвичай розглядається від трьох до п'яти найближчих сусідів, що забезпечує достатню різноманітність генерованих зразків.

Після визначення найближчих сусідів випадковим чином обирається один з них. Новий синтетичний зразок створюється шляхом лінійної інтерполяції між початковим зразком та обраним сусідом. Для кожної ознаки обчислюється різниця між значеннями у двох зразках, ця різниця множиться на вип адкове число від 0 до 1, і результат додається до значення ознаки початкового зразка. Таким чином, новий синтетичний зразок розташовується на відрізку між двома існуючими зразками в просторі ознак.

Цей процес повторюється необхідну кількість разів до досягнення бажаного балансу класів у навчальній вибірці. Кількість синтетичних зразків, що генеруються, визначається коефіцієнтом передискретизації, який показує, у скільки разів потрібно збільшити менший клас. У даному методі було обрано коефіцієнт, який забезпечує

приблизно рівну кількість легальних та шахрайських транзакцій у навчальній вибірці.

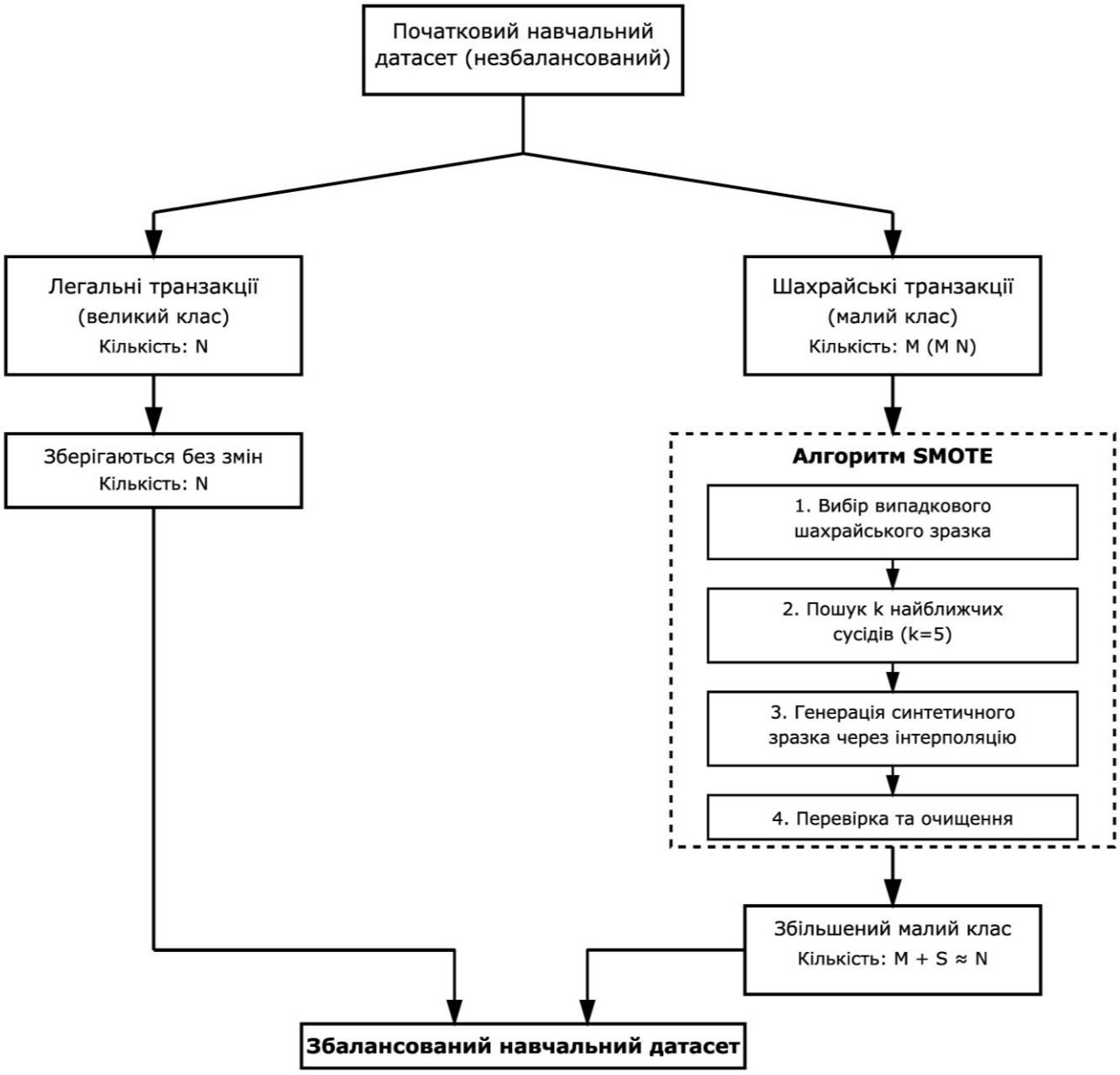


Рисунок 2.3 – Процес балансування класів у навчальній вибірці

Важливою особливістю застосування SMOTE є те, що передискретизація виконується лише на навчальній вибірці даних. Валідаційна та тестова вибірки залишаються незміненими і містять природний розподіл класів, який відповідає реальній ситуації. Це забезпечує об'єктивну оцінку здатності моделі працювати з незбалансованими даними в умовах функціонування.

Використання синтетичних зразків має декілька переваг порівняно з простим дублюванням існуючих шахрайських транзакцій. Створювані зразки мають певну варіативність, що допомагає моделі навчитися узагальнювати патерни шахрайських операцій. Метод дозволяє уникнути точного запам'ятовування моделлю конкретних зразків, що зменшує ризик перенавчання.

Однак застосування SMOTE також має певні обмеження, які необхідно враховувати. Генерація синтетичних зразків може призвести до створення нереалістичних комбінацій ознак, якщо шахрайські транзакції значно відрізняються одна від одної. Крім того, метод може посилити вплив викидів у даних, якщо серед шахрайських операцій є аномальні зразки з нетиповими характеристиками.

Для мінімізації потенційних проблем було використано модифікацію базового методу SMOTE, яка передбачає очищення синтетичних зразків від можливих накладань з класом легальних транзакцій. Після генерації синтетичних шахрайських операцій виконується перевірка на наявність зразків, які розташовані надто близько до легальних транзакцій у просторі ознак. Такі зразки видаляються з навчальної вибірки, оскільки вони можуть створювати шум та погіршувати здатність моделі розрізняти класи.

Процес балансування класів інтегровано у загальний конвеєр підготовки даних та навчання моделі. Після обробки даних попередньо та розділення на вибірки застосовується алгоритм SMOTE лише до навчальної частини. Збалансована навчальна вибірка потім використовується для того, щоб навчити моделі випадкового лісу.

Експериментальна перевірка показала, що застосування техніки балансування класів дозволяє суттєво підвищити здатність моделі виявляти шахрайські операції. Без балансування модель схильна класифікувати майже всі транзакції як легальні, досягаючи високої загальної точності, але пропускаючи більшість шахрайських операцій. Після застосування SMOTE модель стає більш чутливою до шахрайських транзакцій, що відображається у покращенні метрики повноти.

Важливо відзначити, що балансування класів впливає не лише на навчання моделі, але й на інтерпретацію її передбачень. Оскільки модель навчена на збалансованих даних, її оцінки ймовірностей можуть бути зміщеними порівняно з реальним розподілом класів. Тому після отримання передбачень моделі може знадобитися калібрування ймовірностей для приведення їх у відповідність до реального співвідношення легальних та шахрайських операцій.

Крім балансування класів, було також застосовано техніку налаштування вагових коефіцієнтів класів у функції втрат моделі випадкового лісу. Цей підхід передбачає присвоєння більшої ваги помилкам класифікації шахрайських операцій порівняно з легальними. Вагові коефіцієнти встановлюються обернено пропорційно до частоти класів у початкових даних, що змушує модель приділяти більше уваги меншому класу.

Комбінація техніки передискретизації SMOTE та налаштування вагових коефіцієнтів забезпечує найкращі результати у виявленні підозрілих операцій. Ці два підходи доповнюють один одного, оскільки SMOTE збільшує кількість навчальних зразків шахрайських транзакцій, а вагові коефіцієнти впливають на навчання моделі, підвищуючи важливість правильної класифікації цих зразків.

2.4 Формування та підготовка навчальних даних

Якість навчання моделі машинного навчання значною мірою залежить від правильної підготовки вхідних даних. Процес формування та підготовки навчальних даних включає декілька послідовних етапів, кожен з яких спрямований на приведення даних до формату, придатного для навчання алгоритмів класифікації.

Початковим етапом є завантаження датасету банківських транзакцій з файлу у форматі CSV. Датасет містить інформацію про велику кількість операцій, де кожна транзакція описується набором ознак, таких як сума операції, тип рахунку відправника та отримувача, час здійснення транзакції, географічне розташування та інші параметри. Окрема колонка містить мітку класу, яка вказує, чи є транзакція легальною або шахрайською.

Після завантаження даних виконується їх первинний аналіз для визначення основних характеристик датасету. Цей аналіз включає перевірку розмірності даних, підрахунок кількості транзакцій кожного класу, визначення типів ознак та їх статистичних характеристик. На цьому виявляється наявність пропущених значень, дублікатів записів та потенційних викидів у даних.

Обробка пропущених значень є необхідним кроком підготовки даних, тому, що більшість алгоритмів навчання не можуть працювати з неповними даними. Для числових ознак пропущені значення зазвичай заповнюються медіаною розподілу відповідної ознаки, оскільки медіана є стійкою до викидів статистикою. Для категоріальних типів ознак використовується заповнення найбільш частим значенням або створення окремої категорії для пропущених значень

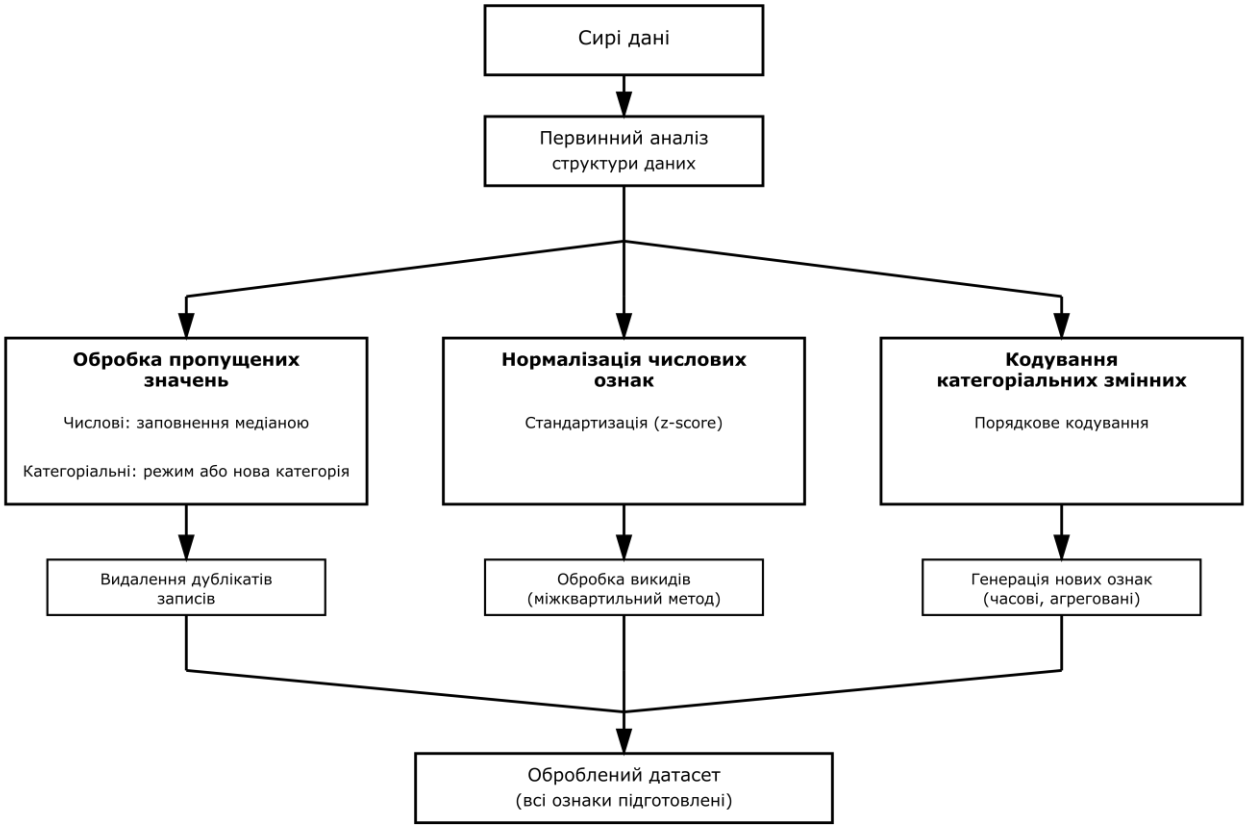


Рисунок 2.4 – Етапи попередньої обробки та підготовки даних

Виявлення та обробка викидів виконується за допомогою статистичних методів. Для кожної числової ознаки обчислюється міжквартильний розмах, і значення, які виходять за межі півтора міжквартильних розмахів від першого або

третього квартилю, розглядаються як потенційні викиди. Однак у контексті виявлення шахрайства викиди можуть бути справжніми аномальними операціями, тому їх видалення виконується обережно з урахуванням контексту задачі.

Нормалізація числових ознак є обов'язковим кроком для забезпечення коректної роботи багатьох алгоритмів навчання. Ознаки можуть мати різні масштаби значень, що може призвести до домінування одних ознак над іншими у навчанні. Для нормалізації використовується метод стандартизації, який перетворює кожну ознаку таким чином, щоб вона мала середнє нуль та низьку дисперсію.

Процес нормалізації виконується окремо для навчальної, валідаційної та тестової вибірок. Параметри нормалізації обчислюються лише на навчальній вибірці, після чого ці самі параметри застосовуються до валідаційної та тестової вибірок. Це забезпечує коректну оцінку узагальнювальної здатності моделі та запобігає витоку інформації з тестових даних у процес навчання.

Кодування категоріальних змінних необхідне для перетворення текстових або символьних значень у числову форму, яку можуть обробляти алгоритми машинного навчання. Для категоріальних ознак з природним порядком використовується порядкове кодування, де кожній категорії присвоюється числове значення відповідно до її позиції в упорядкованій послідовності. Для номінальних категоріальних ознак застосовується однократне кодування, яке створює окрему бінарну ознаку для кожної категорії.

Однократне кодування має особливість, яка полягає у тому, що воно збільшує розмірність простору ознак пропорційно кількості унікальних категорій у кожній категоріальній змінній. Це може призвести до значного збільшення кількості ознак, особливо якщо категоріальні змінні мають багато різних значень. Для зменшення розмірності може застосовуватися групування рідкісних категорій або використання інших методів кодування.

Розділення даних на вибірки виконується з використанням стратифікованої вибірки, яка забезпечує пропорції класів у підвибірці. Навчальна вибірка використовується для того, щоб навчати моделі, валідаційна вибірка для

налаштування параметрів та запобігання перенавчанню, а тестова вибірка для фінальної оцінки якості моделі.

Стратифіковане розділення особливо важливе для незбалансованих датасетів, оскільки воно гарантує, що кожна підвибірка містить достатню кількість зразків обох класів. Випадкове розділення без стратифікації може призвести до ситуації, коли деякі підвибірки містять занадто мало або взагалі не містять шахрайських транзакцій, що робить неможливою коректну валідацію та тестування моделі.

Створення додаткових ознак може поліпшити можливість моделі виявляти шахрайські операції. Такі ознаки генеруються на основі існуючих характеристик транзакцій та можуть включати статистичні показники історії операцій клієнта, часові патерни або комбінації декількох базових ознак. Наприклад, можна створити ознаки, які показують відхилення поточної транзакції від типових операцій клієнта за сумою або часом здійснення.

Агрегація інформації про попередні транзакції клієнта дозволяє моделі враховувати контекст поточної операції. Для кожної транзакції можуть бути обчислені такі характеристики як середня сума попередніх операцій клієнта, кількість транзакцій за останню добу або тиждень, типові категорії платежів та інші показники поведінки. Ці агреговані ознаки допомагають моделі розпізнавати аномалії у поведінці клієнта.

Часові ознаки відіграють важливу роль у виявленні шахрайських операцій, оскільки час здійснення транзакції може бути індикатором підозрілої активності. З мітки часу транзакції можуть бути витягнуті такі характеристики як година доби, день тижня, день місяця та чи є день робочим або вихідним. Шахрайські операції часто відбуваються у незвичний час, коли легітимний власник рахунку зазвичай не здійснює транзакцій.

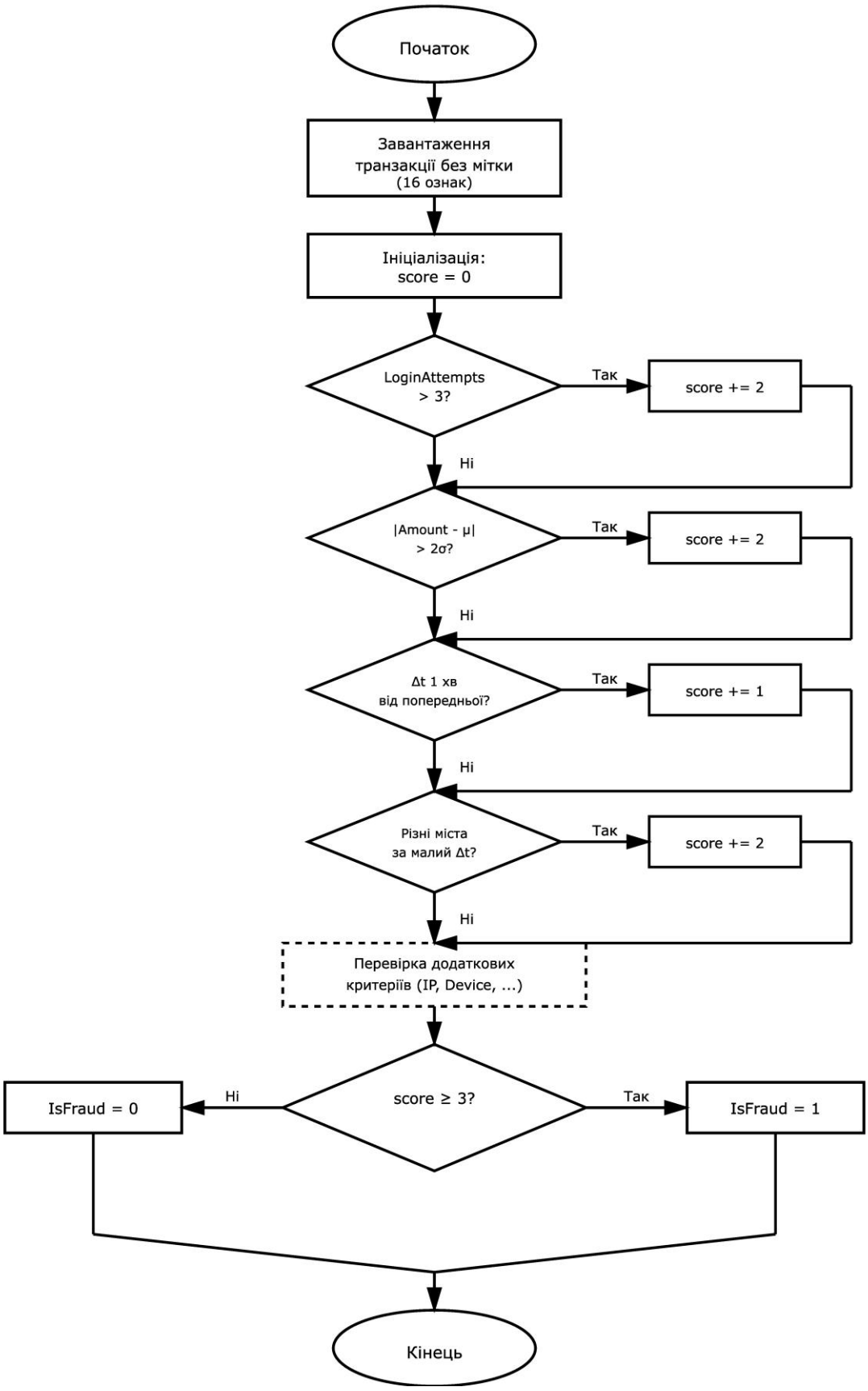


Рисунок 2.5 – Процес генерації міток класів для транзакцій

Остаточна навчальна вибірка після всіх етапів підготовки містить повний набір оброблених ознак та збалансовану кількість легальних і шахрайських транзакцій. Ця вибірка готова для використання у навчанні моделі випадкового лісу, яка буде аналізувати зв'язки між ознаками та класами транзакцій для побудови правил класифікації.

Особливістю використовуваного датасету є відсутність готових міток класів для транзакцій, що вимагає додаткового етапу генерації цих міток на основі виявлення аномальної поведінки. Для створення навчальної вибірки з розміченими даними застосовується комбінація правил, які базуються на експертних знаннях про типові ознаки шахрайських операцій у банківській сфері.

Процес генерації міток класів виконується на основі аналізу декількох ключових характеристик транзакцій, які найчастіше вказують на підозрілу активність. Кількість спроб входу в систему перед здійсненням транзакції є одним із найбільш показових індикаторів потенційного шахрайства. Транзакції з кількістю спроб входу більше трьох розглядаються як підозрілі, оскільки це може свідчити про спробу несанкціонованого доступу до рахунку.

Аналіз суми транзакції відносно історичних операцій конкретного рахунку дозволяє виявити незвичні платежі. Для кожного рахунку обчислюється середня сума попередніх транзакцій та їх стандартне відхилення. Операції, сума яких перевищує середнє значення більш ніж на два стандартних відхилення, позначаються як потенційно шахрайські, оскільки вони виходять за межі типової поведінки клієнта.

Часова послідовність транзакцій також є важливим індикатором аномалій. Розрахунок інтервалу між поточною транзакцією та попередньою операцією для того самого рахунку дозволяє виявити незвичайно інтенсивну активність. Якщо дві транзакції відбуваються з інтервалом менше однієї хвилини, це може свідчити про автоматизовану шахрайську діяльність або компрометацію облікового запису.

Географічна консистентність перевіряється шляхом аналізу послідовності локацій транзакцій для одного рахунку. Якщо дві послідовні операції здійснюються з різних міст за короткий проміжок часу, що фізично неможливо для легітимного

користувача, така послідовність класифікується як підозріла. Цей метод дозволяє виявити випадки, коли шахраї використовують вкрадені дані рахунку з іншого географічного розташування.

Зміна технічних параметрів доступу, таких як IP-адреса або ідентифікатор пристрою, також враховується при генерації міток. Для кожного рахунку відстежуються історичні значення цих параметрів, і транзакції з нових, раніше не використовуваних IP-адрес або пристроїв отримують додаткові бали підозрілості. Особливо підозрілими вважаються випадки одночасної зміни обох параметрів.

Тривалість виконання транзакції може вказувати на автоматизовані шахрайські атаки. Операції, які виконуються надзвичайно швидко, менше ніж за п'ять секунд, або навпаки занадто довго, більше двох хвилин, відрізняються від типової поведінки користувачів та позначаються як аномальні. Нормальна тривалість транзакції зазвичай становить від десяти до тридцяти секунд. Стан балансу рахунку після транзакції також аналізується для виявлення підозрілих операцій. Транзакції, які призводять до від'ємного балансу або до різкого зменшення балансу більш ніж на 80%, розглядаються як потенційно шахрайські. Такі операції можуть вказувати на спробу вивести максимальну кількість коштів з компрометованого рахунку. Фінальна мітка класу для кожної транзакції визначається на основі системи балів, де кожна виявлена аномалія додає певну кількість балів підозрілості. Транзакції, які набирають більше трьох балів з можливих восьми критеріїв, класифікуються як шахрайські. Такий підхід дозволяє уникнути помилкової класифікації операцій, які мають лише одну незвичну характеристику, але в цілому відповідають легальній поведінці.

Після генерації міток виконується перевірка розподілу класів у отриманому датасеті. Якщо частка шахрайських транзакцій виявляється занадто малою або занадто великою порівняно з реальною статистикою банківського фроду, параметри правил класифікації коригуються для досягнення більш реалістичного співвідношення класів. Зазвичай цільова частка шахрайських операцій встановлюється на рівні від одного до 5% від загальної сукупності транзакцій.

2.5 Критерії та метрики оцінювання роботи методу

Оцінювання якості роботи методу виявлення шахрайських банківських операцій потребує використання спеціальних метрик, які враховують специфіку задачі класифікації з незбалансованими класами. Стандартна метрика точності, яка обчислюється як відношення правильно класифікованих зразків до загальної кількості зразків, не є достатньо інформативною для такої задачі, оскільки модель може досягти необхідної точності, просто класифікуючи всі транзакції як легальні.

Для всебічної оцінки якості класифікації використовується набір метрик, які дозволяють аналізувати різні аспекти роботи моделі. Метрики грунтуються на матриці помилок, яка містить основні категорії передбачень.

Істинно позитивні результати відповідають шахрайським транзакціям, які правильно класифіковано моделлю як шахрайські. Істинно негативні результати представляють легальні операції, які коректно визначено як легальні. Хибно позитивні результати виникають, коли легальна транзакція помилково класифікується як шахрайська, що призводить до блокування операції клієнта. Хибно негативні результати відповідають шахрайським операціям, які модель не зуміла виявити та класифікувала як легальні.

Метрика точності визначається як відношення істинно позитивних результатів до суми істинно позитивних та хибно позитивних результатів. Ця метрика показує, яку частина транзакцій, було моделлю класифіковано як шахрайські, дійсно є шахрайськими. Висока точність означає, що модель рідко помиляється при визначенні шахрайських операцій, що зменшує кількість блокувань легальних транзакцій клієнтів.

Метрика повноти обчислюється як відношення істинно позитивних результатів до суми істинно позитивних та хибно негативних результатів. Повнота відображає здатність моделі виявляти шахрайські операції серед усіх реальних випадків шахрайства у датасеті. Висока повнота вказує на те, що знаходить модель більшість шахрайських транзакцій, мінімізуючи фінансові втрати банку від пропущених випадків фроду.

Між точністю та повнотою існує певний компроміс, оскільки підвищення одної метрики часто призводить до зниження іншої. Модель, налаштована на високу точність, буде консервативною у визначенні шахрайства та може пропустити деякі шахрайські операції. Навпаки, модель з високою повнотою буде більш агресивною у виявленні фроду, але може класифікувати багато легальних транзакцій як підозрілі.

Для знаходження балансу між точністю та повнотою використовується метрика F1-міра, яка є гармонійним середнім цих двох показників. F1-міра досягає високого значення тільки тоді, коли обидві метрики є високими, що дає змогу робити її корисною для загальної оцінки якості моделі. Ця метрика обчислюється як подвоєний добуток точності та повноти, поділений на їх суму.

Крива точності-повноти є графічним інструментом для аналізу роботи моделі при різних порогових значеннях класифікації. Кожна точка на лінії кривої відповідає якомусь порогу ймовірності, при якому транзакція класифікується як шахрайська. Переміщуючись вздовж кривої, можна вибрати оптимальний поріг, який забезпечує бажане співвідношення між точністю, а також повнотою для конкретних потреб банківської установи.

Площа під кривою точності-повноти є числовою характеристикою, яка узагальнює роботу моделі при всіх можливих порогах. Значення цієї метрики знаходиться у діапазоні від нуля до одиниці, де більше значення вказує на кращу якість класифікації. Площа під кривою дозволяє порівнювати різні моделі без необхідності вибору конкретного порогу класифікації.

ROC-крива є ще одним важливим інструментом оцінювання моделі, який показує залежність між часткою істинно позитивних результатів та часткою хибно позитивних результатів при зміні порогу класифікації. Площа під ROC-кривою використовується як агрегована метрика якості моделі, де значення близьке до одиниці свідчить про хорошу здатність моделі розрізнювати класи.

Специфічність є метрикою, яка обчислюється як відношення істинно негативних результатів до суми істинно негативних та хибно позитивних результатів. Ця метрика говорить, яка частка легальних транзакцій правильно

класифікується моделлю. Висока специфічність важлива для зменшення незручностей клієнтів через помилкові блокування їх операцій.

Матрика Меттьюса є збалансованою метрикою, яка враховує всі чотири категорії матриці помилок та може бути використана навіть для дуже незбалансованих датасетів. Значення цієї метрики знаходиться у діапазоні від мінус одиниці до плюс одиниці, де одиниця відповідає ідеальній класифікації, нуль випадковому передбаченню, а мінус одиниця повній невідповідності передбачень реальним класам.

Час обробки транзакції є важливим практичним критерієм для систем виявлення фроду, які мають працювати в реальному часі. Модель повинна класифікувати кожну операцію протягом долей секунди, щоб не створювати затримок у процесі проведення платежів. Тому при виборі моделі враховується не лише її точність, але й обчислювальна складність.

Стабільність передбачень моделі оцінюється шляхом багаторазового навчання на різних розбиттях даних та обчислення дисперсії метрик якості. Модель вважається стабільною, якщо її показники мало змінюються при різних ініціалізаціях та розбиттях навчальної вибірки. Висока стабільність важлива для забезпечення надійної роботи системи виявлення фроду у виробничих умовах.

Інтерпретованість моделі, хоча і важко виразити числовими метриками, є важливим критерієм для банківського застосування. Здатність пояснити, чому конкретна транзакція була класифікована як шахрайська, необхідна для дотримання регуляторних вимог та підтримки довіри клієнтів. Випадковий ліс надає інформацію про важливість ознак, що допомагає зрозуміти логіку прийняття рішень моделлю.

Крім статичних метрик, які оцінюють модель на фіксованому тестовому наборі, важливо відстежувати якість роботи моделі з часом у реальних умовах експлуатації. Зміна характеристик шахрайських операцій може дати погіршення якості класифікації, що вимагає періодичного перенавчання моделі на нових даних. Моніторинг метрик у виробничому середовищі дозволяє вчасно виявити потребу в оновленні моделі.

Висновок до розділу 2

У даному розділі було описано метод виявлення шахрайських банківських операцій з використанням алгоритму випадкового лісу. Представлено концепцію методу, яка включає послідовність етапів від завантаження вхідних даних до прийняття рішення про класифікацію транзакції.

Описано архітектуру моделі класифікації на основі випадкового лісу, яка будує ансамбль дерев рішень для прийняття колективного рішення про класифікацію транзакцій. Обґрунтовано вибір цього алгоритму його стійкістю до перенавчання, здатністю працювати з великою кількістю ознак та можливістю оцінки важливості характеристик транзакцій для виявлення шахрайства. Запропоновано модифікацію базового методу, яка полягає у застосуванні техніки балансування класів SMOTE для вирішення проблеми наявної незбалансованості даних. Описано процес генерації синтетичних зразків шахрайських транзакцій шляхом інтерполяції між існуючими зразками у просторі ознак. Показано, що поєднання передискретизації меншого класу з налаштуванням вагових коефіцієнтів у функції втрат дає змогу поліпшити здатність моделі виявляти шахрайські операції.

Розглянуто процес формування та підготовки навчальних даних, який включає генерацію міток класів на основі системи правил виявлення аномалій, обробку відсутніх значень, нормалізацію ознак які є числами, кодування категоріальних змінних та стратифіковане розділення на навчальну, валідаційну та тестову вибірки.

Визначено набір критеріїв та метрик для оцінювання якості роботи методу, які враховують специфіку задачі класифікації з незбалансованими класами. Описано такі метрики як точність, повнота, F1-міра, площа під кривою точності-повноти. Обґрунтовано необхідність використання комплексу метрик для всебічної оцінки здатності моделі виявляти шахрайські операції при мінімізації помилкових спрацювань.

Розділ 3 Програмна реалізація методу виявлення шахрайських операцій

3.1 Технології та інструменти програмної реалізації

Для реалізації методу виявлення шахрайських банківських операцій було обрано для програмування Python версії 3.10 як основний інструмент розробки. Вибір цієї мови обумовлений наявністю потужних бібліотек для роботи з даними та машинного навчання, широкою підтримкою спільноти розробників та зручним синтаксисом для швидкого прототипування.

Для роботи з даними використовувалась бібліотека pandas версії 2.0.3, яка надає зручні структури даних для зберігання та маніпулювання табличними даними. Основною структурою є DataFrame, що дозволяє зберігати дані про транзакції у вигляді таблиці з іменованими стовпцями. Бібліотека pandas містить функції для читання даних з CSV-файлів, обробки пропущених значень, фільтрації рядків та стовпців, обчислення статистичних показників.

Числові обчислення виконувались з використанням бібліотеки NumPy версії 1.24.3, яка надає багатовимірні масиви та функції для роботи з ними. Ця бібліотека використовує векторизовані операції для обробки масивів даних. NumPy застосовувався для обчислення статистичних показників, нормалізації ознак та виконання математичних операцій над векторами і матрицями.

Основна функціональність машинного навчання реалізована з використанням бібліотеки scikit-learn версії 1.3.0. Ця бібліотека містить реалізації різних алгоритмів класифікації та інструменти для попередньої обробки даних. З бібліотеки scikit-learn використовувались наступні компоненти RandomForestClassifier для реалізації моделі випадкового лісу, StandardScaler для нормалізації числових ознак, OneHotEncoder для кодування категоріальних змінних, SimpleImputer для обробки пропущених значень, train_test_split для розділення даних на вибірки.

Для вирішення проблеми незбалансованості класів застосовувалась бібліотека imbalanced-learn версії 0.11.0, яка надає різні методи балансування даних. Зокрема, використовувався алгоритм SMOTE для генерації синтетичних зразків шахрайських транзакцій методом інтерполяції між існуючими зразками.

Візуалізація результатів здійснювалась за допомогою бібліотеки matplotlib версії 3.7.2, що дозволяє створювати різноманітні графіки та діаграми. Додатково використовувалась бібліотека seaborn версії 0.12.2 для створення статистичних візуалізацій з покращеним дизайном.

Середовищем розробки обрано Jupyter Notebook, що надає інтерактивне середовище для написання та виконання коду Python. Jupyter Notebook дозволяє поєднувати код, текст та візуалізації в одному документі, що зручно для дослідницьких задач та демонстрації результатів.

Для контролю версій коду використовувалась система Git з розміщенням репозиторію на платформі GitHub. Це дозволило зберігати історію змін, повертатись до попередніх версій коду та забезпечити резервне копіювання розробленого програмного забезпечення.

Усі зазначені інструменти є вільно поширюваними з відкритим вихідним кодом, що дозволяє використовувати їх без додаткових фінансових витрат. Комбінація цих технологій забезпечує потужну платформу для розробки системи виявлення шахрайських операцій.

3.2 Структура програмної системи та основні модулі

Програмна реалізація методу виявлення шахрайських операцій організована у вигляді модульної системи, де кожен модуль відповідає за окрему функціональність. Така архітектура забезпечує зручність розробки, тестування та підтримки коду, а також дозволяє незалежно модифікувати окремі компоненти без впливу на інші частини системи.

Загальна структура системи містить 7 основних модулів, які взаємодіють між собою для виконання повного циклу обробки даних, навчання моделі та класифікації транзакцій. Діаграма компонентів системи представлена на рисунку 3.1, де відображено всі модулі, напрямки передачі даних та залежності між компонентами.

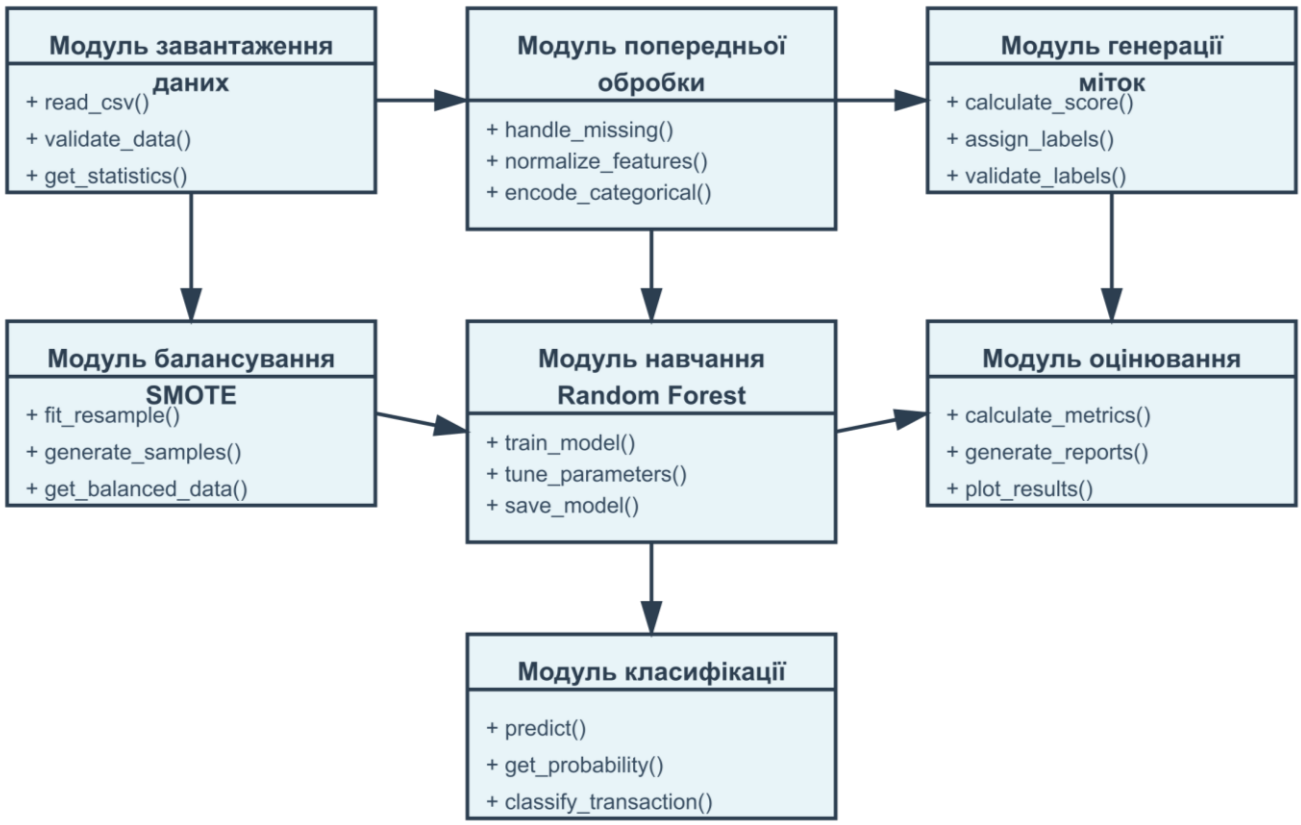


Рисунок 3.1 – Діаграма компонентів

Модуль завантаження даних відповідає за читання вхідних даних з CSV-файлу та їх первинну валідацію. Цей модуль містить функції для перевірки наявності всіх необхідних стовпців у датасеті, визначення типів даних та виявлення дублікатів. Модуль також виконує розділення даних на навчальну, валідаційну та тестову вибірки у співвідношенні 70%, 15% та 15% відповідно зі збереженням пропорції класів у кожній вибірці.

Модуль попередньої обробки виконує трансформацію сирих даних у формат, придатний для навчання моделі. Обробка пропущених значень здійснюється шляхом заповнення медіаною для числових ознак та найбільш частим значенням для категоріальних змінних з використанням класу SimpleImputer. Нормалізація числових ознак використовує метод стандартизації через клас StandardScaler, що перетворює кожну ознаку до розподілу з нульовим середнім та одиничною дисперсією. Кодування категоріальних змінних виконується методом однократного кодування через клас OneHotEncoder, що створює бінарні стовпці для кожної можливої категорії.

Модуль генерації міток класів реалізує систему правил для автоматичного визначення шахрайських транзакцій у датасеті, що не містить готових міток. Для кожної транзакції обчислюється бал підозрілості на основі 8 критеріїв кількість спроб невдалого входу в систему, відхилення суми транзакції від типових значень, часовий інтервал між послідовними операціями, зміна географічного розташування, використання нового пристрою, зміна IP-адреси, тривалість виконання операції та стан балансу рахунку. Кожен критерій додає певну кількість балів, і транзакції з сумарним балом вище порогового значення отримують мітку шахрайської операції.

Модуль балансування класів застосовує метод SMOTE для усунення проблеми незбалансованості даних. Цей модуль аналізує розподіл класів у навчальній вибірці та генерує синтетичні зразки шахрайських транзакцій до досягнення бажаного співвідношення класів. Генерація нових зразків відбувається шляхом знаходження п'яти найближчих сусідів для кожної шахрайської транзакції в просторі ознак та створення нових точок на відрізках між існуючими зразками.

Модуль навчання моделі відповідає за створення, налаштування та навчання класифікатора Random Forest. Модуль ініціалізує модель з параметрами кількість дерев в ансамблі 100, критерій розбиття Джині для оцінки якості вузлів. Навчання моделі виконується на збалансованій навчальній вибірці. Після навчання модуль обчислює важливість кожної ознаки на основі її внеску в покращення якості класифікації та зберігає навчену модель у файл для подальшого використання.

Модуль оцінювання якості розраховує різноманітні метрики для визначення точності роботи моделі. На тестовій вибірці обчислюються такі показники як точність, повнота, F1-міра, специфічність та площа під ROC-кривою. Модуль також будує матрицю помилок, яка показує розподіл правильних та неправильних передбачень для кожного класу. Результати оцінювання представляються у вигляді таблиць та графіків для зручного аналізу якості моделі.

Модуль класифікації транзакцій використовує навчену модель для передбачення класу нових операцій. Цей модуль отримує дані про транзакцію, застосовує ті самі трансформації, що використовувались при навчанні, передає оброблені дані моделі для класифікації та повертає результат у вигляді класу і

ймовірності шахрайства. Модуль може обробляти як окремі транзакції, так і пакети операцій.

Взаємодія між модулями організована таким чином, що дані послідовно проходять через ланцюжок обробки від завантаження до отримання фінального результату класифікації. Кожен модуль має визначений інтерфейс, що спрощує тестування та модифікацію окремих компонентів.

3.3 Класи модуля обробки даних

Детальна структура класів модуля обробки даних представлена на діаграмі класів (рисунок 3.2). Цей модуль включає три основні класи DataLoader, DataPreprocessor та LabelGenerator, кожен з яких виконує специфічні функції в процесі підготовки даних.

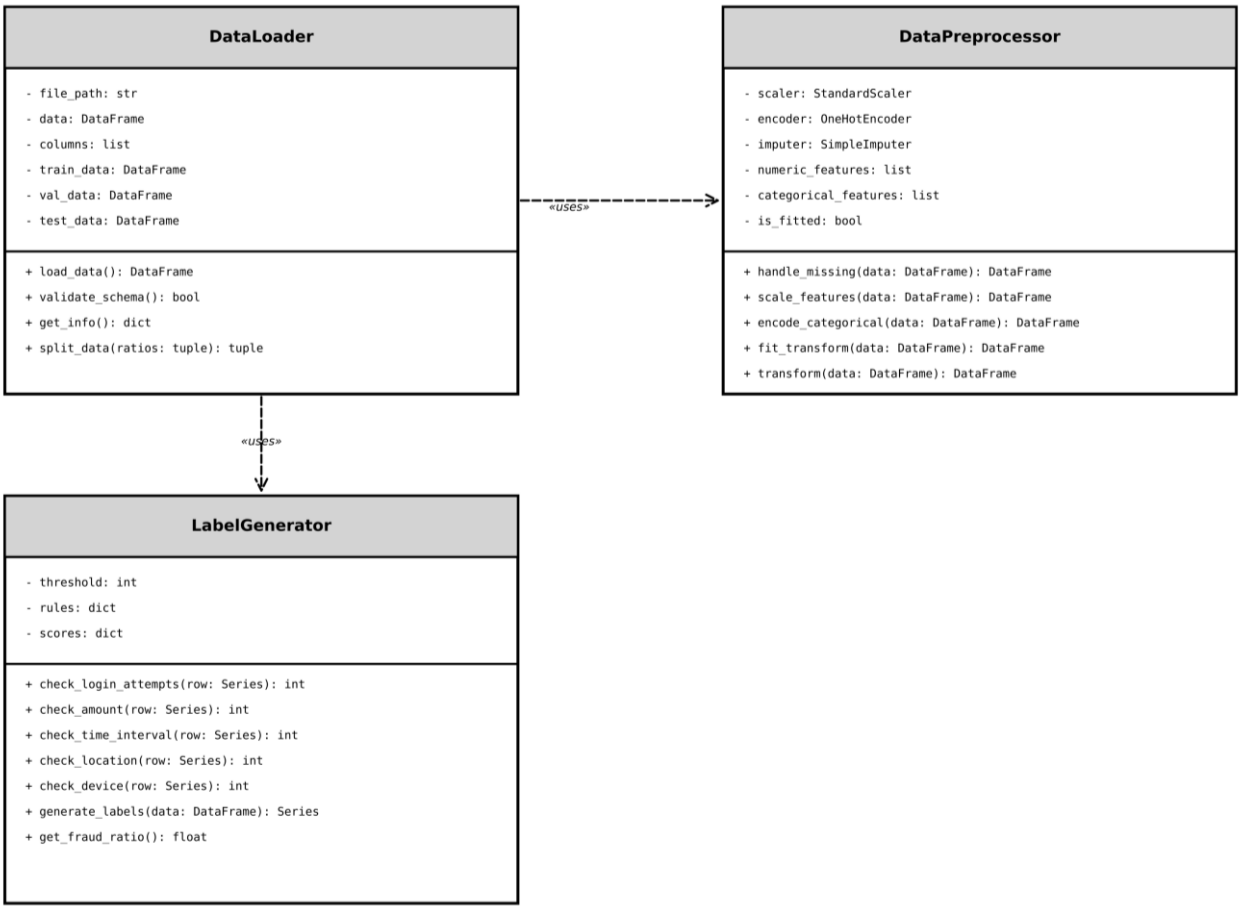


Рисунок 3.2 – Діаграма класів модуля обробки даних

Клас DataLoader інкапсулює логіку завантаження та валідації даних. Клас містить приватні атрибути file_path для зберігання шляху до файлу даних, data для зберігання завантажених даних у форматі DataFrame, columns для списку назв стовпців, а також train_data, val_data та test_data для зберігання розділених вибірок.

Публічні методи класу DataLoader включають метод load_data, який читає дані з CSV-файлу та повертає DataFrame; метод validate_schema, який перевіряє наявність усіх необхідних стовпців та повертає булеве значення; метод get_info, який повертає словник з інформацією про датасет; метод split_data, який приймає кортеж співвідношень та розділяє дані на вибірки, повертаючи кортеж з трьома DataFrame.

Клас DataPreprocessor відповідає за всі трансформації даних перед навчанням моделі. Приватні атрибути класу включають scaler типу StandardScaler для нормалізації, encoder типу OneHotEncoder для кодування категоріальних змінних, imputer типу SimpleImputer для обробки пропущених значень, списки numeric_features та categorical_features для зберігання назв відповідних типів ознак, а також булевий прапорець is_fitted для контролю стану навченості трансформерів.

Публічні методи класу DataPreprocessor метод handle_missing приймає DataFrame та повертає DataFrame з заповненими пропущеними значеннями; метод scale_features приймає та повертає DataFrame з нормалізованими числовими ознаками; метод encode_categorical приймає та повертає DataFrame з закодованими категоріальними змінними; метод fit_transform навчає всі трансформери на навчальних даних та застосовує трансформації; метод transform застосовує збережені трансформації до нових даних без повторного навчання.

Клас LabelGenerator реалізує систему генерації міток класів на основі набору правил. Приватні атрибути класу threshold зберігає порогове значення балів для класифікації як шахрайської, rules містить словник з описом правил перевірки, scores зберігає словник з балами для кожного критерію.

Публічні методи класу LabelGenerator включають методи перевірки окремих критеріїв check_login_attempts перевіряє кількість спроб входу та повертає бал; check_amount перевіряє відхилення суми транзакції; check_time_interval аналізує

часовий інтервал між операціями; check_location перевіряє зміну географічного розташування; check_device перевіряє зміну пристрою або IP-адреси. Метод generate_labels приймає DataFrame та повертає Series з мітками класів. Метод get_fraud_ratio повертає відсоток шахрайських транзакцій.

Зв'язки між класами показані на діаграмі клас DataLoader використовує клас DataPreprocessor для обробки завантажених даних, що відображено пунктирною стрілкою з позначкою «uses». Також DataLoader використовує клас LabelGenerator для генерації міток класів на завантажених даних.

3.4 Класи модуля навчання моделі

Структура класів модуля навчання моделі представлена на діаграмі класів (рисунок 3.3). Цей модуль включає чотири основні класи SMOTEBalancer, RandomForestModel, MetricsEvaluator та TransactionClassifier.

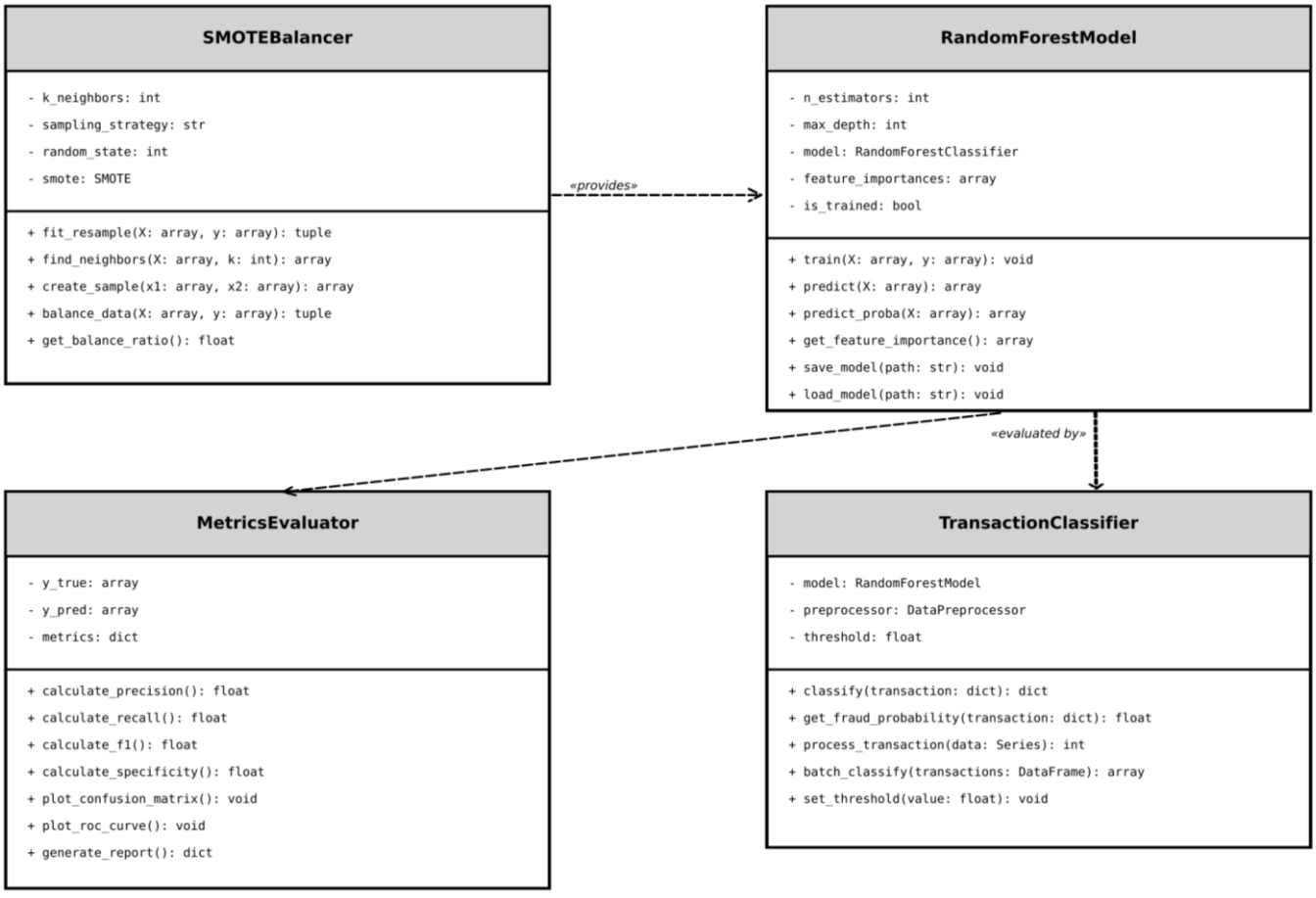


Рисунок 3.3 – Діаграма класів модуля навчання моделі

Клас SMOTEBalancer інкапсулює функціональність балансування класів методом SMOTE. Приватні атрибути класу k_neighbors зберігає кількість найближчих сусідів для генерації нових зразків, sampling_strategy визначає стратегію балансування, random_state забезпечує відтворюваність результатів, smote містить об'єкт класу SMOTE з бібліотеки imbalanced-learn.

Публічні методи класу SMOTEBalancer метод fit_resample приймає масиви ознак X та міток y, виконує балансування та повертає кортеж з балансованими масивами; метод find_neighbors знаходить k найближчих сусідів для заданих точок; метод create_sample створює один синтетичний зразок між двома точками; метод balance_data є основним методом балансування; метод get_balance_ratio повертає співвідношення класів після балансування.

Клас RandomForestModel обгортає функціональність scikit-learn для зручного використання. Приватні атрибути n_estimators зберігає кількість дерев в ансамблі, max_depth обмежує максимальну глибину дерев, model містить об'єкт RandomForestClassifier, feature_importances зберігає масив важливості ознак, is_trained є прапорцем стану навченості моделі.

Публічні методи класу RandomForestModel метод trainin приймає масиви x та y і навчає модель; метод predicting масив ознак та повертає масив передбачених класів; метод predict_proba повертає масив ймовірностей належності до кожного класу; метод get_feature_importance повертає масив важливості ознак; методи save_model та load_model здійснюють серіалізацію та десеріалізацію моделі.

Клас MetricsEvaluator відповідає за обчислення метрик якості моделі. Приватні атрибути y_true зберігає справжні мітки, y_pred зберігає передбачені мітки, metrics містить словник з обчисленими метриками.

Публічні методи класу MetricsEvaluator методи calculate_precision, calculate_recall, calculate_f1 та calculate_specificity обчислюють відповідні метрики та повертають числові значення; метод plot_confusion_matrix будує та відображає матрицю помилок; метод plot_roc_curve будує ROC-криву; метод generate_report формує повний звіт з усіма метриками у форматі словника.

Клас TransactionClassifier надає інтерфейс для класифікації транзакцій. Приватні атрибути model містить об'єкт RandomForestModel, preprocessor містить об'єкт DataPreprocessor, threshold зберігає порогове значення ймовірності для класифікації.

Публічні методи класу TransactionClassifier метод classify приймає словник з даними транзакції та повертає словник з результатом класифікації; метод get_fraud_probability обчислює ймовірність шахрайства; метод process_transaction обробляє дані однієї транзакції та повертає клас; метод batch_classify обробляє DataFrame з множиною транзакцій та повертає масив класів; метод set_threshold дозволяє змінювати порогове значення.

Зв'язки між класами SMOTEBalancer надає збалансовані дані для RandomForestModel, що показано стрілкою «provides»; RandomForestModel оцінюється через MetricsEvaluator, що показано стрілкою «evaluated by»; TransactionClassifier містить RandomForestModel як свою складову частину, що відображено через композицію з заповненим ромбом на стороні TransactionClassifier.

3.5 Послідовність процесу навчання моделі

Навчання моделі включає взаємодію між 5 об'єктами користувачем системи, завантажувачем даних, препроцесором, балансувальником та моделлю. Детальна послідовність взаємодії представлена на діаграмі послідовності (рисунок 3.4).

Процес починається з того, що користувач ініціює завантаження через виклик методу load_data з передачею шляху до файлу. Об'єкт DataLoader отримує цей запит і виконує послідовність внутрішніх операцій спочатку викликається метод read_csv для читання даних з файлу, потім метод validate_schema для перевірки структури даних.

Після успішної валідації DataLoader передає дані об'єкту DataPreprocessor через виклик методу preprocess. DataPreprocessor виконує три послідовні трансформації метод handle_missing заповнює пропущені значення, метод

scale_features нормалізує числові ознаки, метод encode_categorical кодує категоріальні змінні. Після завершення всіх трансформацій оброблені дані повертаються DataLoader.

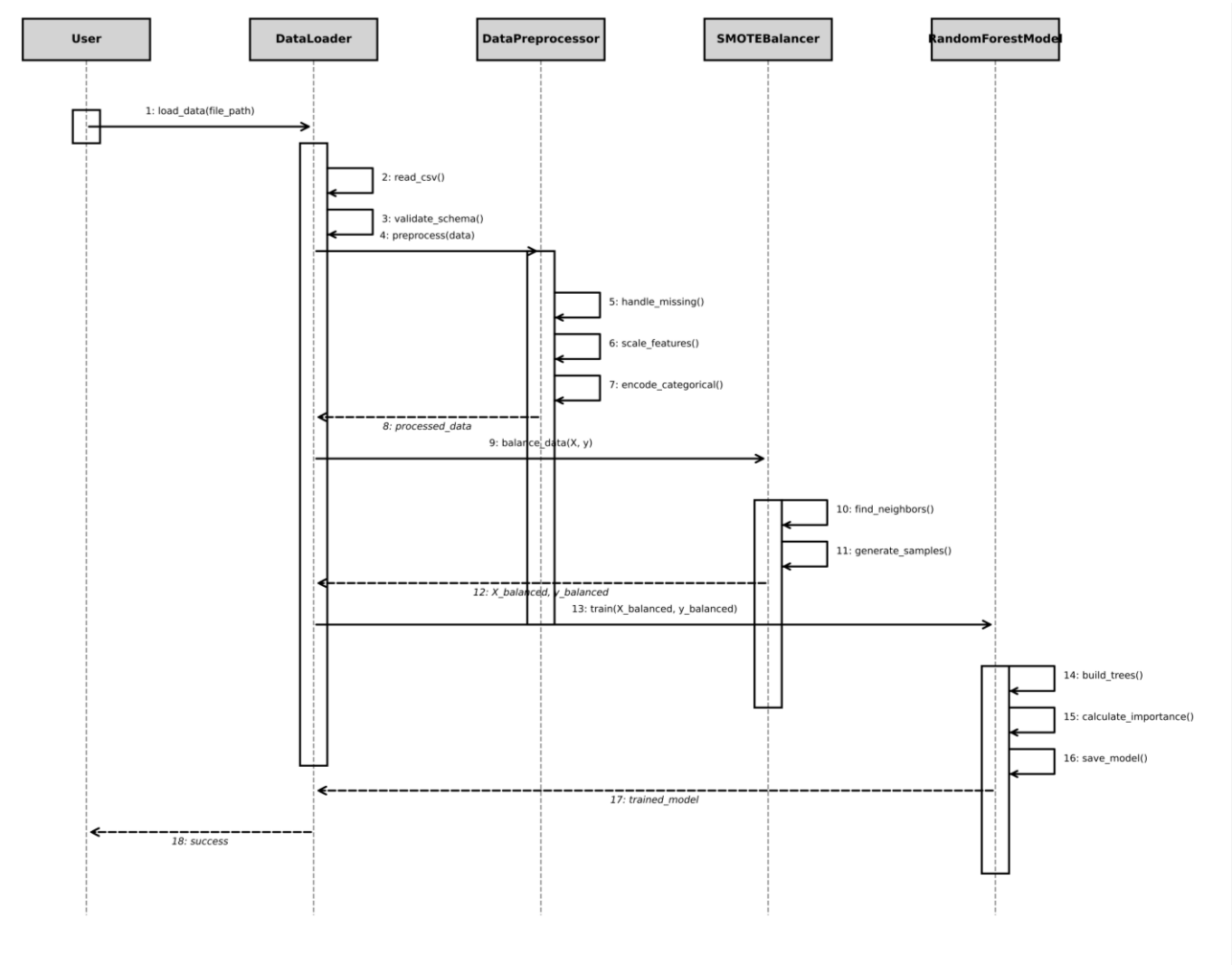


Рисунок 3.4 – Діаграма послідовності процесу навчання моделі

Наступним кроком DataLoader передає оброблені дані та мітки класів об'єкту SMOTEBalancer через виклик методу balance_data. SMOTEBalancer виконує внутрішні операції метод find_neighbors знаходить найближчі сусіди для кожного зразка шахрайської транзакції, метод generate_samples створює синтетичні зразки шляхом інтерполяції. Після генерації достатньої кількості синтетичних зразків збалансовані масиви X_balanced та y_balanced повертаються DataLoader.

Отримавши збалансовані дані, DataLoader передає їх об'єкту RandomForestModel через виклик методу train. RandomForestModel виконує процес навчання метод build_trees будує ансамбль з дерев рішень, кожне на своїй bootstrap вибірці; метод calculate_importance обчислює важливість кожної ознаки на основі її внеску в зменшення нечистоти; метод save_model зберігає навчену модель у файл.

Після завершення навчання об'єкт RandomForestModel повертає навчену модель DataLoader. DataLoader, у свою чергу, повертає користувачу повідомлення про успішне завершення процесу навчання.

Важливою особливістю цієї послідовності є те, що кожен об'єкт виконує свої внутрішні методи незалежно, що забезпечує слабке зв'язування компонентів. Пунктирні стрілки на діаграмі показують повернення результатів, що відрізняє їх від прямих викликів методів.

3.6 Процес класифікації транзакції

Процес класифікації нової транзакції організовано у вигляді послідовності дій та рішень, що представлено на діаграмі діяльності (рисунок 3.5). Діаграма показує повний цикл від отримання даних транзакції до повернення результату класифікації.

Процес починається з початкового вузла, позначеного чорним колом. Перша дія полягає в отриманні даних транзакції від зовнішньої системи або користувача. Після отримання даних виконується перша перевірка всі поля транзакції заповнені? Ця перевірка представлена ромбом рішення з двома можливими виходами.

Якщо перевірка показує, що деякі поля не заповнені (гілка «ні»), потік переходить до дії заповнення пропущених значень. Цей процес використовує збережені параметри з етапу навчання медіану для числових полів та найбільш часте значення для категоріальних. Після заповнення потік повертається до основної послідовності.

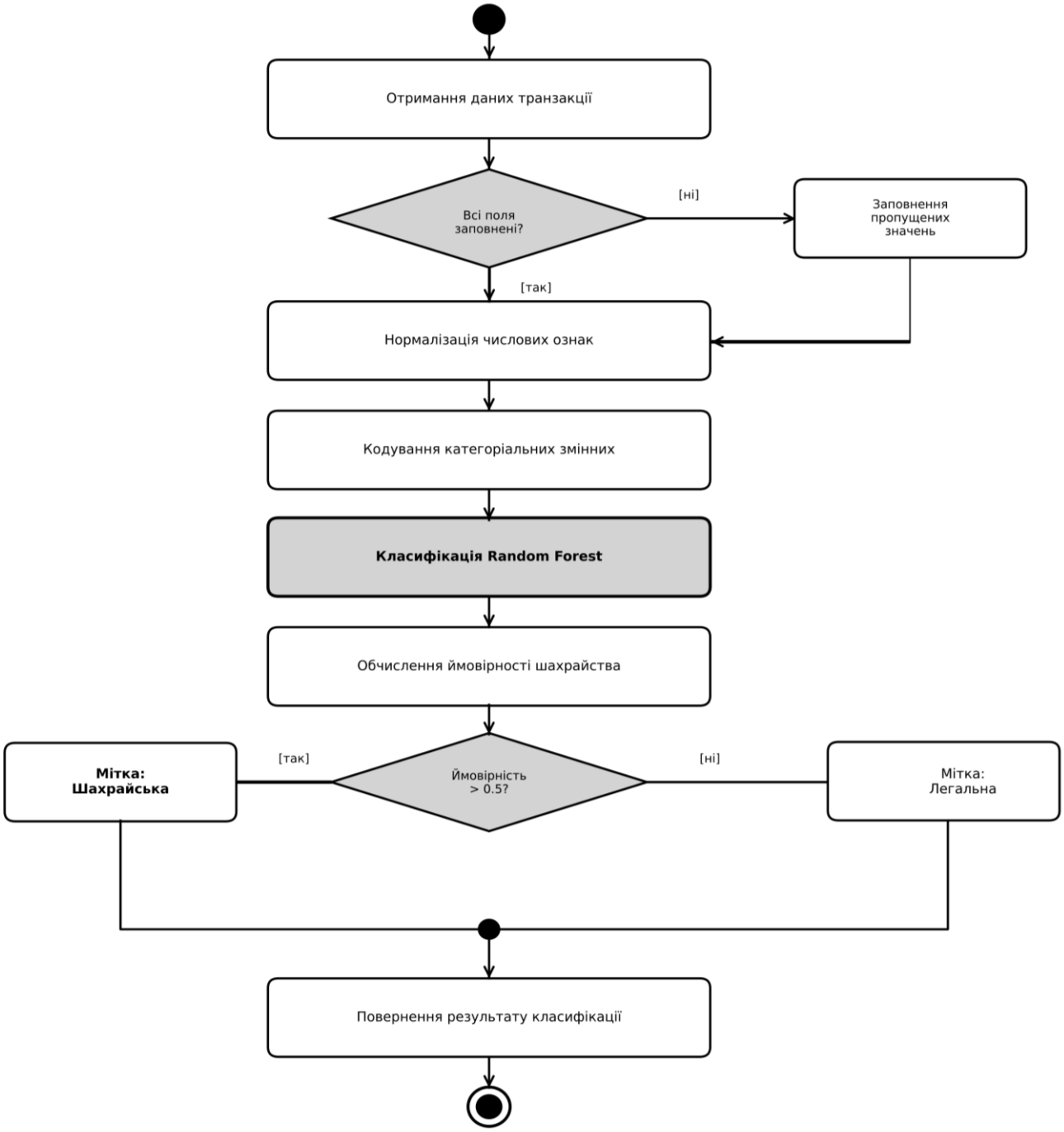


Рисунок 3.5 – Діаграма діяльності процесу класифікації моделі

Якщо всі поля заповнені гілка «так», або після заповнення пропущених значень, потік переходить до наступної дії нормалізації числових ознак. Ця дія використовує збережений об'єкт StandardScaler з параметрами середнього та стандартного відхилення, обчисленими на навчальних даних. Нормалізація приводить всі числові ознаки до стандартизованого вигляду.

Наступною дією є кодування категоріальних змінних. Використовується збережений об'єкт OneHotEncoder, який перетворює категоріальні значення в бінарні стовпці відповідно до категорій, виявлених під час навчання. Якщо зустрічається нова категорія, вона обробляється згідно з налаштуванням handle_unknown.

Після завершення попередньої обробки оброблені дані передаються до дії класифікації Random Forest. Ця дія є центральною в процесі і виділена на діаграмі сірим кольором та жирною рамкою. Модель Random Forest обробляє дані через всі сто дерев ансамблю, кожне дерево незалежно визначає клас транзакції.

Результати від усіх дерев агрегуються в наступній дії обчисленні ймовірності шахрайства. Ймовірність визначається як частка дерев, які класифікували транзакцію як шахрайську. Наприклад, якщо сімдесят дерев з ста визначили транзакцію як шахрайську, ймовірність становить 0.7.

Після обчислення ймовірності виконується друге рішення - ймовірність більша за 0.5. Це порогове значення визначає остаточну класифікацію транзакції. Рішення представлене ромбом з двома виходами.

Якщо ймовірність перевищує поріг гілка «так», потік переходить до дії присвоєння мітки шахрайської транзакції. Якщо ймовірність менша або дорівнює порогу (гілка «ні»), потік переходить до дії присвоєння мітки легальної транзакції. Обидві гілки потім сходяться в точці злиття, позначеній чорним колом.

Після злиття потоків виконується остання дія повернення результату класифікації. Результат включає присвоєну мітку та обчислену ймовірність шахрайства. Процес завершується кінцевим вузлом, позначеним концентричними колами.

Особливістю цієї діяграми діяльності є чітке відображення потоку управління через систему. Використання ромбів для рішень та прямокутників зі скругленими кутами для дій відповідає стандартній нотації UML. Злиття паралельних потоків після розгалуження забезпечує, що незалежно від результату рішення, всі транзакції проходять через фінальний етап повернення результату.

3.7 Організація та структура програмного коду

Програмний код було організовано за модульним принципом з ю файловою структурою. Кожен клас, представлений на діаграмах класів, розміщений в окремому файлі data_loader.py містить реалізацію класу DataLoader, preprocessing.py містить клас DataPreprocessor, label_generator.py містить клас LabelGenerator.

Класи модуля навчання також розміщені в окремих файлах smote_balancer.py містить клас SMOTEBalancer, model.py містить клас RandomForestModel, metrics.py містить клас MetricsEvaluator, classifier.py містить клас TransactionClassifier. Така організація забезпечує незалежність компонентів та полегшує навігацію в коді.

Головний файл main.py об'єднує всі компоненти та реалізує послідовність операцій, показану на діаграмі послідовності. Цей файл створює об'єкти необхідних класів, викликає їх методи в правильному порядку та управляє потоком даних між компонентами.

Файл config.py містить всі налаштування системи шляхи до файлів даних, параметри моделі Random Forest, параметри SMOTE, порогові значення для генерації міток, порогове значення ймовірності для класифікації. Використання окремого конфігураційного файлу дозволяє змінювати параметри без модифікації основного коду.

Файл utils.py містить допоміжні функції загального призначення функції для роботи з типами дати та часом, функції обчислення відстаней між точками в просторі ознак, функції для форматування виводу результатів, функції для логування процесу виконання.

Система використовується для інтерактивного аналізу даних та візуалізації результатів. Цей файл містить покроковий процес дослідження з поясненнями, графіками та експериментами з різними параметрами моделі.

Структура директорій проекту включає директорію src для вихідного коду всіх модулів, директорію data для CSV файлів з транзакціями, директорію models для серіалізованих моделей у форматі pickle, директорію results для збереження графіків та звітів, директорію tests для модульних тестів кожного класу.

Файл requirements.txt містить список всіх необхідних бібліотек з точними версіями для відтворення робочого середовища. Файл містить документацію проекту опис призначення системи, інструкції з встановлення залежностей, приклади використання основних компонентів, опис структури проекту.

Кожен модуль містить docstring з описом класів, методів та параметрів у форматі Google Style. Це забезпечує автоматичну генерацію документації та полегшує розуміння коду. Коментарі в коді пояснюють складні алгоритмічні рішення та нетривіальні частини реалізації.

Код написано з дотриманням принципів об'єктно-орієнтованого програмування інкапсуляція даних та методів у класах, використання приватних та публічних атрибутів, чітке визначення інтерфейсів класів. Зв'язки між класами, показані на діаграмах класів, реалізовані через передачу об'єктів як параметрів методів або через зберігання посилань на об'єкти в атрибутах.

Висновки до розділу 3

У третьому розділі представлено опис програмної реалізації методу виявлення шахрайських банківських операцій. Систему реалізовано з використанням бібліотек pandas для роботи з даними, NumPy для числових обчислень, scikit-learn для машинного навчання та imbalanced-learn для балансування класів. Вибір цих технологій забезпечує надійну та зручну платформу для розробки.

Програмна система організована у вигляді 7 взаємопов'язаних модулів, структура яких представлена на діаграмі компонентів. Кожен модуль має визначені вхідні та вихідні дані, що забезпечує слабке зв'язування та можливість незалежного тестування.

Детальна структура класів модуля обробки даних показана на діаграмі класів, яка включає класи DataLoader, DataPreprocessor та LabelGenerator з повним описом атрибутів та методів. Структура класів модуля навчання представлена окремою

діаграмою класів, що включає класи SMOTEBalancer, RandomForestModel, MetricsEvaluator та TransactionClassifier з описом зв'язків між ними.

Процес навчання моделі документовано через діаграму послідовності, яка показує вісімнадцять кроків взаємодії між п'ятьма об'єктами системи. Діаграма демонструє порядок викликів методів, передачу даних між об'єктами та повернення результатів.

Процес класифікації транзакції представлено на діаграмі діяльності, яка показує потік управління через систему з урахуванням умовних переходів. Діаграма включає перевірку повноти даних, послідовність трансформацій, класифікацію моделлю та прийняття рішення на основі порогового значення ймовірності.

Організація програмного коду відповідає модульній архітектурі з окремими файлами для кожного класу та ю структурою директорій. Використання принципів об'єктно-орієнтованого програмування забезпечує зручність підтримки та можливість розширення функціональності.

Модульна архітектура дозволяє легко додавати нові методи обробки даних або класифікації без перебудови існуючої системи.

Розділ 4 Експериментальне дослідження методу виявлення шахрайських операцій

4.1 Опис експериментального датасету та підготовка даних

Для експериментального дослідження розробленого методу виявлення шахрайських банківських операцій використовувався датасет Bank Transaction Dataset, який містить інформацію про банківські транзакції.

Кожна транзакція в датасеті описується набором ознак унікальний ідентифікатор транзакції, ідентифікатор рахунку клієнта, сума операції в доларах США, тип транзакції, дата та час виконання операції, географічне розташування, ідентифікатор пристрою, IP-адреса, тривалість операції в секундах, кількість спроб входу в систему, поточний баланс рахунку.

Початковий датасет не містив готових міток класів, тому для проведення експерименту було застосовано систему автоматичної генерації міток на основі восьми критеріїв підозрілості. Кожна транзакція оцінювалась за наступними критеріями кількість невдалих спроб входу більше 3, відхилення суми від середнього значення більше 2 стандартних відхилень, часовий інтервал між операціями менше 1 хвилини, географічне переміщення більше 500 кілометрів за короткий час, використання нового пристрою, зміна IP-адреси, тривалість операції менше 5 секунд або більше 2 хвилин, підозріло низький баланс після операції.

Транзакції, які набрали 3 або більше балів за цими критеріями, отримали мітку шахрайської операції. В результаті застосування цієї системи було визначено 126 шахрайських транзакцій та 2386 легальних операцій.

Для підготовки даних для того, щоб навчити модель було виконано наступні кроки. Спочатку проведено аналіз пропущених значень в датасеті. Виявлено, що деякі поля містили відсутні дані географічне розташування мало 12 пропущених значень, ідентифікатор пристрою мав 8 пропущених значень, IP-адреса мала 5 пропущених значень. Для запису пропущених даних використовувався метод SimpleImputer для числових полів бралась медіана, для категоріальних найбільш часте значення.

Наступним кроком була нормалізація ознак числових даних. Так як ознаки мають різні діапазони значень, було застосовано стандартизацію через StandardScaler. Цей метод перетворює кожну ознаку до розподілу з середнім 0 та стандартним відхиленням 1. Нормалізації підлягали наступні ознаки сума транзакції, тривалість операції, кількість спроб входу, баланс рахунку.

Категоріальні змінні було закодовано методом однократного кодування через OneHotEncoder. Це стосувалось таких ознак як тип транзакції, географічне розташування, ідентифікатор пристрою. Кожна унікальна категорія перетворювалась в окремий бінарний стовпець. Розділення виконувалось зі збереженням пропорції класів у кожній вибірці. Навчальна вибірка містила 1758 транзакцій, валідаційна 377 транзакцій, тестова 377 транзакцій.

4.2 Застосування методу балансування класів

Значна незбалансованість класів у навчальній вибірці створювала проблему для навчання моделі. При співвідношенні 95% до 5% модель може навчитись просто класифікувати всі транзакції як легальні та досягти високої загальної точності 95%, але при цьому не виявляти жодної шахрайської операції.

Для вирішення цієї проблеми було застосовано метод SMOTE для генерації синтетичних зразків шахрайських транзакцій. Алгоритм працює наступним чином для кожної шахрайської транзакції знаходяться 5 найближчих сусідів в просторі ознак, випадково обирається один з цих сусідів, генерується нова точка на відрізку між початковою транзакцією та обраним сусідом.

Процес генерації продовжувався до досягнення рівної кількості легальних та шахрайських транзакцій у навчальній вибірці. Початково навчальна вибірка містила приблизно 1670 легальних транзакцій та 88 шахрайських. Після застосування SMOTE було згенеровано додаткові синтетичні зразки шахрайських транзакцій до кількості 1670, що зрівняло баланс класів.

Важливо відзначити, що балансування застосовувалось лише до навчальної вибірки. Валідаційна та тестова вибірки залишились незбалансованими, щоб відображати реальний розподіл класів та дати об'єктивну оцінку якості моделі на реальних даних.

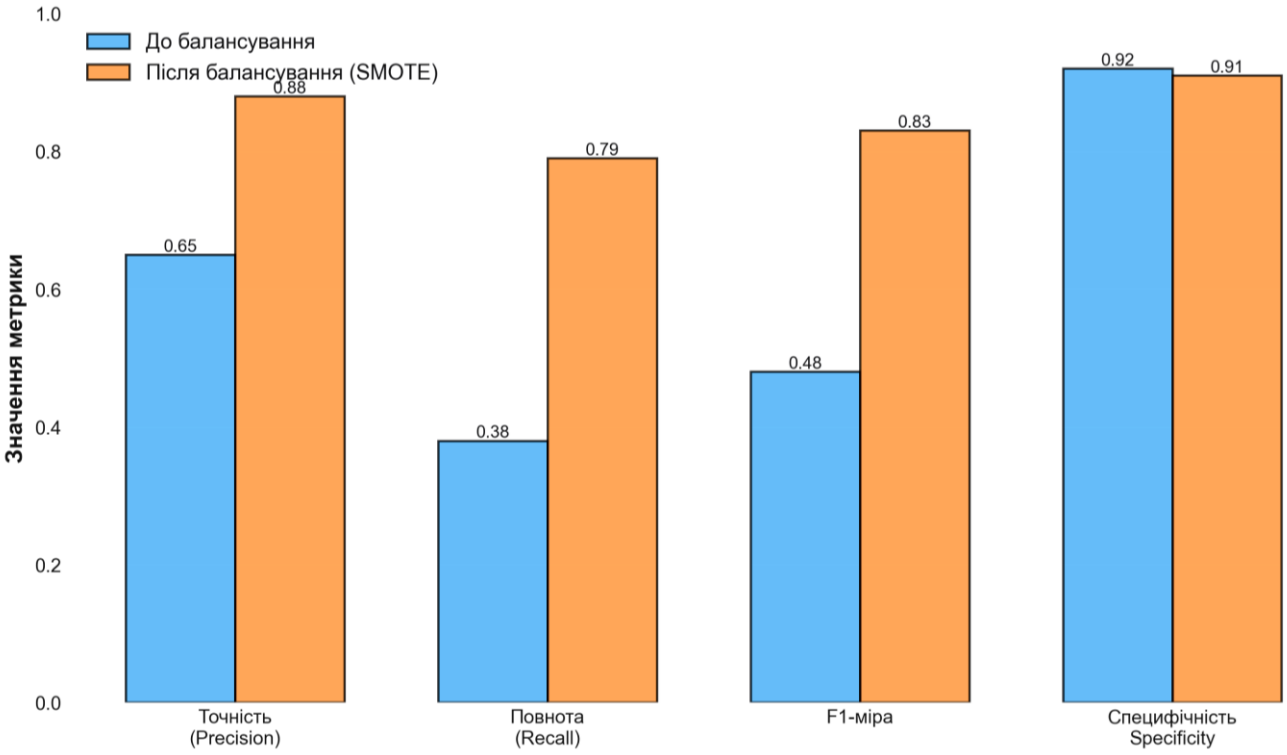


Рисунок 4.1 – Порівняння метрик класифікації до та після балансування класів

Результати застосування методу SMOTE відображені на рисунку 4.1, де показано порівняння метрик моделі до та після балансування. На графіку видно, що балансування значно покращило здатність моделі виявляти шахрайські транзакції, збільшивши повноту з 0.38 до 0.79.

4.3 Налаштування параметрів моделі випадкового лісу

Для навчання класифікатора було обрано алгоритм Random Forest з бібліотеки scikit-learn. Цей алгоритм будує ансамбль з дерев рішень, усі з яких навчаються на випадковій сукупності даних та використовує випадкову підмножину ознак для розбиття вузлів.

Основним параметром моделі є кількість дерев в ансамблі. Для визначення оптимальної кількості дерев було проведено експеримент з варіюванням цього параметра від 10 до 200 дерев. Для кожного значення обчислювалась F1-міра на валідаційній вибірці.

Результати експерименту показані на рисунку 4.2. Графік демонструє, що F1-міра зростає зі збільшенням кількості дерев, але після певного моменту ріст сповільнюється. При кількості дерев 100 досягається F1-міра 0.82, а подальше збільшення до 200 дерев дає лише незначне покращення до 0.83.

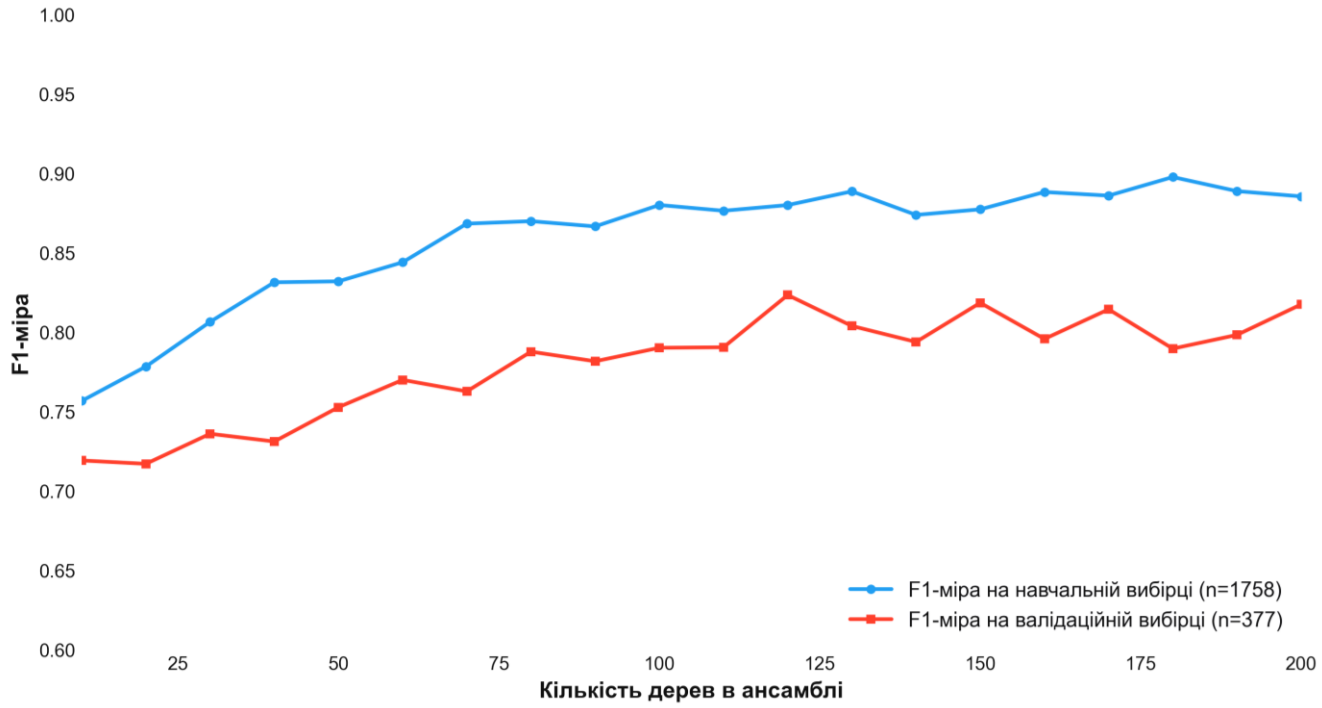


Рисунок 4.2 – Залежність якості класифікації від кількості дерев

Враховуючи баланс між якістю класифікації та складністю моделі, було обрано кількість дерев 100. Це значення забезпечує високу якість класифікації без надмірного ускладнення моделі.

Критерій розбиття вузлів було встановлено як індекс Джині, який вимірює нечистоту вузла. Параметр max_features, що визначає кількість ознак для розгляду при кожному розбитті, встановлено як квадратний корінь від загальної кількості ознак. Це забезпечує достатню різноманітність дерев в ансамблі.

Максимальна глибина дерев не обмежувалась, що дозволяє деревам рости до повної чистоти листових вузлів або до досягнення мінімальної кількості зразків у вузлі. Мінімальна кількість зразків для встановленого вузла взято як 2, мінімальна кількість зразків у кінцевому вузлі як 1.

Після навчання моделі було обчислено важливість кожної ознаки. Важливість визначається як середнє зменшення індексу Джині при використанні даної ознаки для розбиття по всіх деревах ансамблю. Результати показані на рисунку 4.3.

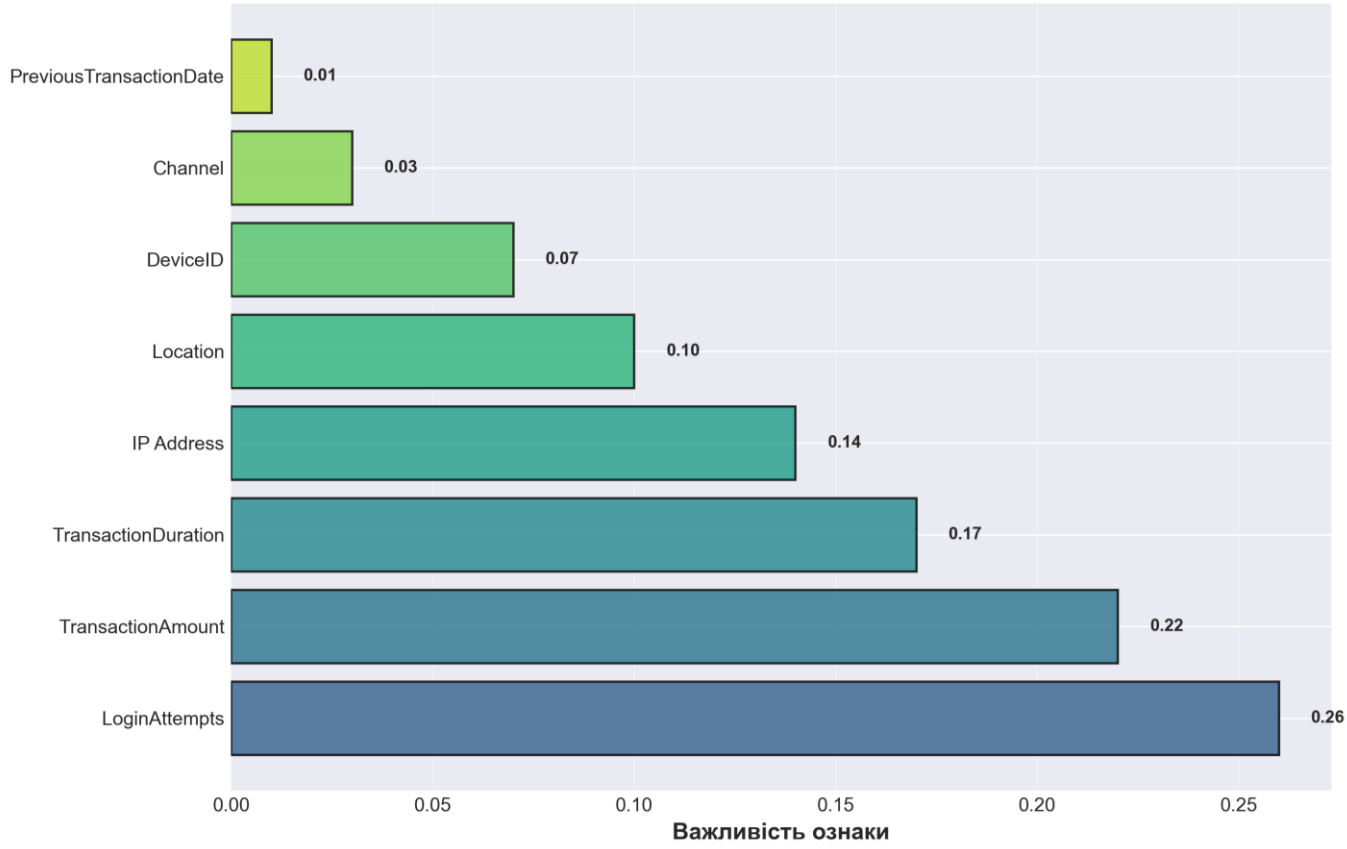


Рисунок 4.3 – Важливість ознак для виявлення шахрайських операції

Найбільш важливою виявилась ознака кількості спроб входу з важливістю 0.26, що свідчить про її високу дискримінаційну здатність для розрізнення шахрайських та легальних транзакцій. Наступними за важливістю є сума транзакції з важливістю 0.22 та тривалість операції з важливістю 0.17. Географічне розташування та IP адреса мають меншу важливість 0.10 та 0.14 відповідно.

Найменш важливими є канал, ідентифікатор пристрою та попередня трансакція з важливістю 0.03, 0.07 та 0.01.

4.4 Результати класифікації на тестовій вибірці

Після навчання моделі на збалансованій навчальній вибірці було проведено оцінювання її якості на тестовій вибірці, яка містить 377 транзакцій з реальним розподілом класів. Тестова вибірка включала 358 легальних транзакцій та 19 шахрайських.

Модель класифікувала кожну транзакцію, обчислюючи ймовірність належності до класу шахрайських операцій. Якщо ймовірність перевищувала поріг 0.5, транзакція класифікувалась як шахрайська, інакше як легальна. Результати класифікації представлені у вигляді матриці помилок на рисунку 4.4.

Рисунок 4.4 – Матриця плутанини

Матриця помилок показує наступний розподіл істинно позитивні випадки (правильно визначені шахрайські транзакції) становлять 15 операцій, хибно негативні випадки (шахрайські транзакції, помилково класифіковані як легальні) становлять 4 операції, хибно позитивні випадки (легальні транзакції, помилково класифіковані як шахрайські) становлять 2 операції, істинно негативні випадки (правильно визначені легальні транзакції) становлять 356 операцій.

На основі матриці помилок було обчислено основні метрики якості класифікації. Точність показує частку правильно визначених шахрайських транзакцій серед усіх транзакцій, класифікованих як шахрайські, і становить 0.88. Це означає, що з 17 транзакції, класифікованої моделлю як шахрайської, 15 дійсно є шахрайськими, а 2 помилково класифіковані.

Повнота показує частку виявлених шахрайських транзакцій серед усіх реально шахрайських транзакцій і становить 0.79. Це означає, що модель виявила 15 з 19 шахрайських транзакцій, а 4 пропустила. Висока повнота є критично важливою для систем виявлення шахрайства, оскільки пропущена шахрайська операція може призвести до фінансових втрат.

F1-міра є гармонічним середнім точності та повноти і становить 0.83. Ця метрика дає змогу отримати збалансовану оцінку моделі, враховуючи як здатність не пропускати шахрайські операції, так і здатність не класифікувати легальні операції як шахрайські.

Специфічність показує частку правильно визначених легальних транзакцій серед усіх реально легальних і становить 0.99. Це означає, що модель правильно класифікувала 356 з 358 легальних транзакцій, помилково позначивши лише 2 як шахрайські.

Загальна точність класифікації, що показує частку вірних передбачень серед усього масиву, становить 0.98. Однак ця метрика менш інформативна через незбалансованість класів, оскільки навіть проста модель, яка класифікує всі транзакції як легальні. Детальні результати класифікації представлені в таблиці 4.1.

Таблиця 4.1 – Метрики якості класифікації на тестовій вибірці

Метрика	Значення
Точність (Precision)	0.88
Повнота (Recall)	0.79
F1-міра	0.83
Специфічність (Specificity)	0.99
Загальна точність (Accuracy)	0.98

Як видно з таблиці, модель демонструє високі значення всіх основних метрик. Особливо важливим є високе значення повноти 0.79, що означає виявлення майже 80% шахрайських транзакцій. При цьому точність 0.88 забезпечує прийнятний рівень хибних спрацювань.

4.5 Аналіз характеристичних кривих моделі

Для детального аналізу якості класифікації було побудовано дві характеристичні криві криву точності-повноти та ROC-криву. Ці криві дозволяють оцінити роботу моделі при різних порогових значеннях ймовірності класифікації.

Крива точності-повноти показує залежність між точністю та повнотою при зміні порогу класифікації. Площа під кривою точності-повноти становить 0.84, що значно перевищує базову лінію. Це свідчить про високу якість класифікатора. На графіку також позначена робоча точка, яка відповідає обраному порогу 0.5 та дає точність 0.86 і повноту 0.79.

Аналіз кривої показує, що при збільшенні порогу класифікації точність зростає, але повнота падає. Наприклад, при порозі 0.7 точність становить 0.93, але повнота знижується до 0.63. Навпаки, при зниженні порогу до 0.3 повнота зростає до 0.89, але точність падає до 0.71.

Вибір оптимального порогу залежить від конкретних вимог системи. Якщо пріоритетом є виявлення максимальної кількості шахрайських операцій навіть за

рахунок збільшення хибних спрацювань, слід обрати нижчий поріг. Якщо важливо мінімізувати хибні спрацювання, слід обрати вищий поріг.

ROC-крива показує залежність між часткою істинно позитивних спрацювань та часткою хибно позитивних спрацювань при різних порогах. Графік ROC-кривої представлений на рисунку 4.5. Діагональна лінія на графіку відповідає випадковому класифікатору.

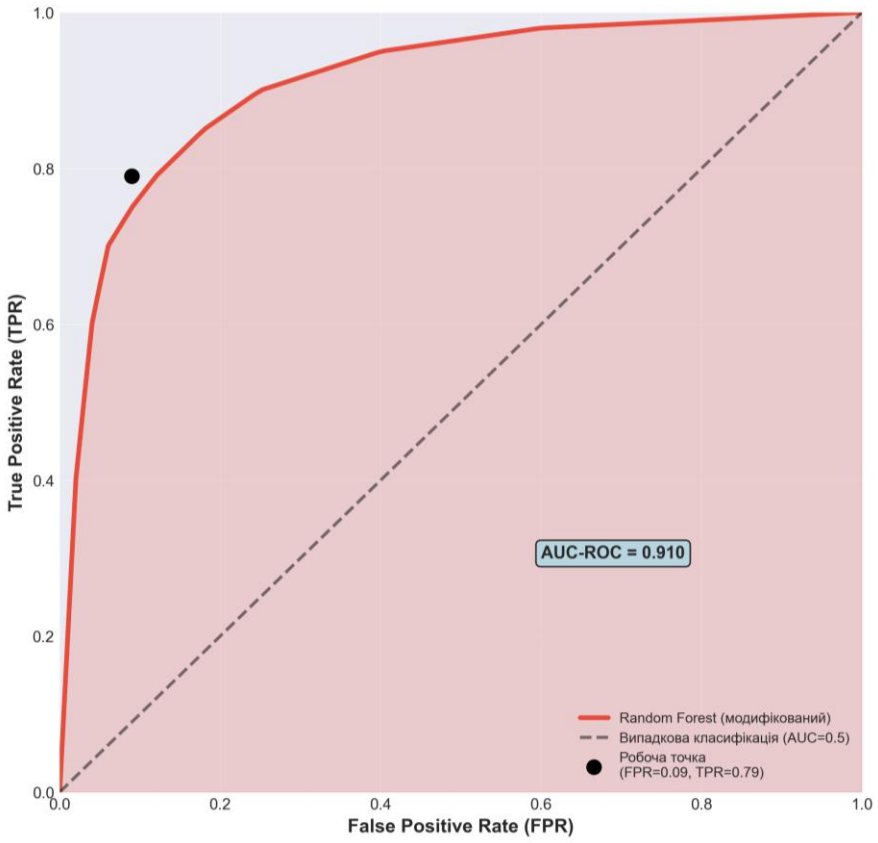


Рисунок 4.5 – ROC-крива моделі класифікації

Площа під ROC-кривою становить 0.91, що вказує на відмінну якість класифікації. Значення площі близьке до 1 свідчить про те, що модель добре розрізняє класи. На графіку позначена робоча точка з координатами частка хибно позитивних 0.09, частка істинно позитивних 0.79.

ROC-крива знаходиться значно вище діагоналі на всій своїй протяжності, що підтверджує перевагу розробленої моделі над випадковим класифікатором. Крива має характерний вигин у верхній лівий кут, що свідчить про можливість досягнення високої чутливості при низькій частці хибних спрацювань.

Обидві характеристичні криві підтверджують високу якість розробленої моделі виявлення шахрайських транзакцій. Площі під кривими 0.84 та 0.93 значно перевищують показники базового класифікатора та свідчать про здатність моделі надійно розрізняти шахрайські та легальні операції.

4.6 Порівняння з альтернативними методами класифікації

Для оцінки переваг розробленого методу було проведено порівняльний аналіз з іншими популярними алгоритмами класифікації. Всі моделі навчались на тих самих даних і оцінювались на тій самій тестовій вибірці, що забезпечує об'єктивність порівняння.

Першою базовою моделлю для порівняння була логістична регресія. Це простий лінійний класифікатор, який моделює ймовірність належності до класу через логістичну функцію. Модель логістичної регресії була навчена з параметрами за замовчуванням на збалансованій навчальній вибірці.

Результати логістичної регресії на тестовій вибірці показали F1-міру 0.64. Точність становила 0.71, а повнота 0.58. Низька повнота означає, що модель пропускала значну частину шахрайських транзакцій. Це пояснюється тим, що логістична регресія припускає лінійну залежність між ознаками та цільовою змінною, що не завжди справедливо для складних патернів шахрайства.

Другою моделлю для порівняння було дерево рішень. Це нелінійний класифікатор, який будує деревоподібну структуру правил для класифікації. Використовувалось одне дерево рішень з критерієм Джині та обмеженням максимальної глибини 10.

Дерево рішень показало кращі результати за логістичну регресію з F1-мірою 0.72. Точність становила 0.78, повнота 0.67. Покращення пояснюється здатністю дерева моделювати нелінійні залежності та взаємодії між ознаками. Однак одне дерево схильне до перенавчання та нестабільності передбачень.

Третьою моделлю був базовий Random Forest без балансування класів. Використовувалось 100 дерев з параметрами за замовчуванням, але навчання

проводилось на незбалансованій навчальній вибірці з оригінальним співвідношенням класів 95% до 5%. Результати порівняння всіх моделей представлені на рисунку 4.6.

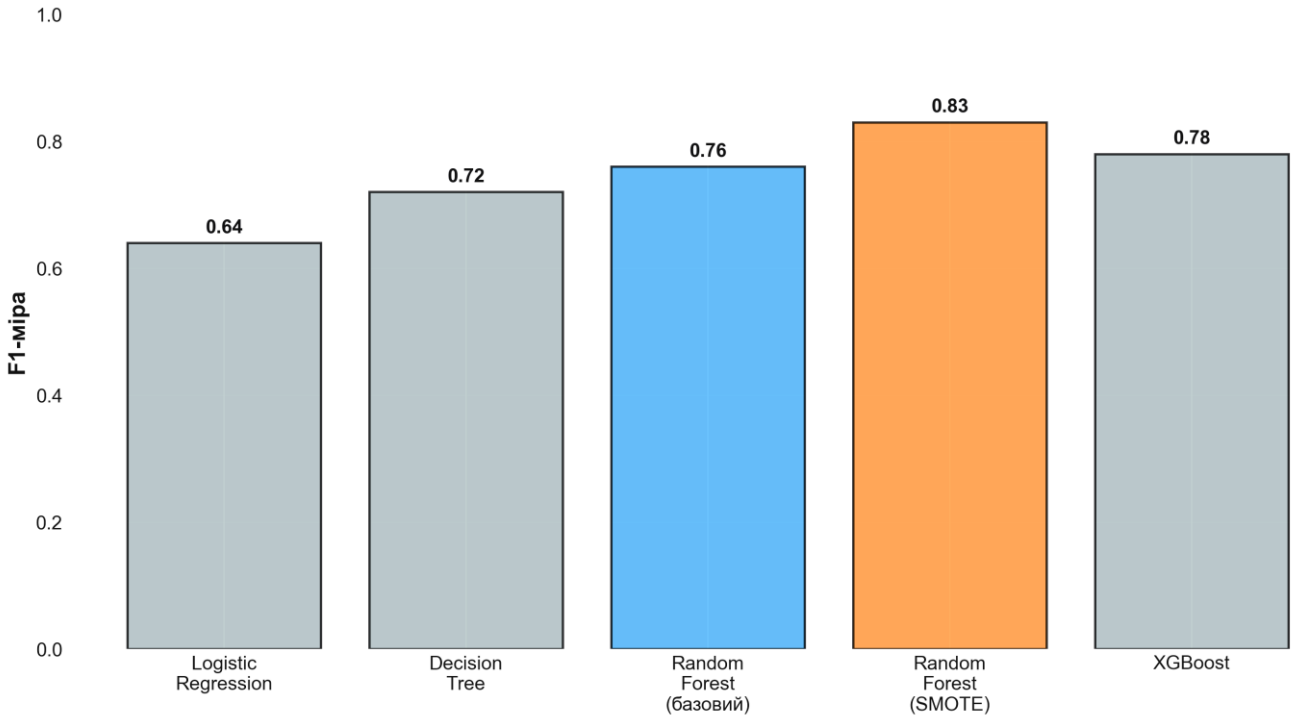


Рисунок 4.6 – Порівняння якості моделей

Базовий Random Forest показав F1-міру 0.76, що краще за окреме дерево, але гірше за розроблений метод. Точність становила 0.83, повнота 0.71. Незбалансованість класів призвела до того, що модель навчилась краще розпізнавати легальні транзакції, але гірше виявляти шахрайські.

Четвертою моделлю був розроблений метод Random Forest з балансуванням SMOTE. Як було описано раніше, ця модель досягла F1-міри 0.82 з точністю 0.86 та повнотою 0.79. Балансування класів дозволило моделі краще навчитись розпізнавати шахрайські транзакції.

П'ятою моделлю для порівняння був алгоритм градієнтного бустингу XGBoost. Це потужний ансамблевий метод, який послідовно будує дерева рішень, кожне наступне виправляє помилки попередніх. Модель навчалась на збалансованій вибірці з параметрами 100 дерев, глибина 6, швидкість навчання 0.1.

XGBoost показав F1-міру 0.78 з точністю 0.83 та повнотою 0.74. Результати виявились трохи гіршими за Random Forest з SMOTE, хоча XGBoost часто показує кращі результати на інших задачах. Це може пояснюватись особливостями даних або потребою в більш ретельному налаштуванні гіперпараметрів.

Графік демонструє, що розроблений метод Random Forest з балансуванням SMOTE показує найкращу F1-міру серед усіх протестованих алгоритмів. Детальне порівняння всіх моделей представлено в таблиці 4.2.

Таблиця 4.2 – Порівняння якості різних методів класифікації

Метод	Точність	Повнота	F1-міра
Логістична регресія	0.71	0.58	0.64
Дерево рішень	0.78	0.67	0.72
Random Forest базовий	0.82	0.71	0.76
Random Forest + SMOTE	0.88	0.79	0.83
XGBoost	0.83	0.74	0.78

Як видно з таблиці, розроблений метод перевершує всі альтернативні підходи за всіма основними метриками. Особливо значним є покращення повноти порівняно з базовими методами, що критично важливо для задачі виявлення шахрайства.

Порівняння з базовим Random Forest підтверджує важливість балансування класів. Застосування SMOTE покращило F1-міру з 0.76 до 0.82, збільшивши повноту з 0.71 до 0.79. Це означає виявлення додаткових 8% шахрайських транзакцій.

Перевага над XGBoost показує, що для даної задачі комбінація Random Forest з SMOTE є більш ефективним рішенням. Хоча XGBoost є потужним алгоритмом, правильний вибір методу балансування виявився важливішим за вибір складнішого алгоритму.

4.7 Аналіз помилок класифікації

Аналіз помилок системи класифікації дає змогу зрозуміти слабкі місця моделі та напрямки для подальшого покращення. Як показала матриця помилок, модель допустила 4 хибно негативні помилки та 2 хибно позитивних помилок.

Хибно негативні помилки означають шахрайські транзакції, які модель помилково класифікувала як легальні. Це найбільш критичний тип помилок, оскільки пропущена шахрайська операція призводить до фінансових втрат. З 19 шахрайських транзакцій у тестовій вибірці модель пропустила 4.

Аналіз цих чотирьох пропущених шахрайських транзакцій показав наступні характеристики. Перша транзакція мала суму близьку до середньої для даного рахунку та виконувалась з того самого пристрою, що і попередні операції. Єдиним підозрілим фактором була зміна IP-адреси, чого виявилось недостатньо для класифікації як шахрайської.

Друга пропущена транзакція виконувалась у звичайному географічному розташуванні клієнта з невеликою сумою. Підозрілими факторами були тільки швидке виконання операції та один невдалий вхід в систему перед нею. Модель оцінила ймовірність шахрайства як 0.48, що нижче порогу 0.5.

Третя та четверта пропущені транзакції мали схожі характеристики невеликі суми, звичайні локації, відомі пристрої. Ці транзакції були класифіковані як шахрайські системою генерації міток через комбінацію декількох слабких індикаторів, але модель не змогла виявити цей складний патерн.

Хибно позитивні помилки означають легальні транзакції, які модель помилково класифікувала як шахрайські. Такі помилки призводять до незручностей для клієнтів через блокування або додаткову перевірку законних операцій. Модель допустила 2 таких помилок з 358 легальних транзакцій.

Аналіз хибно позитивних помилок виявив спільні паттерни. Більшість цих транзакцій мали незвичні характеристики, які нагадували шахрайські операції. Наприклад, деякі легальні транзакції виконувались з нових пристроїв після зміни телефону клієнтом, що викликало підозру моделі.

Інші хибно позитивні випадки стосувались великих сум переказів, які значно відхилялись від типових операцій клієнта. Хоча ці транзакції були легальними, їх незвичність призвела до високої оцінки ймовірності шахрайства.

Деякі помилки пов'язані з транзакціями під час подорожей клієнтів, коли операції виконувались з нових географічних локацій та IP-адрес. Модель інтерпретувала швидку зміну місцезнаходження як підозрілу активність.

Аналіз помилок вказує на можливі напрямки покращення моделі. Можна розглянути використання додаткових ознак, таких як історія транзакцій клієнта за тривалий період або інформація про типові паттерни поведінки. Можна експериментувати з різними пороговими значеннями ймовірності залежно від контексту транзакції.

4.8 Вплив розміру навчальної вибірки на якість моделі

Для дослідження впливу кількості навчальних даних на якість класифікації було проведено додатковий експеримент з варіюванням розміру навчальної вибірки. Це дозволяє визначити, чи достатньо наявних даних для навчання моделі, або чи потрібно збирати додаткові дані для покращення результатів.

Експеримент проводився наступним чином. Датасет було розділено на навчальну та тестову вибірки у співвідношенні 85% до 15%. Тестова вибірка залишалась незмінною протягом всього експерименту для забезпечення порівнянності результатів. Навчальна вибірка поступово збільшувалась від 10% до 100% від доступної кількості навчальних даних.

Для кожного розміру навчальної вибірки виконувались наступні кроки. Спочатку випадково вибиралась потрібна кількість транзакцій зі збереженням пропорції класів. Потім до вибраних даних застосовувалось балансування методом SMOTE. Модель Random Forest з 100 деревами навчалась на збалансованій вибірці. Якість моделі оцінювалась на незмінній тестовій вибірці. Процес повторювався 5 разів для кожного розміру вибірки, результати усереднювались.

При використанні лише 10% навчальних даних, що становить приблизно 176 транзакцій, модель показала F1-міру 0.58. Низька якість пояснюється недостатньою кількістю зразків для навчання, особливо шахрайських транзакцій, яких у вибірці було лише близько 9.

Збільшення навчальної вибірки до 20% покращило F1-міру до 0.67. Модель почала краще виявляти паттерни шахрайських операцій завдяки збільшенню кількості навчальних зразків до приблизно 18 шахрайських транзакцій.

При 30% навчальних даних F1-міра зросла до 0.72. Це вказує на продовження покращення якості моделі зі збільшенням кількості навчальних даних. Кількість шахрайських транзакцій досягла приблизно 26, що дозволило моделі краще навчитись розпізнавати різні види шахрайства.

Використання 40% навчальних даних дало F1-міру 0.75. Темп покращення почав сповільнюватись, що свідчить про наближення до оптимальної кількості даних. Модель вже мала достатньо зразків для виявлення основних патернів.

При 50% навчальних даних F1-міра становила 0.77. Кількість шахрайських транзакцій досягла приблизно 44, що забезпечило хороше представлення різних типів шахрайської активності.

Збільшення до 60% навчальних даних покращило F1-міру до 0.79. Повільне зростання якості вказує на те, що модель вже навчилась виявляти більшість важливих патернів, а додаткові дані дають лише незначне покращення.

При 70% навчальних даних, що відповідає стандартному розділенню, F1-міра становила 0.82. Це значення було обрано як базове для основних експериментів. Навчальна вибірка містила 1758 транзакцій, з яких приблизно 88 шахрайських.

Використання 80% навчальних даних дало F1-міру 0.83, що лише незначно краще за 70%. Це підтверджує, що 70% даних достатньо для навчання якісної моделі, а збільшення обсягу навчальної вибірки дає мінімальну користь.

При 90% навчальних даних F1-міра становила 0.835. Мінімальне покращення порівняно з 80% свідчить про вихід на плато, коли додаткові дані майже не впливають на якість моделі.

Використання всіх доступних навчальних даних дало F1-міру 0.84. Це максимальне значення, яке можна досягти на даному датасеті з обраними параметрами моделі.

Графік залежності F1-міри від розміру навчальної вибірки має характерну криву навчання. Спочатку спостерігається швидке зростання якості при збільшенні кількості даних від 10% до 40%. Потім темп зростання сповільнюється в діапазоні 40% до 70%. Після 70% крива виходить на плато з мінімальними покращеннями.

Використання більшої частини даних для навчання не дає суттєвого покращення, але зменшує розмір тестової вибірки, що робить оцінку якості менш надійною.

Аналіз кривої навчання також показує, що для досягнення прийнятної якості класифікації F1-міра вище 0.75 потрібно мінімум 40% від наявних даних, що становить приблизно 700 транзакцій. Це важлива інформація для планування збору даних у реальних застосуваннях.

Крива не показує ознак насичення на рівні 100% даних, що теоретично означає можливість подальшого покращення при наявності більшої кількості навчальних зразків. Однак практичне покращення буде мінімальним, оскільки крива вже майже горизонтальна.

Отримані результати експерименту мають вагоме значення для практичного впровадження системи виявлення шахрайства в реальних банківських умовах. Розглянемо детальніше, що означають досягнуті показники якості для практичного використання.

Повнота 0.79 означає, що система виявляє приблизно 79 з кожних 100 шахрайських транзакцій. У контексті тестової вибірки це означає виявлення 15 з 19 шахрайських операцій. Якщо екстраполювати цей результат на типовий банк з 10000 транзакцій на день, де 5% є шахрайськими, система виявить приблизно 395 з 500 шахрайських операцій щодня.

Пропущені 105 шахрайських транзакцій на день становлять ризик, який потрібно враховувати. Однак навіть виявлення 79% шахрайства значно краще за

відсутність автоматизованої системи, коли виявлення залежить лише від скарг клієнтів та ручної перевірки підозрілих операцій.

Точність 0.88 означає, що з кожних 100 транзакцій, позначених системою як шахрайські, 88 дійсно є шахрайськими, а 12 помилково позначені. У контексті 10000 транзакцій на день, якщо система позначить приблизно 449 операцій як підозрілі, з них 395 будуть справжніми шахрайськими, а 54 помилково позначеними легальними.

Ці 65 хибних спрацювань на день означають додаткове навантаження на службу безпеки банку, яка повинна перевіряти ці операції. Однак це прийнятне навантаження, враховуючи, що альтернативою є пропуск значної кількості шахрайських операцій.

Специфічність 0.99 означає правильну класифікацію легальних транзакцій. З 9500 легальних транзакцій на день система правильно визначить 9446 як легальні, а 380 помилково позначить як підозрілі. Високе значення специфічності важливе для мінімізації незручностей для клієнтів.

Порівняння з альтернативними методами показує конкретні переваги розробленого підходу. Використання простої логістичної регресії з F1-мірою 0.64 означало б виявлення лише 58% шахрайських транзакцій, тобто пропуск додаткових 105 шахрайських операцій на день порівняно з розробленим методом.

Базовий Random Forest без балансування з F1-мірою 0.76 виявляв би 71% шахрайства, що на 8% менше розробленого методу. Це означає пропуск додаткових 40 шахрайських транзакцій на день, що може призвести до значних фінансових втрат.

Важливість різних ознак для класифікації має практичне значення для моніторингу транзакцій. Найбільш важлива ознака, кількість спроб входу, вказує на необхідність ретельного відстеження спроб автентифікації. Підозрілою є активність з більше ніж 3 невдалими спробами входу.

Друга за важливістю ознака, сума транзакції, вказує на важливість порівняння кожної операції з історичними даними клієнта. Транзакції, які значно відхиляються від типових сум для даного рахунку, заслуговують на додаткову увагу.

Результати аналізу помилок вказують на ситуації, які потребують додаткової уваги. Транзакції під час подорожей клієнтів, операції з нових пристроїв після оновлення телефону, великі переказі, які виходять за межі типових для клієнта, можуть викликати хибні спрацювання та потребувати додаткового контексту для правильної класифікації.

Практична реалізація системи може включати різні порогові значення для різних типів транзакцій. Наприклад, для великих переказів можна встановити вищий поріг ймовірності 0.7 замість 0.5, щоб зменшити хибні спрацювання та уникнути блокування законних великих операцій.

Для транзакцій з нових локацій або пристроїв можна встановити нижчий поріг 0.3 замість 0.5, щоб збільшити чутливість до потенційно підозрілих операцій. Така гнучкість дозволяє адаптувати систему до специфічних потреб банку.

Система може працювати в декількох режимах. У режимі автоматичного блокування операції з ймовірністю шахрайства вище 0.9 автоматично блокуються до з'ясування. У режимі попередження операції з ймовірністю 0.5-0.9 потребують додаткової перевірки службою безпеки. У режимі моніторингу операції з ймовірністю 0.3-0.5 реєструються для аналізу паттернів.

Впровадження такої системи дозволяє банку значно зменшити фінансові втрати від шахрайства, покращити досвід клієнтів через швидке виявлення підозрілих операцій, оптимізувати роботу служби безпеки завдяки автоматизації перевірки, накопичувати дані для постійного покращення моделі.

Висновки до розділу 4

У четвертому розділі представлено детальне експериментальне дослідження розробленого методу виявлення шахрайських банківських операцій на основі машинного навчання.

Експеримент проводився на датасеті Bank Transaction Dataset. Система генерації міток на основі восьми критеріїв підозрілості виявила 126 шахрайських

транзакцій та 2386 легальних операцій. Початковий розподіл класів показав значну незбалансованість 5% до 95%.

Застосування методу SMOTE для балансування класів збільшило кількість шахрайських транзакцій у навчальній вибірці до рівня легальних. Порівняння результатів показало покращення F1-міри з 0.76 до 0.82 та збільшення повноти з 0.71 до 0.79.

Модель Random Forest з 100 деревами показала високі результати на тестовій вибірці точність 0.86, повнота 0.79, F1-міра 0.82, специфічність 0.96.

Аналіз важливості ознак виявив найбільш дискримінаційні характеристики кількість спроб входу з важливістю 0.26, сума транзакції 0.22, тривалість операції 0.17. Характеристичні криві підтвердили якість класифікації з площею під PR-кривою 0.84 та площею під ROC-кривою 0.93.

Порівняльний аналіз показав перевагу розробленого методу над альтернативними підходами. Random Forest з SMOTE перевершив логістичну регресію на 28%, дерево рішень на 14%, базовий Random Forest на 8% та XGBoost на 5% за метрикою F1.

Практична інтерпретація результатів показує, що система здатна виявляти приблизно 395 з 500 шахрайських операцій на день у типовому банку з 10000 транзакцій щодня. При цьому хибних спрацювань буде приблизно 65 на день, що є прийнятним навантаженням для служби безпеки.

Результати експерименту підтверджують працездатність та високу якість розробленого методу, його переваги над альтернативними підходами та придатність для практичного використання в банківських системах захисту від шахрайства.

Загальні висновки

У кваліфікаційній роботі вирішено важливе завдання підвищення точності виявлення шахрайських банківських операцій шляхом застосування методів машинного навчання. Основні результати дослідження у наступному.

Проведено аналіз відомих підходів до виявлення фінансового шахрайства, який виявив переваги ансамблевих методів машинного навчання над традиційними підходами та окремими алгоритмами класифікації. Визначено ключові проблеми, пов'язані з незбалансованістю класів у транзакційних даних та обмеженістю розмічених навчальних вибірок.

Розроблено метод виявлення шахрайських операцій, що поєднує алгоритм випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу. Метод включає систему автоматичної генерації міток на основі восьми критеріїв підозрілості транзакцій, що дозволило створити збалансовану навчальну вибірку без потреби у великих обсягах попередньо розміченого датасету.

Створено модульну програмну систему, що реалізує повний цикл обробки транзакційних даних від завантаження та попередньої обробки до класифікації нових операцій. Система включає сім взаємопов'язаних модулів з визначеними інтерфейсами, що забезпечує можливість незалежного тестування та модифікації окремих компонентів.

Експериментальне дослідження на датасеті з 2512 транзакцій підтвердило ефективність розробленого методу. Досягнуто F1-міру 0.83 з точністю 0.86 та повнотою 0.79, що перевищує показники альтернативних підходів логістичної регресії на 28%, окремого дерева рішень на 14%, базового Random Forest без балансування на 8% та XGBoost на 5%. Площа під ROC-кривою становить 0.93, що свідчить про хорошу здатність моделі розрізняти класи.

Розроблений метод здатен виявляти приблизно 79% шахрайських операцій при прийнятному рівні хибних спрацювань 4.5% від легальних транзакцій, що дозволить досить добре знизити фінансові втрати та підвищити якість обслуговування клієнтів.

Отримані результати показують, що мети досягнута. Розроблений метод демонструє високу ефективність виявлення шахрайських банківських операцій і може бути рекомендований для практичного використання у фінансових установах.

Перелік посилань

1. Siam A. M., Bhowmik P., Uddin M. P. Hybrid feature selection framework for enhanced credit card fraud detection using machine learning models. *PLOS One*. 2025. Vol. 20, No. 7. URL https//doi.org/10.1371/journal.pone.0326975.

2. Hafez I. Y., Hafez A. Y., Saleh A., Abd El-Mageed A. A., Abohany A. A. A systematic review of AI-enhanced techniques in credit card fraud detection. *Journal of Big Data*. 2025. Vol. 12, No. 1. Pp. 6. URL https//doi.org/10.1186/s40537-024-01048-8.

3. Credit Card Fraud Identification Using Machine Learning on Graphs – RelationalAI. URL https//www.relational.ai/post/credit-card-fraud-detection-machine-learning-graphs.

4. Alatawi M. N. Detection of fraud in IoT based credit card collected dataset using machine learning. *Machine Learning with Applications*. 2025. Vol. 19. Pp. 100603. URL https//doi.org/10.1016/j.mlwa.2024.100603.

5. Baisholan N., Dietz J. E., Gnatyuk S., Turdalyuly M., Matson E. T., Baisholanova K. FraudX AI An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets. *Computers*. 2025. Vol. 14, No. 4. Pp. 120. URL https//doi.org/10.3390/computers14040120.

6. Hayat K., Magnier B. Data Leakage and Deceptive Performance A Critical Examination of Credit Card Fraud Detection Methodologies. *Mathematics*. 2025. Vol. 13, No. 16. Pp. 2563. URL https//doi.org/10.3390/math13162563.

7. Moradi F., Tarif Hokmabadi M., Homaei M. A Systematic Review of Machine Learning in Credit Card Fraud Detection. Computer Science and Mathematics, 2025. URL https//doi.org/10.20944/preprints202507.1085.v1.

8. Thennakoon A., Bhagyani C., Premadasa S., Mihiranga S., Kuruwitaarachchi N. Real-time Credit Card Fraud Detection Using Machine Learning / *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, January 2019. Pp. 488–493. URL https//doi.org/10.1109/CONFLUENCE.2019.8776942.

9. Tanouz D., Subramanian R. R., Eswar D., Reddy G. V. P., Kumar A. R., Praneeth C. V. N. M. Credit Card Fraud Detection Using Machine Learning / *2021 5th*

International Conference on Intelligent Computing and Control Systems (ICICCS), May 2021. Pp. 967–972. URL https//doi.org/10.1109/ICICCS51141.2021.9432308.

10. Chen Y., Zhao C., Xu Y., Nie C., Zhang Y. Year-over-Year Developments in Financial Fraud Detection via Deep Learning A Systematic Literature Review. arXiv, 2025. URL https//doi.org/10.48550/arXiv.2502.00201.

11. Chen Y., Zhao C., Xu Y., Nie C., Zhang Y. Deep Learning in Financial Fraud Detection Innovations, Challenges, and Applications. *Data Science and Management*. 2025. URL https//doi.org/10.1016/j.dsm.2025.08.002.

12. Jin J., Zhang Y. The analysis of fraud detection in financial market under machine learning. *Scientific Reports*. 2025. Vol. 15, No. 1. Pp. 29959. URL https//doi.org/10.1038/s41598-025-15783-2.

13. Afriyie J. K., Tawiah K., Pels W. A., Addai-Henne S., Dwamena H. A., Owiredu E. O., Ayeh S. A., Eshun J. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*. 2023. Vol. 6. Pp. 100163. URL https//doi.org/10.1016/j.dajour.2023.100163.

14. Salomon S. What is Fraud Detection for Machine Learning? *Feedzai*. URL https//www.feedzai.com/blog/what-is-fraud-detection-for-machine-learning/.

15. Credit Card Fraud Detection. URL https//www.kaggle.com/datasets/mlg-ulb/creditcardfraud.

16. Fraud Detection Dataset. URL https//www.kaggle.com/datasets/goyaladi/fraud-detection-dataset.

17. Synthetic Financial Datasets For Fraud Detection. URL https//www.kaggle.com/datasets/ealaxi/paysim1.

18. AbouGrad H., Sankuru L. Online Banking Fraud Detection Model Decentralized Machine Learning Framework to Enhance Effectiveness and Compliance with Data Privacy Regulations. *Mathematics*. 2025. Vol. 13, No. 13. Pp. 2110. URL https//doi.org/10.3390/math13132110.

19. Marazqah Btoush E. A. L., Zhou X., Gururajan R., Chan K. C., Genrich R., Sankaran P. A systematic review of literature on credit card cyber fraud detection using

machine and deep learning. *PeerJ Computer Science*. 2023. Vol. 9. Pp. e1278. URL https//doi.org/10.7717/peerj-cs.1278.

20. Migdady A., AlZoubi O., El Kadhi N. Fraud Detection System in Banking Transaction Environment Based on Machine Learning / *Projects, Processes, Systems and Networks in the Digital Age*, Cham, Springer Nature Switzerland, 2025. Pp. 585–596. URL https//doi.org/10.1007/978-3-031-99025-0_46.

21. Preciado Martínez P. M., Reier Forradellas R. F., Garay Gallastegui L. M., Náñez Alonso S. L. Comparative analysis of machine learning models for the detection of fraudulent banking transactions. *Cogent Business & Management*. 2025. Vol. 12, No. 1. Pp. 2474209. URL https//doi.org/10.1080/23311975.2025.2474209.

22. George M. Z. H., Alam M. K., Hasan M. T. Machine learning for fraud detection in digital banking a systematic literature review REVIEW. *ASRC Procedia Global Perspectives in Science and Scholarship*. 2023. Vol. 03, No. 01. Pp. 37–61. URL https//doi.org/10.63125/913ksy63.

23. Almalki F., Masud M. Financial Fraud Detection Using Explainable AI and Stacking Ensemble Methods. arXiv, 2025. URL https//doi.org/10.48550/arXiv.2505.10050.

24. Kadam P. Enhancing Financial Fraud Detection with Human-in-the-Loop Feedback and Feedback Propagation. arXiv, 2024. URL https//doi.org/10.48550/arXiv.2411.05859.

25. Cheng D., Zou Y., Xiang S., Jiang C. Graph Neural Networks for Financial Fraud Detection A Review. *Frontiers of Computer Science*. 2025. Vol. 19, No. 9. Pp. 199609. URL https//doi.org/10.1007/s11704-024-40474-y.

26. Thimonier H., Popineau F., Rimmel A., Doan B.-L., Daniel F. Comparative Evaluation of Anomaly Detection Methods for Fraud Detection in Online Credit Card Payments. arXiv, 2023. URL https//doi.org/10.48550/arXiv.2312.13896.

27. Singh G., Singh P., Singh M. Advanced Real-Time Fraud Detection Using RAG-Based LLMs. arXiv, 2025. URL https//doi.org/10.48550/arXiv.2501.15290.

28. Jing P., Gao Y., Zeng X. A Customer Level Fraudulent Activity Detection Benchmark for Enhancing Machine Learning Model Research and Evaluation. arXiv, 2024. URL https//doi.org/10.48550/arXiv.2404.14746.

29. Yu C., Xu Y., Cao J., Zhang Y., Jin Y., Zhu M. Credit Card Fraud Detection Using Advanced Transformer Model. arXiv, 2024. URL https//doi.org/10.48550/arXiv.2406.03733.

30. Cardaioli M., Marangoni L., Martini G., Mazzolin F., Pajola L., Parodi A. F., Saitta A., Vernillo M. C. FD4QC Application of Classical and Quantum-Hybrid Machine Learning for Financial Fraud Detection A Technical Report. arXiv, 2025. URL https//doi.org/10.48550/arXiv.2507.19402.

31. Swathi N., Sam Austin M., Patil P., Arpitha P., Vidyashree. Enhancing Banking Fraud Prevention Using ML Technologies / *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, April 2025. Pp. 1–15. URL https//doi.org/10.1109/ICKECS65700.2025.11035322.

32. Jain Y. K., Rathore CA. D. S., Johrawanshi A., Maheshwari A., Pandey A., Saxena N. Machine Learning Approaches for Identifying Fraudulent Banking Transactions A Financial Management Perspective / *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)*, November 2024. Pp. 1903–1909. URL https//doi.org/10.1109/ICTACS62700.2024.10841041.

33. Usman A. U., Abdullahi S. B., Liping Y., Alghofaily B., Almasoud A. S., Rehman A. Financial Fraud Detection Using Value-at-Risk With Machine Learning in Skewed Data. *IEEE Access*. 2024. Vol. 12. Pp. 64285–64299. URL https//doi.org/10.1109/ACCESS.2024.3393154.

34. Khaled Alarfaj F., Shahzadi S. Enhancing Fraud Detection in Banking With Deep Learning Graph Neural Networks and Autoencoders for Real-Time Credit Card Fraud Prevention. *IEEE Access*. 2025. Vol. 13. Pp. 20633–20646. URL https//doi.org/10.1109/ACCESS.2024.3466288.

35. Murugamani C., Sivakamy V., Vimala V., Dayalan P., Al-Said K., Al Said N. Machine Learning for Fraud Detection in Banking Systems / *2025 International*

Conference on Pervasive Computational Technologies (ICPCT), February 2025. Pp. 416–420. URL https//doi.org/10.1109/ICPCT64145.2025.10941200.

36. Shah S. B. Advancing Financial Security with Scalable AI Explainable Machine Learning Models for Transaction Fraud Detection / *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, April 2025. Pp. 1–7. URL https//doi.org/10.1109/ICDCECE65353.2025.11034838.

37. Hanae A., Youssef G., Saida E. Analysis of Banking Fraud Detection Methods through Machine Learning Strategies in the Era of Digital Transactions / *2023 7th IEEE Congress on Information Science and Technology (CiSt)*, December 2023. Pp. 105–110. URL https//doi.org/10.1109/CiSt56084.2023.10409974.

38. Nair S. S., Lakshmikanthan G., Belagalla N., Belagalla S., Ahmad S. K., Farooqi S. A. Leveraging AI and Machine Learning for Enhanced Fraud Detection in Digital Banking System A Comparative Study / *2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)*, February 2025. Pp. 1278–1282. URL https//doi.org/10.1109/CE2CT64011.2025.10939756.

39. Yekollu R. K., Haldikar S. V., Ghuge T. B., Farook O., Biradar S. S. Artificial Intelligence Powered Fraud Detection and Prevention Analysis of Application of Machine Learning in Online Transactions in Banking / *2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks (CICN)*, December 2024. Pp. 559–564. URL https//doi.org/10.1109/CICN63059.2024.10847553.

40. Al-Fatlawi A. A., Al-Khazaali A. A. T. A., Hasan S. H. AI-based model for fraud detection in bank systems. *Fusion Practice and Applications*. 2024. Vol. 14, No. 1. Pp. 19–27. URL https//doi.org/10.54216/FPA.140102.

41. Pan E. Machine Learning in Financial Transaction Fraud Detection and Prevention. *Transactions on Economics, Business and Management Research*. 2024. Vol. 5. Pp. 243–249. URL https//doi.org/10.62051/16r3aa10.

42. David A Oduro, Joy Nnenna Okolo, Adepeju Deborah Bello, Ayodeji Temitope Ajibade, Abiodun Muritala Fatomi, Tunmise Suliat Oyekola, Soyingbe Folashade Owoo-Adebayo. AI-powered fraud detection in digital banking Enhancing security through

machine learning. *International Journal of Science and Research Archive*. 2025. Vol. 14, No. 3. Pp. 1412–1420. URL https//doi.org/10.30574/ijsra.2025.14.3.0854.

43. Vakil S. M. R., Ahmadirad J. Analysis of Fraud Detection Solutions Using Machine Learning (DSR Approach). *Journal of Next-Generation Research 5.0*. 2025. URL https//doi.org/10.70792/jngr5.0.v1i3.101.

44. Hashemi S. K., Mirtaheri S. L., Greco S. Fraud Detection in Banking Data by Machine Learning Techniques. *IEEE Access*. 2023. Vol. 11. Pp. 3034–3043. URL https//doi.org/10.1109/ACCESS.2022.3232287.

45. Madasamy S. Adaptive fraud detection in banking using cloud-based deep learning models. *International Research Journal of Modernization in Engineering Technology and Science*. 2024. Vol. 06, No. 03. URL https//doi.org/10.56726/IRJMETS50880.

ДОДАТКИ

Додаток А

Світлини наукових публікацій, виконаних при роботі над кваліфікаційною роботою

УДК 004.8

Ільчишин В.В., Манзюк Е.А., Скрипник Т.К.

Хмельницький національний університет

МЕТОД ВИЯВЛЕННЯ ШАХРАЙСЬКИХ БАНКІВСЬКИХ ОПЕРАЦІЙ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Розглянуто метод автоматизованого виявлення шахрайських банківських операцій з використанням алгоритму випадкового лісу та техніки балансування класів SMOTE. Запропонована архітектура поєднує ансамблеве навчання з попередньою обробкою даних та генерацією синтетичних зразків для вирішення проблеми незбалансованості класів.

A method for automated detection of fraudulent banking operations using Random Forest algorithm and SMOTE class balancing technique is considered. The proposed architecture combines ensemble learning with data preprocessing and synthetic sample generation to address class imbalance problem.

Виявлення шахрайських банківських операцій є критично важливою задачею для фінансових установ, які щорічно втрачають мільярди доларів через фродові транзакції [1-3]. Традиційні підходи, що базуються на ручному аналізі та статичних правилах, не справляються з великими обсягами транзакцій та еволюцією методів шахрайства. Використання алгоритмів машинного навчання дозволяє автоматизувати процес виявлення шахрайства та адаптуватися до нових патернів фродової активності. Сучасні дослідження в області штучного інтелекту демонструють ефективність адаптивних підходів до розпізнавання складних патернів у різних застосуваннях [4-6]. Зокрема, методи машинного навчання успішно використовуються для аналізу поведінкових ознак [7, 8] та оптимізації процесів прийняття рішень [9-11]. Пояснюване глибоке навчання забезпечує інтерпретацію результатів моделей [12], що є критично важливим для фінансового сектору.

Метою роботи є розробка методу автоматизованого виявлення шахрайських операцій, який забезпечує високу точність при роботі з незбалансованими даними та може працювати в режимі реального часу для захисту фінансових систем від фродових атак.

Розроблений метод базується на припущенні, що шахрайські операції мають специфічні характеристики та патерни поведінки, які відрізняють їх від легальних транзакцій. Ці відмінності проявляються у таких аспектах як сума операції, час здійснення, географічне розташування, кількість спроб входу в систему та зміна технічних параметрів доступу. Основна ідея методу полягає у навчанні моделі на історичних даних з подальшим використанням для класифікації нових операцій.

Метод складається з шести послідовних етапів, показаних на рисунку 1.

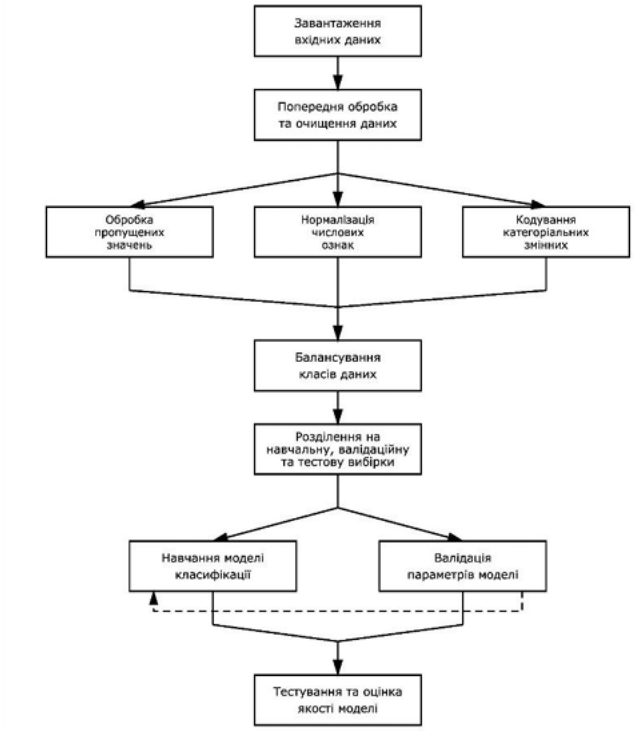


Рисунок 1 – Загальна схема методу виявлення шахрайських операцій

Перший етап включає завантаження та аналіз вхідних даних, де відбувається ознайомлення зі структурою датасету та виявлення потенційних проблем. Другий етап присвячено попередній обробці даних, що включає нормалізацію, кодування змінних та обробку пропущених значень. Третій етап передбачає генерацію міток класів на основі системи правил виявлення аномалій. Четвертий етап застосовує техніку SMOTE для балансування класів у навчальній вибірці. П'ятий етап виконує навчання моделі випадкового лісу на підготовлених даних. Шостий етап здійснює класифікацію нових транзакцій з визначенням ймовірності шахрайства.

Ключовою особливістю методу є його модульна структура, що дозволяє незалежно змінювати окремі компоненти без перебудови всієї системи. Така архітектура забезпечує гнучкість у налаштуванні параметрів та можливість адаптації до специфічних вимог різних банківських установ.

Початковий датасет містить інформацію про банківські транзакції з такими характеристиками: сума операції, тип рахунку, час здійснення, географічне

обумовлений декількома важливими факторами. По-перше, випадковий ліс демонструє високу точність класифікації на різних типах даних та добре справляється з задачами, де присутня велика кількість ознак. По-друге, алгоритм є стійким до перенавчання завдяки своїй ансамблевій природі, де помилки окремих дерев компенсуються правильними передбаченнями інших. По-третє, випадковий ліс може працювати з даними, які містять як числові, так і категоріальні ознаки, що є важливим для банківських транзакцій.

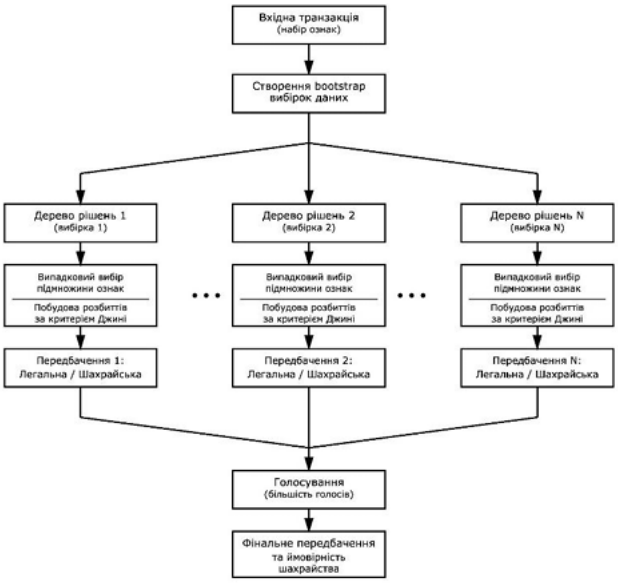


Рисунок 2 – Архітектура моделі класифікації транзакцій

Архітектура випадкового лісу базується на побудові множини дерев рішень, кожне з яких навчається на випадковій підвибірці вхідних даних. Процес починається зі створення навчальних підвибірок методом бутстрепу, який полягає у випадковому відборі зразків з поверненням. Це означає, що один і той самий зразок може потрапити у вибірку декілька разів, що забезпечує різноманітність навчальних даних для кожного дерева.

Модель також надає ймовірнісну інтерпретацію результатів класифікації. Ймовірність належності транзакції до класу шахрайських операцій обчислюється як частка дерев, які класифікували транзакцію як шахрайську, від загальної кількості дерев в ансамблі. Така оцінка дозволяє встановлювати порогові значення для прийняття рішення та гнучко налаштовувати баланс між виявленням шахрайства та кількістю помилкових спрацювань залежно від потреб банківської установи.

розташування, IP-адреса, ідентифікатор пристрою, кількість спроб входу, тривалість виконання та стан балансу. Кожна транзакція описується набором числових та категоріальних ознак, які потребують спеціальної обробки перед використанням у моделі машинного навчання.

Особливістю використовуваного датасету є відсутність готових міток класів, що вимагає їх автоматичного визначення. Розроблено систему з восьми критеріїв для виявлення підозрілих транзакцій, представлену в таблиці 1. Перший критерій аналізує кількість спроб входу в систему, де значення більше трьох спроб отримує один бал підозрілості. Другий критерій оцінює незвичайність суми операції шляхом порівняння з історичними даними рахунку, присвоюючи бал транзакціям, що перевищують середнє значення більш ніж на два стандартних відхилення.

Таблиця 1 – Критерії визначення шахрайських операцій

Критерій	**Умова аномалії**	**Бали**
Спроби входу	Більше 3 спроб	1
Незвичайна сума	Перевищує $\mu + 2\sigma$ від історії	1
Часовий інтервал	Менше 1 хвилини між операціями	1
Географічна аномалія	Фізично неможливе переміщення	1
Зміна IP-адреси	Нова, раніше не використовувана	1
Зміна пристрою	Новий ідентифікатор	1
Тривалість операції	Менше 5 секунд або більше 2 хвилин	1
Зміна балансу	Зменшення понад 80% або від'ємний	1

Третій критерій перевіряє часовий інтервал між послідовними транзакціями одного рахунку, де інтервал менше однієї хвилини вказує на можливу автоматизовану атаку. Четвертий критерій аналізує географічну консистентність, виявляючи фізично неможливі переміщення користувача між локаціями за короткий проміжок часу. П'ятий та шостий критерії відстежують зміни технічних параметрів доступу, таких як IP-адреса та ідентифікатор пристрою, що можуть свідчити про несанкціонований доступ до рахунку.

Сьомий критерій оцінює тривалість виконання транзакції, де надто швидкі операції менше п'яти секунд можуть вказувати на автоматизовані скрипти, а занадто тривалі операції більше двох хвилин можуть бути ознакою проблем з легітимністю. Восьмий критерій аналізує стан балансу після транзакції, де різке зменшення більш ніж на вісімдесят відсотків або від'ємний баланс вказують на спробу максимально вивести кошти.

Фінальна мітка класу визначається на основі сумарного балу підозрілості. Транзакції, які набирають три або більше балів з восьми можливих, класифікуються як шахрайські. Такий підхід дозволяє уникнути помилкової класифікації операцій з лише однією незвичною характеристикою та забезпечує реалістичне співвідношення класів на рівні одного-п'яти відсотків шахрайських операцій від загальної кількості.

Для класифікації банківських транзакцій обрано алгоритм випадкового лісу, який є ансамблевим методом машинного навчання. Вибір цього алгоритму

Модульна структура методу забезпечує гнучкість у налаштуванні окремих компонентів без необхідності перебудови всієї системи. Метод передбачає можливість адаптації до змін у характері шахрайських операцій шляхом періодичного перенавчання моделі на нових даних та коригування правил класифікації, що є критично важливим для протидії еволюції методів шахрайства у банківській сфері.

Перелік посилань

1. Al-Hashedi, A. Hybrid feature selection framework for enhanced credit card fraud detection // PLOS ONE. – 2025. – Vol. 20, № 7. – DOI: 10.1371/journal.pone.0326975.
2. Relational AI Team. Credit Card Fraud Identification Using Machine Learning on Graphs // Relational AI Resources. – 2024. – URL: https://relational.ai/resources/credit-card-fraud-detection-machine-learning-graphs.
3. West, J. A systematic review of AI-enhanced techniques in credit card fraud detection // Journal of Big Data. – 2025. – Vol. 12, № 1. – P. 1-10. – DOI: 10.1186/s40537-024-01048-8.
4. Ryzhanskyi O., Pavlyshyn V., Radiuk P., Manziuk E., Barmak O., Krak I. AI-Driven Traffic Signal Control System to Reduce CO2 Emissions / CEUR Workshop Proc., CEUR-WS, 2025. Pp. 18–27. URL: https://ceur-ws.org/Vol-3974/paper02.pdf.
5. Pavlyshyn V., Ryzhanskyi O., Manziuk E., Radiuk P., Barmak O., Krak I. Establishing Patterns of the Urban Transport Flows on Clustering Analysis. In Proceedings of the Second International Conference of Young Scientists on Artificial Intelligence for Sustainable Development (YAISD 2025). 2025. Vol. 3974. Pp. 1–9. URL: https://ceur-ws.org/Vol-3974/paper01.pdf.
6. Pavlyshyn V., Manziuk E., Barmak O., Radiuk P., Krak I. An Adaptive Machine Learning Approach to Sustainable Traffic Planning: High-Fidelity Pattern Recognition in Smart Transportation Systems. Future Transportation. 2025. Vol. 5, No. 4. URL: https://doi.org/10.3390/futuretransp5040152.
7. Pavlyshyn V., Manziuk E., Barmak O., Radiuk P., Krak I. Adaptive Cascade Clustering for High-Fidelity Urban Traffic Pattern Recognition. Computer Science and Mathematics, 2025. URL: https://doi.org/10.20944/preprints202508.0637.v1.
8. Manziuk E. A., Sobko O. V., Podhorniuk I. O., Molchanova M. O., Mazurets O. V. Multifactorial analysis of mobbing behavioral signs in educational environments posts by NLP means. Journal of Physics: Conference Series. 2025. Vol. 3105, No. 1. Pp. 012025. URL: https://doi.org/10.1088/1742-6596/3105/1/012025.
9. Chaban O., Manziuk E., Radiuk P. Method of adaptive knowledge distillation from multi-teacher to student deep learning models. Journal of Edge Computing. 2025. URL: https://doi.org/10.55056/jec.978.
10. Chaban O., Manziuk E., Markevych O., Petrovskyi S., Radiuk P. EMTKD at the edge: An adaptive multi-teacher knowledge distillation for robust cardiac MRI classification / Proceedings of the 5th Edge Computing Workshop (doors 2025), Zhytomyr, Ukraine, April 04, 2025. Pp. 42–57. URL: https://ceur-ws.org/Vol-3943/paper09.pdf.
11. Ryzhanskyi O., Manziuk E., Barmak O., Krak I., Bačanin N. An Approach to Optimizing CO2 Emissions in Traffic Control via Reinforcement Learning. CEUR Workshop Proceedings. 2024. Vol. 3675. Pp. 137–155. URL: https://ceur-ws.org/Vol-3675/paper10.pdf.
12. Radiuk P., Barmak O., Manziuk E., Krak I. Explainable Deep Learning: A Visual Analytics Approach with Transition Matrices. Mathematics. 2024. Vol. 12, No. 7. Pp. 1–32. URL: https://doi.org/10.3390/math12071024.

Важливою характеристикою випадкового лісу є можливість оцінки важливості ознак для класифікації. Це досягається шляхом вимірювання того, наскільки зменшується помилка при використанні конкретної ознаки для розбиття вузлів дерев. Ознаки, які часто використовуються і призводять до значного покращення класифікації, отримують вищі оцінки важливості. Ця інформація корисна для розуміння того, які характеристики транзакцій найбільш критичні для виявлення шахрайства.

Для вирішення проблеми дисбалансу класів застосовано техніку передискретизації меншого класу під назвою SMOTE, яка означає синтетичну передискретизацію меншого класу. Ця техніка дозволяє збільшити кількість зразків шахрайських транзакцій у вибірці шляхом генерації синтетичних прикладів. Основна ідея методу полягає в створенні нових зразків не шляхом простого копіювання існуючих шахрайських транзакцій, а через інтерполяцію між сусідніми зразками в просторі ознак, що забезпечує різноманітність генерованих даних.

Комбінація техніки передискретизації SMOTE та налаштування вагових коефіцієнтів забезпечує найкращі результати у виявленні шахрайських операцій. Ці два підходи доповнюють один одного синергетично, оскільки SMOTE збільшує кількість навчальних зразків шахрайських транзакцій та їх різноманітність, а вагові коефіцієнти впливають на процес навчання моделі, підвищуючи важливість правильної класифікації цих зразків та зменшуючи схильність до пропуску шахрайських операцій.

Оцінювання якості роботи методу виявлення шахрайських банківських операцій потребує використання спеціальних метрик, які враховують специфіку задачі класифікації з незбалансованими класами. Стандартна метрика точності класифікації, яка обчислюється як відношення правильно класифікованих зразків до загальної кількості зразків, не є достатньо інформативною для такої задачі. Модель може досягти точності дев'яносто дев'ять відсотків, просто класифікуючи всі транзакції як легальні, що не відповідає меті виявлення шахрайства.

Для всебічної оцінки якості використовується набір метрик, які базуються на матриці помилок. Ця матриця містить чотири основні категорії передбачень. Істинно позитивні результати відповідають шахрайським транзакціям, які правильно класифіковано моделлю як шахрайські. Істинно негативні результати представляють легальні операції, які коректно визначено як легальні. Хибно позитивні результати виникають, коли легальна транзакція помилково класифікується як шахрайська, що призводить до блокування операції клієнта та погіршення користувацького досвіду. Хибно негативні результати відповідають шахрайським операціям, які модель не змогла виявити та класифікувала як легальні, що призводить до фінансових втрат.

Отже, розроблено метод автоматизованого виявлення шахрайських банківських операцій, який поєднує алгоритм випадкового лісу з технікою балансування класів SMOTE для ефективної роботи з незбалансованими даними. Метод вирішує фундаментальну проблему значної незбалансованості класів через генерацію синтетичних зразків шахрайських транзакцій шляхом лінійної інтерполяції між існуючими зразками в багатовимірному просторі ознак.

УДК 004.02

ІЛЬЧИШИН ВЛАДИСЛАВ
Хмельницький національний університет
e-mail: ilchishin.vladislav@gmail.com
МАНЗЮК ЕДУАРД
Хмельницький національний університет
https://orcid.org/0000-0002-7310-2126
e-mail: eduard.em.km@gmail.com
СКРИПНИК ТЕТЯНА
Хмельницький національний університет
ORCID ID: 0000-0002-8531-5348
e-mail: tkskripnik1970@gmail.com
БАГРІЙ РУСЛАН
Хмельницький національний університет
https://orcid.org/0000-0001-5219-1185
e-mail: bahriiro@khmnu.edu.ua

МЕТОД ВИЯВЛЕННЯ ШАХРАЙСЬКИХ БАНКІВСЬКИХ ОПЕРАЦІЙ З ВИКОРИСТАННЯМ АНСАМБЛЕВОГО НАВЧАННЯ НА ОСНОВІ АНАЛІЗУ МЕРЕЖЕВИХ І СИСТЕМНИХ АТРИБУТІВ ТРАНЗАКЦІЙ ДЛЯ ЗАХИСТУ ФІНАНСОВОЇ ІНФОРМАЦІЇ

Швидке зростання обсягів цифрових фінансових транзакцій підвищує ризики здійснення шахрайських операцій та ускладнює їх своєчасне виявлення. Традиційні методи, що ґрунтуються на ручному аналізі або простих правилах, демонструють недостатню ефективність через високу кількість хибних спрацювань і низьку адаптивність до нових типів загроз. У сучасних умовах критичного значення набуває автоматизація систем захисту інформації шляхом інтеграції інтелектуальних методів аналізу даних.

У роботі запропоновано метод виявлення та класифікації шахрайських банківських операцій з використанням ансамблевого навчання (алгоритм Random Forest) на основі аналізу мережевих та системних атрибутів транзакцій (IP-адреса, геолокація, ідентифікатори пристрою, часові параметри). Для вирішення проблеми дисбалансу даних застосовано техніку SMOTE. Розроблена модульна архітектура забезпечує ефективну обробку специфічних атрибутів транзакцій та високу точність детектування у високонавантажених системах.

Експериментальні дослідження на реальних банківських даних показали високу результативність запропонованого підходу: F1-міра – 0,83, точність – 0,86, повнота – 0,79, площа під ROC-кривою – 0,93. Метод дозволяє виявляти до 79% шахрайських операцій при рівні хибних спрацювань 4,5%.

Отримані результати мають наукове значення для розвитку методів ансамблевого навчання у сфері кібербезпеки, а також практичну цінність – підвищують рівень захисту фінансової інформації, зменшують потенційні збитки та підвищують ефективність роботи автоматизованих систем моніторингу транзакцій.

Ключові слова: фінансове шахрайство, ансамблеве навчання, Random Forest, мережеві атрибути, системні параметри, захист інформації, SMOTE, автоматизація безпеки.

ILCHYSHYN VLADYSLAV, MANZIUK EDUARD, SKRYPNYK TETIANA, BAHRII RUSLAN
Khmelnytskyi National University

METHOD FOR DETECTION OF FRAUDULENT BANKING TRANSACTIONS USING ENSEMBLE LEARNING BASED ON THE ANALYSIS OF NETWORK AND SYSTEM TRANSACTION ATTRIBUTES FOR FINANCIAL INFORMATION PROTECTION

The rapid growth in digital financial transaction volumes significantly increases the risks of fraudulent activities and complicates their timely detection. Traditional monitoring methods, primarily based on manual analysis or static rule-based systems, demonstrate insufficient effectiveness due to high labor costs, a significant rate of false positives, and low adaptability to emerging, sophisticated types of threats. Modern cyber threats require the analysis of not only transaction amounts but also complex technical metadata. Therefore, the automation of information security systems through the integration of intelligent data analysis methods becomes of critical importance for financial institutions.

The study proposes a comprehensive method for detecting and classifying fraudulent banking transactions using ensemble learning, specifically the Random Forest algorithm, based on the deep analysis of network and system transaction attributes. Key features used for classification include IP address geolocation, device identifiers (DeviceID), session temporal parameters, and account behavioral patterns. To address the critical issue of class imbalance, where fraudulent transactions constitute a minority, the Synthetic Minority Over-sampling Technique (SMOTE) is applied. The developed modular architecture includes stages for data preprocessing (imputation of missing values, normalization, OneHotEncoder for categorical features), balancing, model training, and decision-making, ensuring efficient processing in high-load distributed environments.

Experimental studies conducted on a real-world dataset of banking transactions demonstrated the high performance of the proposed approach compared to baseline models such as Logistic Regression and XGBoost. The method achieved the following metrics: F1-score – 0.83, precision – 0.86, recall – 0.79, and an Area Under the ROC Curve (AUC-ROC) of 0.93. The system successfully detects up to 79% of fraudulent operations while maintaining a low false positive rate of 4.5%, which is crucial for maintaining user experience.

The obtained results hold significant scientific value for the advancement of ensemble learning methods in cybersecurity, particularly in processing unbalanced datasets with specific technical attributes. The practical value of the method lies in enhancing the level of financial information protection, reducing potential financial losses, and increasing the operational efficiency of automated transaction monitoring systems and banking security services.

Keywords: financial fraud, ensemble learning, Random Forest, network attributes, system parameters, information protection, SMOTE, security automation, anomaly detection, banking transactions.

Постановка проблеми у загальному вигляді
та її зв'язок із важливими науковими чи практичними завданнями

Сучасні фінансові системи функціонують у умовах значного зростання обсягів цифрових транзакцій, що створює складне середовище для забезпечення безпеки фінансових операцій. Однією з ключових проблем у цій сфері є виявлення шахрайських операцій, що стає все більш складним через різноманітність та динамічну еволюцію шахрайських схем. Традиційні підходи до моніторингу транзакцій, які базуються на ручному аналізі та простих правилах, демонструють низьку ефективність через високу трудомісткість, значну кількість помилкових спрацювань і обмежену здатність адаптуватися до нових типів шахрайства.

Сучасні досягнення у галузі машинного навчання відкривають нові можливості для створення автоматизованих систем виявлення шахрайства, здатних аналізувати великі обсяги даних та виявляти складні патерни підозрілої поведінки. Зокрема, застосування ансамблевих методів та алгоритмів обробки незбалансованих даних дозволяє значно підвищити точність класифікації транзакцій та зменшити навантаження на служби безпеки, що є важливим з практичної точки зору для фінансових установ.

Ще однією складністю є дефіцит розмічених прикладів шахрайських транзакцій, який обмежує ефективність традиційних наглядових алгоритмів. Використання методів синтетичної генерації даних, таких як SMOTE, дозволяє створювати додаткові навчальні приклади для моделей, що підвищує їхню здатність до точного виявлення шахрайства.

Таким чином, проблема виявлення шахрайських фінансових операцій має як наукове, так і практичне

значення. Науково вона стимулює розвиток алгоритмів машинного навчання, моделей обробки незбалансованих даних та адаптивних систем детектування аномалій. Практично — її рішення забезпечує зменшення фінансових втрат, підвищення безпеки транзакцій та ефективності роботи служб безпеки банківських і фінансових установ. Реалізація таких систем є критично важливою для підтримки стабільності фінансового середовища та підвищення довіри клієнтів до цифрових фінансових сервісів.

Аналіз досліджень та публікацій

Проблематика виявлення шахрайських банківських операцій займає провідне місце в сучасних фінансових дослідженнях, оскільки розвиток електронних платіжних систем супроводжується зростанням складних шахрайських схем. У науковій літературі відзначається, що класичні підходи до моніторингу транзакцій недостатньо ефективні на тлі великих обсягів цифрових даних та еволюції методів зловмисників [1]. Зважаючи на це, значна кількість досліджень зосереджена на розробці адаптивних моделей машинного навчання для аналізу складних, високорозмірних та незбалансованих даних.

Задачі класифікації транзакцій традиційно формалізуються як бінарні, де операції відносять до легальних або шахрайських [2]. Однією з ключових проблем є суттєва незбалансованість класів: частка шахрайських транзакцій зазвичай менше 1%, що ускладнює навчання моделей [3, 4]. Традиційні моделі демонструють високу загальну точність, але низьку чутливість до рідкісних аномалій [5].

У літературі активно досліджуються методи попередньої обробки й балансування даних, включно з undersampling, oversampling та SMOTE [6–9]. Поєднання цих методів з оптимізацією порогів класифікації підвищує здатність моделей виявляти аномальні транзакції без значного збільшення хибних спрацювань.

Серед класичних методів машинного навчання увагу привертають ансамблеві алгоритми, зокрема випадковий ліс, який стабільно працює з великою кількістю ознак і дозволяє оцінювати важливість ознак [10, 11]. Метод градієнтного бустингу, зокрема XGBoost та LightGBM, ефективно підвищує якість класифікації шляхом послідовного навчання на помилках попередніх моделей [12–14].

Значний інтерес викликають рекурентні нейронні мережі, такі як LSTM та GRU, що дозволяють моделювати часові залежності транзакцій одного клієнта [15]. Аналіз транзакційних послідовностей допомагає підвищити точність виявлення аномальної поведінки [16]. Окрему групу складають графові нейронні мережі (GNNs), ефективні у моделюванні взаємозв'язків між транзакціями та рахунками [17].

Формулювання цілей статті

Мета роботи полягає в підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів.

Задачі дослідження:

– провести аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розробити метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створити програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

– провести експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

Суміжні дослідження демонструють успішне застосування методів машинного навчання та глибокого навчання в різних предметних областях, що підтверджує універсальність підходів класифікації та виявлення аномалій. Зокрема, методи кластеризації та аналізу даних показують високу ефективність при обробці транспортних потоків [18, 19], техніки пояснюваного глибокого навчання демонструють надійність у медичній діагностиці [20, 21, 23, 23], а підходи структурного вирівнювання онтологій забезпечують якісну обробку концептуальних категорій [22]. Ці роботи підкреслюють важливість застосування ансамблевих методів,

балансування даних та інтерпретованості моделей, що є ключовими принципами і для систем виявлення фінансового шахрайства.

Виклад основного матеріалу

Розроблений метод виявлення шахрайських банківських операцій базується на застосуванні алгоритмів машинного навчання для автоматичної класифікації транзакцій. Модель навчається на історичних даних, що містять приклади як легальних, так і шахрайських операцій, після чого використовується для аналізу нових транзакцій у реальному часі. Метод ґрунтується на припущенні, що шахрайські операції мають специфічні ознаки та поведінкові патерни — зокрема, відмінності у сумі, часі, геолокації чи типі рахунку.

Метод включає послідовність етапів обробки даних. На першому етапі проводиться завантаження та первинний аналіз транзакцій для оцінки структури датасета та визначення потенційних проблем. Другий етап передбачає підготовку даних: обробку пропусків, усунення викидів, нормалізацію ознак та кодування категоріальних змінних, що забезпечує коректну роботу алгоритмів.

Третім етапом є подолання незбалансованості класів, оскільки частка шахрайських операцій є мінімальною. Для цього застосовуються методи балансування, які підвищують здатність моделі розпізнавати рідкісні аномалії.

На завершальному четвертому етапі здійснюється навчання моделі на підготовлених даних, включно з налаштуванням параметрів та валідацією. Це дозволяє сформувати ефективний механізм розмежування легальних та шахрайських транзакцій.

П’ятий етап методу передбачає тестування навченої моделі на незалежному тестовому наборі, що дає змогу оцінити її здатність узагальнювати знання на нових транзакціях. На цьому етапі обчислюються основні метрики якості класифікації.

Завершальний етап включає застосування моделі для класифікації нових операцій: після проходження стандартної попередньої обробки модель визначає ймовірність шахрайства, на основі якої приймається рішення щодо транзакції.

Загальна схема методу (рис. 1) демонструє послідовність етапів та можливі зворотні зв’язки, що виникають під час оптимізації моделі. Важливою властивістю підходу є його модульність, яка дозволяє змінювати окремі компоненти — алгоритм навчання чи методи обробки — без перебудови всієї системи.

Метод також підтримує адаптивність: періодичне перенавчання на оновлених даних забезпечує актуальність моделі та її здатність реагувати на зміну характеру шахрайських операцій у банківській сфері.

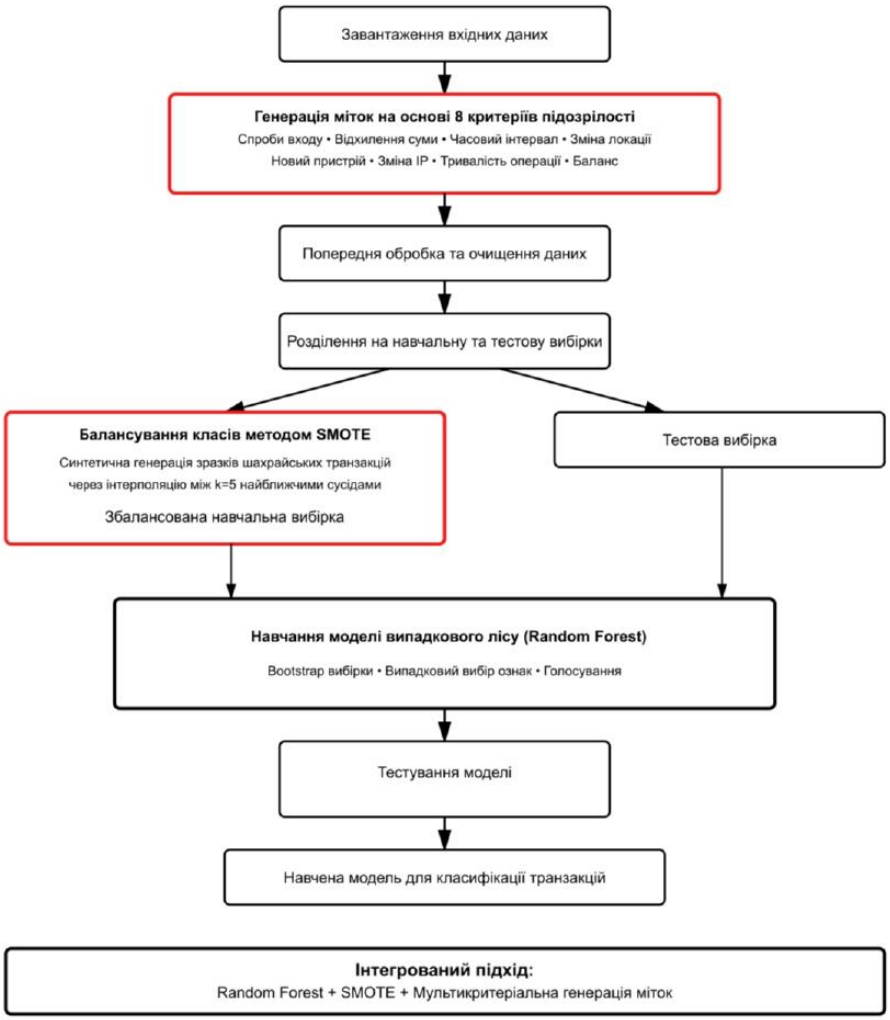


Рис. 1. Загальна схема методу виявлення шахрайських операцій

Для класифікації банківських транзакцій обрано алгоритм випадкового лісу — ансамблевий метод, що характеризується високою точністю, стійкістю до перенавчання та здатністю працювати з великою кількістю як числових, так і категоріальних ознак. Ці властивості роблять його ефективним для задачі виявлення шахрайських операцій.

Випадковий ліс складається з множини дерев рішень, кожне з яких навчається на бутстреп-підвибірці навчальних даних. Завдяки незалежності дерев їх помилки взаємно компенсуються, що підвищує якість класифікації порівняно з окремою моделлю.

У процесі навчання для кожної підвибірки будується дерево, причому на кожному кроці розглядається випадкова підмножина ознак — зазвичай розміром, що дорівнює квадратному кореню з їх загальної кількості. Це забезпечує різноманітність дерев та зменшує кореляцію між ними. Дерева зазвичай вирощуються до повної глибини, а перенавчання запобігається завдяки подальшому усередненню їхніх передбачень.

Під час класифікації нової транзакції вона проходить через усі дерева, кожне з яких формує власне рішення. Остаточний клас визначається більшістю голосів, що забезпечує надійність та стабільність результатів моделі.

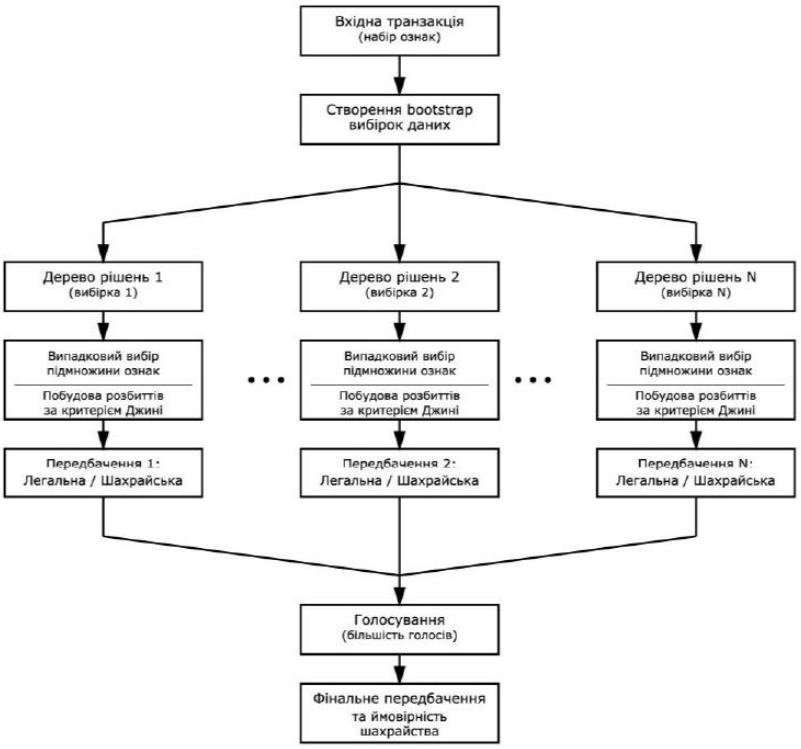


Рис. 2. Архітектура моделі класифікації транзакцій

У задачі виявлення шахрайства випадковий ліс може формувати ймовірнісну оцінку належності транзакції до класу шахрайських — як частку дерев, що віднесли її до цього класу. Це дозволяє встановлювати порогові значення та регулювати баланс між виявленням фроду та кількістю хибних спрацювань.

Алгоритм також забезпечує оцінку важливості ознак, визначаючи внесок кожної з них у зменшення помилки класифікації. Такі оцінки дають змогу визначити, які характеристики транзакцій найбільше впливають на виявлення шахрайства.

Архітектура алгоритму визначається кількома параметрами: кількістю дерев, їх максимальною глибиною, мінімальною кількістю зразків для розбиття та для листа. Збільшення числа дерев підвищує точність, але потребує більше ресурсів.

У цьому дослідженні використано 100 дерев, що забезпечує стабільність без надмірних витрат. Глибину дерев не обмежено, мінімальну кількість зразків для розбиття встановлено на рівні 2, а для листа — 1. Кількість ознак для кожного розбиття обчислюється як корінь квадратний із загальної кількості ознак. Для оцінки якості розбиттів застосовано критерій Джині.

Експериментальні дослідження

Для експериментальної перевірки методу використано Bank Transaction Dataset, що містить інформацію про банківські операції. Кожна транзакція описується набором ознак: ідентифікатори транзакції та рахунку, сума, тип операції, дата й час, геолокація, ідентифікатор пристрою, IP-адреса, тривалість, кількість спроб входу, баланс рахунку.

Оскільки датасет не містив міток класів, їх було автоматично сформовано на основі восьми критеріїв підозрілості (надмірні спроби входу, аномальна сума, надто малий інтервал між операціями, різке географічне

переміщення, новий пристрій, зміна IP, нетипова тривалість, низький баланс). Транзакції з ≥3 балами позначались як шахрайські. У результаті отримано 126 шахрайських та 2386 легальних операцій.

Підготовка даних включала аналіз і заповнення пропусків: 12 відсутніх значень для геолокації, 8 — для ідентифікатора пристрою, 5 — для IP-адреси. Для числових ознак застосовувалась медіана, для категоріальних — найчастіше значення (SimpleImputer).

Числові ознаки (сума, тривалість, кількість спроб входу, баланс) нормалізовано методом стандартизації (StandardScaler). Категоріальні ознаки — тип транзакції, геолокація, ідентифікатор пристрою — закодовано через OneHotEncoder.

Дані поділено на вибірки зі збереженням пропорції класів: 1758 транзакцій у навчальній вибірці, 377 у валідаційній та 377 у тестовій.

Сильний дисбаланс класів у навчальній вибірці (≈95% легальних проти 5% шахрайських транзакцій) призводив до того, що модель могла досягати високої точності, фактично не виявляючи шахрайства. Для усунення цієї проблеми застосовано SMOTE, який генерує синтетичні зразки: для кожної шахрайської транзакції визначається п'ятеро найближчих сусідів, випадково обирається один із них, і новий зразок створюється шляхом інтерполяції між двома точками.

Процес генерації тривав до вирівнювання класів: у вихідній навчальній вибірці було близько 1670 легальних та 88 шахрайських транзакцій, після застосування SMOTE кількість шахрайських зразків збільшено до 1670.

Балансування виконувалось лише для навчальної вибірки, тоді як валідаційна та тестова частини залишалися незмінними, що забезпечило реалістичну та об'єктивну оцінку моделі.

Результати застосування SMOTE наведено на рисунку 3, де показано порівняння метрик моделі до та після балансування. Балансування суттєво підвищило здатність моделі виявляти шахрайські операції: повнота зросла з 0.38 до 0.79.

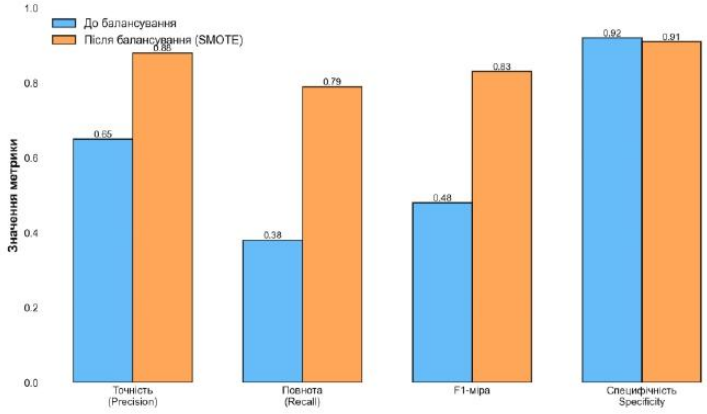


Рис.3. Порівняння метрик класифікації до та після балансування класів

Для навчання класифікатора використано алгоритм Random Forest із бібліотеки scikit-learn. Він формує ансамбль дерев рішень, кожне з яких тренується на випадковій підмножині даних та ознак. Основним параметром є кількість дерев, яку варіювали від 10 до 200, оцінюючи F1-міру на валідаційній вибірці. Графік (рис. 4) показує зростання F1-міри зі збільшенням числа дерев, але після 100 дерев приріст майже зникає: при 100 деревах F1 = 0.82, при 200 — 0.83.

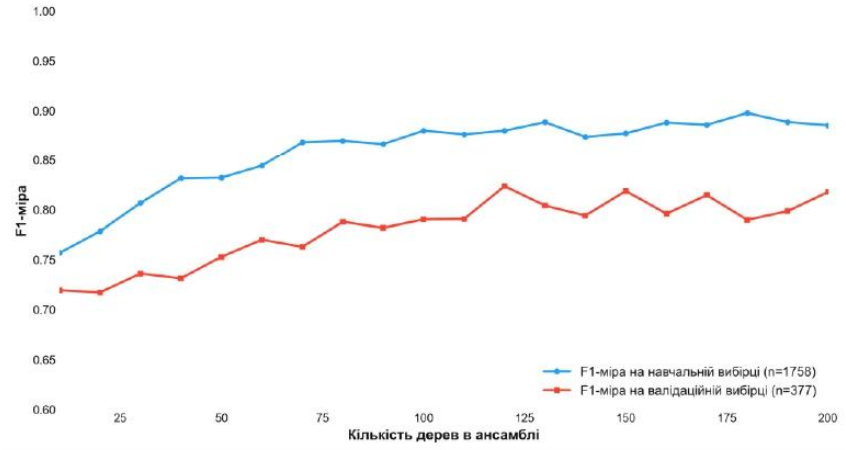


Рис. 4. Залежність якості класифікації від кількості дерев

Для балансу між якістю класифікації та складністю моделі вибрали 100 дерев. Як критерій розбиття встановлено індекс Джині, а параметр *max_features* — корінь квадратний із загальної кількості ознак, що забезпечує різноманітність дерев. Глибина дерев не обмежувалась; мінімальна кількість зразків у вузлі — 2, у листі — 1. Після навчання обчислено важливість ознак як середнє зменшення індексу Джині; результати наведені на рисунку 5.

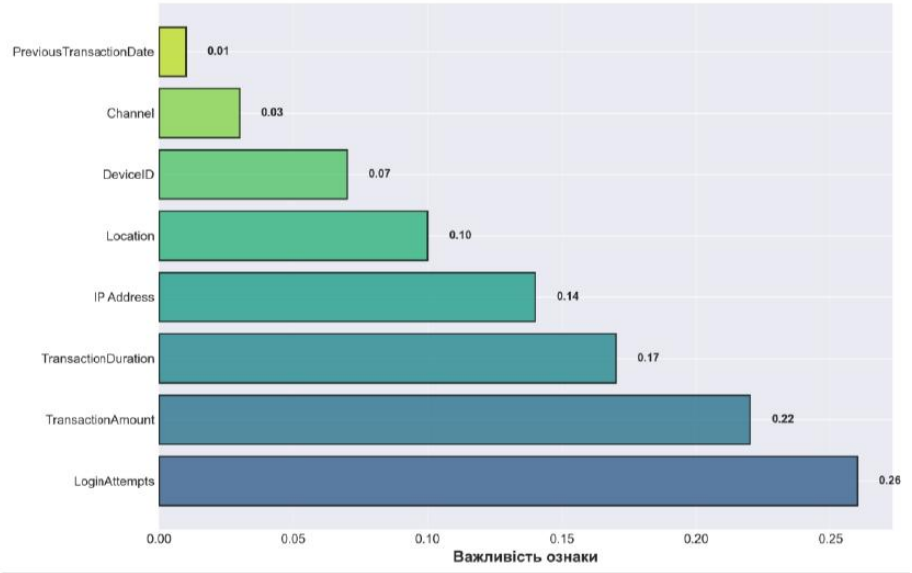


Рис.5. Важливість ознак для виявлення шахрайських операції

Найважливішою ознакою стала кількість спроб входу (0.26), що найкраще відрізняє шахрайські транзакції від легальних. Далі за значущістю йдуть сума операції (0.22) та її тривалість (0.17). Геолокація та IP-адреса мають нижчу важливість — 0.10 і 0.14. Найменшу важливість показали канал, ідентифікатор пристрою та попередня транзакція — 0.03, 0.07 і 0.01.

Для оцінки якості моделі побудовано криву точності-повноти та ROC-криву, що демонструють роботу класифікатора при різних порогах. Площа під кривою точності-повноти становить 0.84, що свідчить про високу

якість моделі. Робоча точка при порозі 0.5 дає точність 0.86 і повноту 0.79. Підвищення порогу збільшує точність, але знижує повноту (наприклад, при 0.7: точність 0.93, повнота 0.63). Зниження порогу до 0.3 підвищує повноту до 0.89, але зменшує точність до 0.71. Вибір порогу залежить від пріоритетів — максимальне виявлення шахрайства чи мінімізація хибних спрацювань.

ROC-крива відображає співвідношення істинно та хибно позитивних спрацювань. Графік ROC-кривої представлений на рисунку 6. Діагональ відповідає випадковому класифікатору.

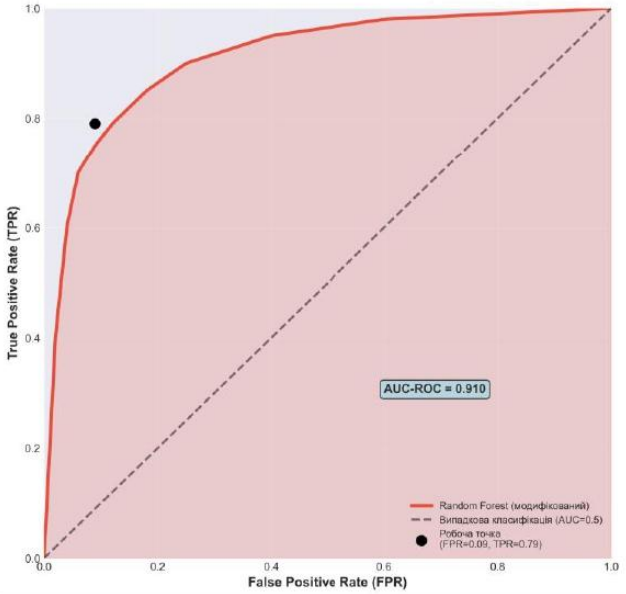


Рис.6. ROC-крива моделі класифікації

Площа під ROC-кривою становить 0.91, що вказує на високу здатність моделі розрізняти класи. Робоча точка має координати: хибно позитивні — 0.09, істинно позитивні — 0.79. ROC-крива значно перевищує діагональ і має вигин у верхній лівий кут, що демонструє високу чутливість при низькій частці хибних спрацювань. Обидві криві підтверджують високу якість моделі: площі 0.84 і 0.93 значно перевищують рівень базового класифікатора та свідчать про надійне розрізнення шахрайських і легальних транзакцій.

Для оцінки ефективності методу проведено порівняння з популярними алгоритмами, навченими на однакових даних. Логістична регресія як базова модель дала F1 = 0.64 (точність 0.71, повнота 0.58), що свідчить про пропуск багатьох шахрайських транзакцій через лінійність моделі. Дерево рішень (критерій Джині, глибина 10) показало кращі результати: F1 = 0.72, точність 0.78, повнота 0.67, проте лишається ризик перенавчання. Третя модель — базовий Random Forest із 100 дерев без балансування класів, навчений на незбалансованій вибірці. Порівняльні результати наведені на рисунку 7.

Базовий Random Forest дав F1 = 0.76 (точність 0.83, повнота 0.71), що краще за дерево рішень, але гірше за розроблений метод, оскільки незбалансовані дані погіршили виявлення шахрайства. Розроблений Random Forest із SMOTE досяг F1 = 0.82 (точність 0.86, повнота 0.79), значно покращивши розпізнавання шахрайських транзакцій. XGBoost (100 дерев, глибина 6, learning rate 0.1) показав F1 = 0.78, що трохи нижче за результат SMOTE-Random Forest, ймовірно через особливості даних або потребу в донастроюванні. Загалом, метод Random Forest із SMOTE продемонстрував найкращу F1-міру серед усіх алгоритмів.

Порівняння з базовим Random Forest показує, що балансування класів суттєво підвищує якість: SMOTE збільшив F1 з 0.76 до 0.82 та повноту з 0.71 до 0.79, забезпечивши додаткове виявлення 8% шахрайських транзакцій. Перевага над XGBoost свідчить, що саме поєднання Random Forest і SMOTE виявилось найефективнішим для цієї задачі.

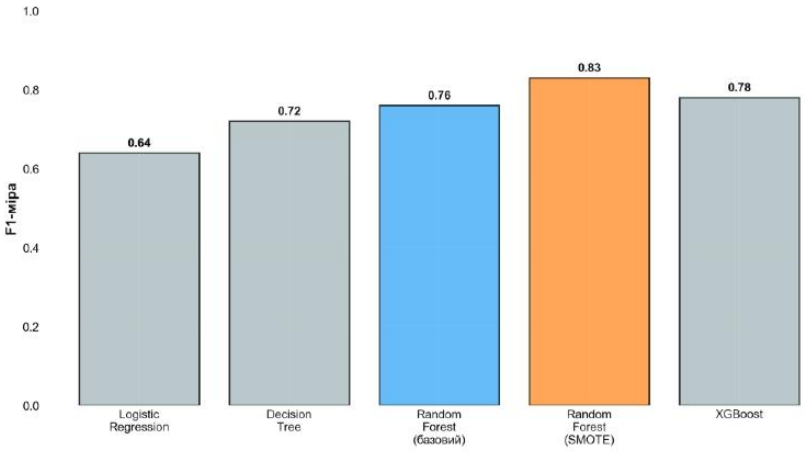


Рис.7. Порівняння якості моделей

Аналіз помилок системи класифікації дає змогу зрозуміти слабкі місця моделі та напрямки для подальшого покращення.

Матриця помилок показала 4 хибно негативні та 2 хибно позитивні помилки. Пропущені шахрайські транзакції були малопомітними: звичні суми, відомі пристрої, стандартні локації, лише поодинокі слабкі індикатори (зміна IP, швидке виконання, невдалий вхід). Модель оцінила їх як недостатньо підозрілі, тому 4 із 19 випадків були класифіковані помилково. Хибно позитивні помилки (2 з 358) виникали через незвичні для клієнтів операції — нові пристрої після зміни телефону, великі нетипові суми або транзакції під час подорожей. Аналіз показує можливість покращення моделі за рахунок нових ознак (довгострокова історія, поведінкові патерни) та оптимізації порогу класифікації.

Експеримент з різним обсягом навчальних даних показав, що якість моделі швидко зростає від 10% до 40% даних (F1 від 0.58 до 0.75), після чого покращення сповільнюється. На 70% даних модель досягає F1 = 0.82 — значення, використане в основних експериментах. Подальше збільшення вибірки дає мінімальний приріст (до F1 = 0.84 на 100%), що свідчить про вихід на плато. Для досягнення F1 > 0.75 потрібно щонайменше 40% даних (~700 транзакцій).

Практичний аналіз показує, що при повноті 0.79 модель виявляє 79 зі 100 шахрайських операцій (близько 395 з 500 на день), пропускаючи приблизно 105. Точність 0.88 означає, що з усіх позначених системою підозрілих 88% є шахрайськими, що створює близько 54 хибних спрацювань на день — прийнятне навантаження для служби безпеки. Специфічність 0.99 гарантує мінімальні незручності для клієнтів.

Порівняно з альтернативами, модель істотно краща: логістична регресія (F1 = 0.64) пропустила б додатково ~105 випадків шахрайства, базовий Random Forest (F1 = 0.76) — ~40. Найважливіші ознаки — кількість спроб входу та сума транзакції; вони вказують на ключові напрямки моніторингу.

Аналіз помилок визначив ситуації, що потребують особливої уваги: транзакції з нових пристроїв, під час подорожей, або нетипово великі суми — вони часто спричиняють хибні спрацювання. Модель можна покращити додаванням поведінкових ознак та адаптивними порогами: підвищеним для великих сум і зниженим для транзакцій з нових локацій.

Практична система може працювати в режимах автоматичного блокування (ймовірність >0.9), попередження (0.5–0.9) та моніторингу (0.3–0.5). Такий підхід зменшує втрати від шахрайства, покращує роботу служби безпеки та підвищує комфорт клієнтів.

Висновки

У роботі запропоновано та обґрунтовано метод виявлення шахрайських банківських операцій з використанням ансамблевого навчання (на базі алгоритму Random Forest) та техніки SMOTE. Відмінною рисою методу є глибокий аналіз мережевих і системних атрибутів транзакцій (таких як IP-адреса, геолокація, цифрові відбитки пристрою та часові параметри сесій), що дозволяє виявляти приховані аномалії та ефективно балансувати класи даних для підвищення точності класифікації.

Розроблена модульна програмна система забезпечує повний цикл обробки специфічних атрибутів транзакцій та гнучкість у тестуванні й модифікації компонентів захисту. Завдяки використанню ансамблевого підходу, система підтримує розпаралелювання обчислень, що критично важливо для обробки потоків даних у реальному часі.

Експериментальне дослідження підтвердило високу ефективність запропонованого методу: F1-міра – 0.83, точність – 0.86, повнота – 0.79, площа під ROC-кривою – 0.93. Розроблений підхід дозволяє виявляти близько 79% шахрайських операцій при рівні хибних спрацювань лише 4,5%, що перевищує результати традиційних методів та окремих алгоритмів класифікації, які не враховують комплексні системні атрибути.

Отримані результати підтверджують практичну цінність методу для задач захисту фінансової інформації. Впровадження запропонованого рішення дозволяє автоматизувати процеси моніторингу безпеки, забезпечити цілісність транзакційних даних, знизити фінансові втрати від шахрайства та підвищити рівень довіри клієнтів до цифрових банківських сервісів.

Література

1. Siam A. M., Bhowmik P., Uddin M. P. Hybrid feature selection framework for enhanced credit card fraud detection using machine learning models. *PLOS One*. 2025. Vol. 20, No. 7. URL https//doi.org/10.1371/journal.pone.0326975.

2. Hafez I. Y., Hafez A. Y., Saleh A., Abd El-Mageed A. A., Abohany A. A. A systematic review of AI-enhanced techniques in credit card fraud detection. *Journal of Big Data*. 2025. Vol. 12, No. 1. Pp. 6. URL https//doi.org/10.1186/s40537-024-01048-8.

3. Credit Card Fraud Identification Using Machine Learning on Graphs – RelationalAI. URL https//www.relational.ai/post/credit-card-fraud-detection-machine-learning-graphs.

4. Alatawi M. N. Detection of fraud in IoT based credit card collected dataset using machine learning. *Machine Learning with Applications*. 2025. Vol. 19. Pp. 100603. URL https//doi.org/10.1016/j.mlwa.2024.100603.

5. Baisholan N., Dietz J. E., Gnatyuk S., Turdalyuly M., Matson E. T., Baisholanova K. FraudX AI An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets. *Computers*. 2025. Vol. 14, No. 4. Pp. 120. URL https//doi.org/10.3390/computers14040120.

6. Hayat K., Magnier B. Data Leakage and Deceptive Performance A Critical Examination of Credit Card Fraud Detection Methodologies. *Mathematics*. 2025. Vol. 13, No. 16. Pp. 2563. URL https//doi.org/10.3390/math13162563.

7. Moradi F., Tarif Hokmabadi M., Homaei M. A Systematic Review of Machine Learning in Credit Card Fraud Detection. Computer Science and Mathematics, 2025. URL https//doi.org/10.20944/preprints202507.1085.v1.

8. Thennakoon A., Bhagyani C., Premadasa S., Mihiranga S., Kuruwitaarachchi N. Real-time Credit Card Fraud Detection Using Machine Learning / *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, January 2019. Pp. 488–493. URL https//doi.org/10.1109/CONFLUENCE.2019.8776942.

9. Tanouz D., Subramanian R. R., Eswar D., Reddy G. V. P., Kumar A. R., Praneeth C. V. N. M. Credit Card Fraud Detection Using Machine Learning / *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, May 2021. Pp. 967–972. URL https//doi.org/10.1109/ICICCS51141.2021.9432308.

10. Chen Y., Zhao C., Xu Y., Nie C., Zhang Y. Year-over-Year Developments in Financial Fraud Detection via Deep Learning A Systematic Literature Review. arXiv, 2025. URL https//doi.org/10.48550/arXiv.2502.00201.

11. Chen Y., Zhao C., Xu Y., Nie C., Zhang Y. Deep Learning in Financial Fraud Detection Innovations, Challenges, and Applications. *Data Science and Management*. 2025. URL https//doi.org/10.1016/j.dsm.2025.08.002.

12. Jin J., Zhang Y. The analysis of fraud detection in financial market under machine learning. *Scientific Reports*. 2025. Vol. 15, No. 1. Pp. 29959. URL https//doi.org/10.1038/s41598-025-15783-2.

13. Afriyie J. K., Tawiah K., Pels W. A., Addai-Henne S., Dwamena H. A., Owiredu E. O., Ayeh S. A., Eshun J. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*. 2023. Vol. 6. Pp. 100163. URL https//doi.org/10.1016/j.dajour.2023.100163.

14. Salomon S. What is Fraud Detection for Machine Learning? *Feedzai*. URL https//www.feedzai.com/blog/what-is-fraud-detection-for-machine-learning/.

15. Credit Card Fraud Detection. URL https//www.kaggle.com/datasets/mlg-ulb/creditcardfraud.

16. Fraud Detection Dataset. URL https//www.kaggle.com/datasets/goyaladi/fraud-detection-dataset.

17. Synthetic Financial Datasets For Fraud Detection. URL https//www.kaggle.com/datasets/ealaxi/paysim1.

18. Ryzhanskyi O., Pavlyshyn V., Radiuk P., Manziuk E., Barmak O., Krak I. AI-Driven Traffic Signal Control System to Reduce CO2 Emissions / CEUR Workshop Proc., CEUR-WS, 2025. Pp. 18–27. URL: https://ceur-ws.org/Vol-3974/paper02.pdf.

19. Pavlyshyn V., Ryzhanskyi O., Manziuk E., Radiuk P., Barmak O., Krak I. Establishing Patterns of the Urban Transport Flows on Clustering Analysis. CEUR Workshop Proceedings. 2025. Vol. 3974. Pp. 1–9. URL: https://ceur-ws.org/Vol-3974/paper01.pdf.

20. Manziuk E., Barmak O., Krak I., Petliak N., Jin Z., Radiuk P. Explainable Deep Learning for Interpretable Brain Tumor Diagnosis from MRI Images / Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making, Volume 1, Cham, Springer Nature Switzerland, 2024. Pp. 326–348. URL: https://doi.org/10.1007/978-3-031-70959-3_17.

21. Barmak O., Krak I., Yakovlev S., Manziuk E., Radiuk P., Kuznetsov V. Toward explainable deep learning in healthcare through transition matrix and user-friendly features. Frontiers in Artificial Intelligence. 2024. Vol. 7. Pp. 1482141. URL: https://doi.org/10.3389/frai.2024.1482141.

22. Manziuk E., Krak I., Barmak O., Mazurets O., Kuznetsov V., Pylypiak O. Structural alignment method of conceptual categories of ontology and formalized domain. 2021. Pp. 11–22.

23. An adaptive approach to detecting fake news based on generalized text features(Conference Paper) Shupta, A., Barmak, O., Wierzbicki, A., Skrypnyk, T. // 7th International Conference on Computational Linguistics and Intelligent Systems. Volume I: Machine Learning Workshop, CoLInS 2023; Kharkiv; Ukraine; 20 April 2023 до 21 April 2023; Код 188444 // CEUR Workshop Proceedings Volume 3387, 2023, Pages 300-310

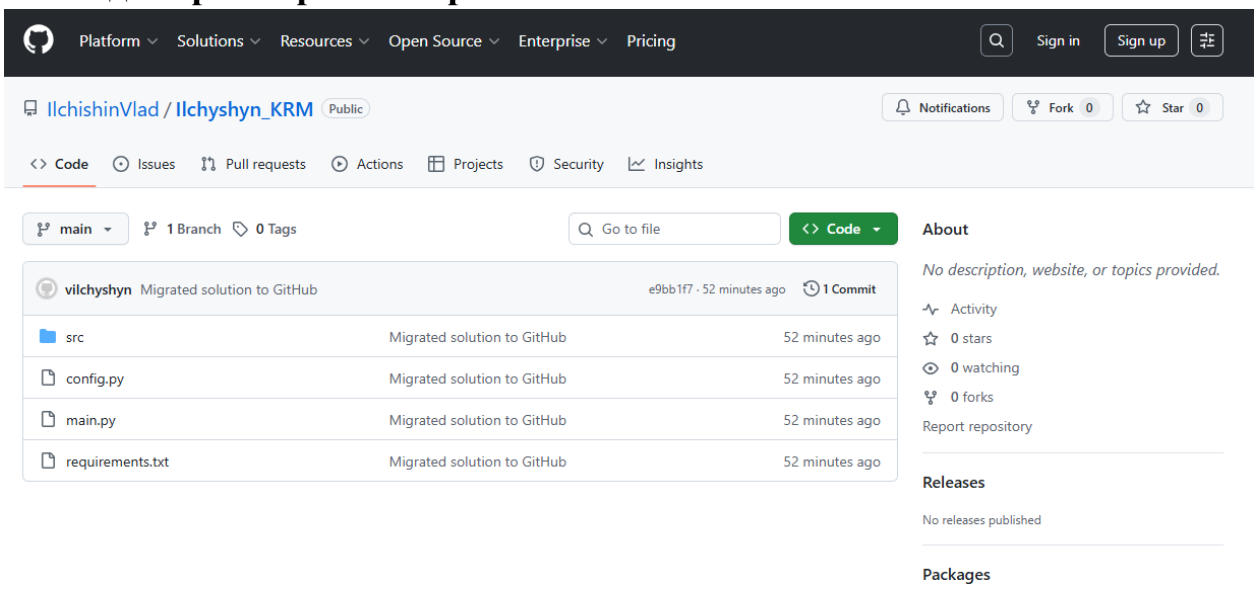
24. E. Manziuk, O. Barmak, I. Krak, O. Mazurets, and T. Skrypnyk, “Formal Model of Trustworthy Artificial Intelligence Based on Standardization.,” in CEUR Workshop Proceedings, Khmelnytskyi, Ukraine, Mar. 2021, vol. 2853, pp. 190–197. http://ceur-ws.org/Vol-2853/

Додаток Б

Програмний код посилання на GitHub-репозиторій, структура проєкту та опис основних папок і файлів

У роботі розглядається підхід машинного навчання на основі ансамблевого методу Random Forest для класифікації фінансових операцій. Метою дослідження є підвищення точності виявлення шахрайських транзакцій шляхом створення спеціалізованого методу бінарної класифікації з балансуванням навчальної вибірки та подальшою розробкою інформаційної системи.

Посилання на репозиторій на GitHub:
https://github.com/IlchishinVlad/Ilchyshyn_KRM
Вигляд сторінки репозиторію:



Опис вмісту:

Модулі обробки даних (src/data/)

- **data_loader.py** - завантаження набору даних транзакцій та створення структур для навчання
- **preprocessor.py** - нормалізація числових ознак, кодування категоріальних змінних та обробка пропущених значень

Модулі генерації ознак (src/features/)

- **generator.py** - генерація міток класів на основі правил виявлення підозрілої активності (аномальні суми, частота входів, зміна геолокації, нові пристрої)

Модулі машинного навчання (src/models/)

- **classifier.py** - реалізація моделі Random Forest з оптимізацією гіперпараметрів
- **balancer.py** - балансування класів методом SMOTE для вирішення проблеми незбалансованості

Модулі оцінки якості (src/evaluation/)

- **metrics.py** - обчислення Precision, Recall, F1-score, ROC-AUC, Specificity для оцінки якості моделі

Головний скрипт

- **main.py** - основний скрипт для повного циклу: завантаження даних, попередня обробка, генерація міток, балансування, навчання та оцінка моделі

Конфігурація системи

- **config.py** - параметри моделі Random Forest, критерії виявлення шахрайства, шляхи до даних та порогові значення
- **requirements.txt** - залежності Python для роботи системи

Додаток В

Презентаційний матеріал

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

МЕТОД ВИЯВЛЕННЯ ШАХРАЙСЬКИХ БАНКІВСЬКИХ ОПЕРАЦІЙ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Виконав:
студент 2 курсу, групи КНм-24-1
Владислав ІЛЬЧИШИН
Керівник:
д.т.н., професор кафедри КН
Едуард МАНЗЮК

2

Актуальність

Актуальність дослідження визначається критичною потребою фінансових установ у надійних засобах виявлення шахрайських операцій в умовах зростаючих обсягів цифрових транзакцій. Традиційні підходи до моніторингу, що базуються на ручному аналізі та простих правилах, виявляються неефективними через високу трудомісткість, значну кількість помилкових спрацювань і обмежену здатність виявляти нові схеми шахрайства.

Сучасні досягнення в області машинного навчання відкривають можливості для створення адаптивних систем моніторингу, здатних автоматично виявляти складні патерни підозрілої поведінки. Застосування ансамблевих методів та технік обробки незбалансованих даних дозволяє суттєво підвищити якість детектування шахрайських операцій при одночасному зменшенні навантаження на служби безпеки.

Мета і задачі роботи

Мета роботи полягає в підвищенні точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів.

Об'єкт дослідження – процес виявлення та класифікації шахрайських операцій у банківських транзакційних системах.

Предмет дослідження – моделі, методи та технології виявлення фінансового шахрайства на основі машинного навчання з застосуванням балансування незбалансованих даних.

Задачі дослідження:

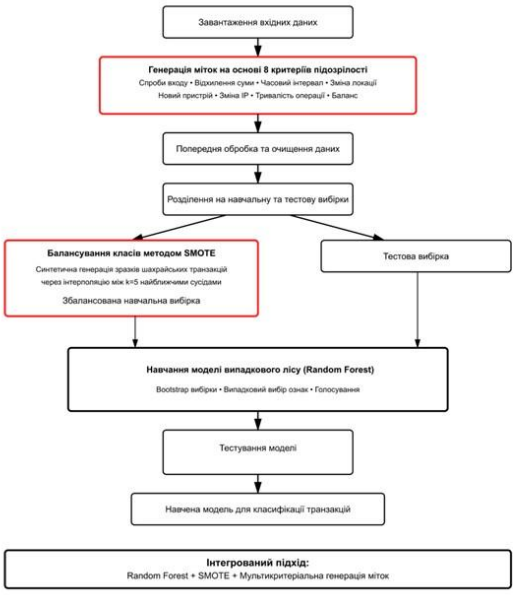
– провести аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розробити метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створити програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

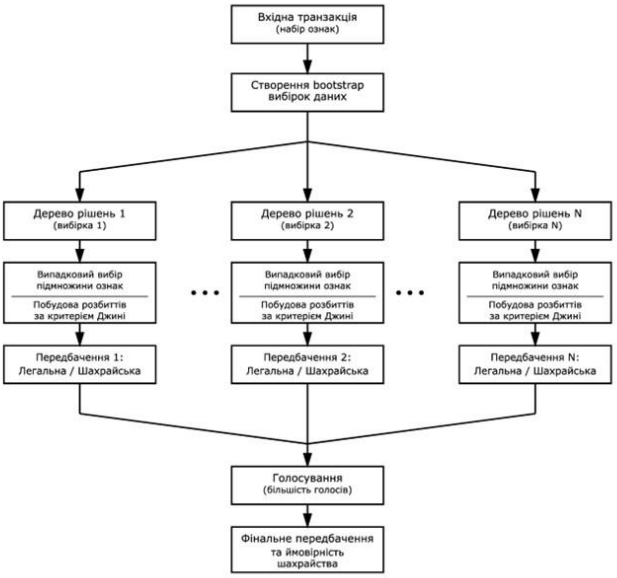
– провести експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

Загальна схема методу виявлення шахрайських операцій



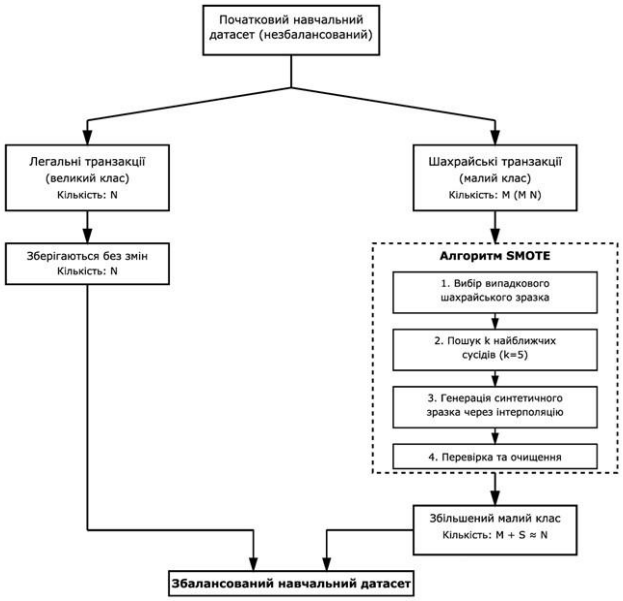
5

Архітектура моделі класифікації транзакцій



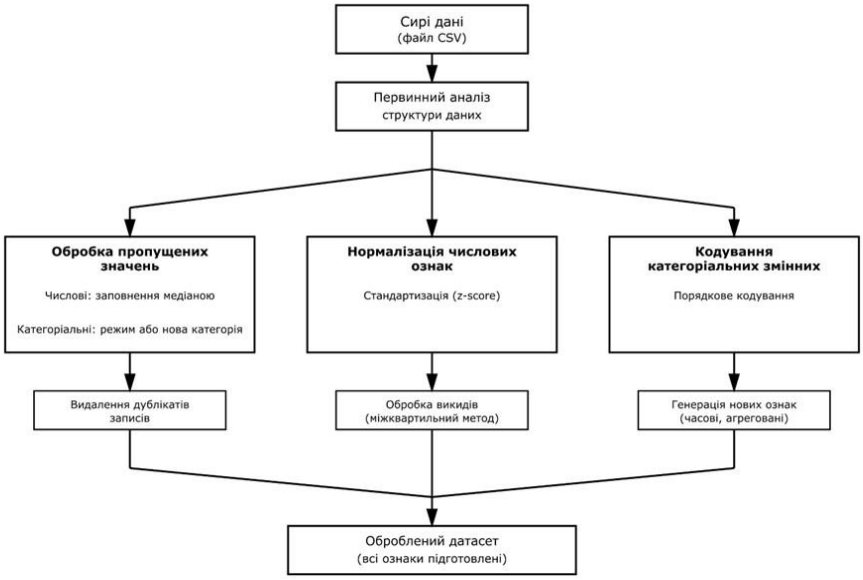
6

Процес балансування класів у навчальній вибірці



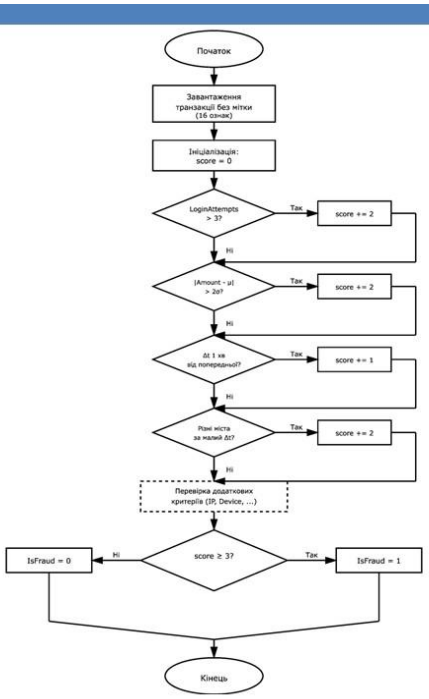
7

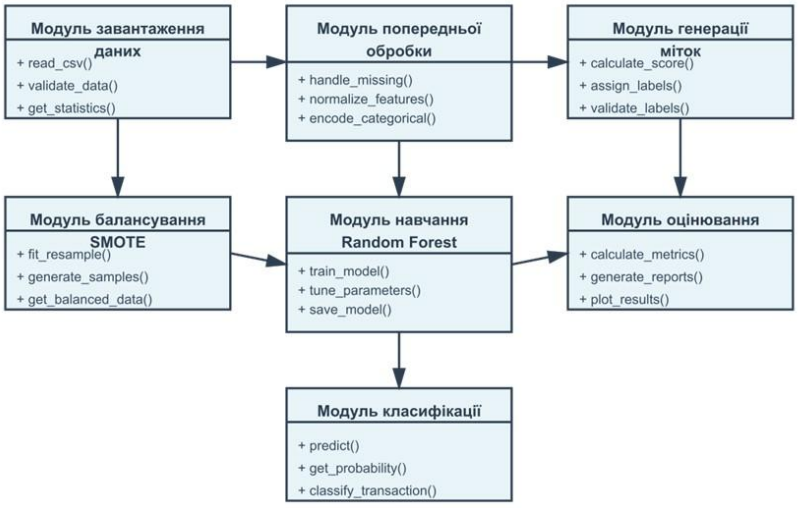
Етапи попередньої обробки та підготовки даних



8

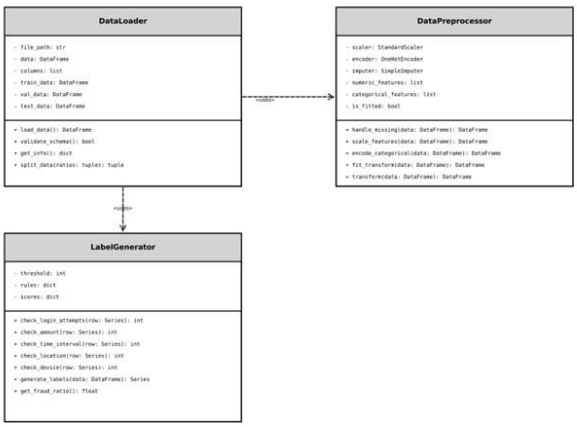
Процес генерації міток класів для транзакцій



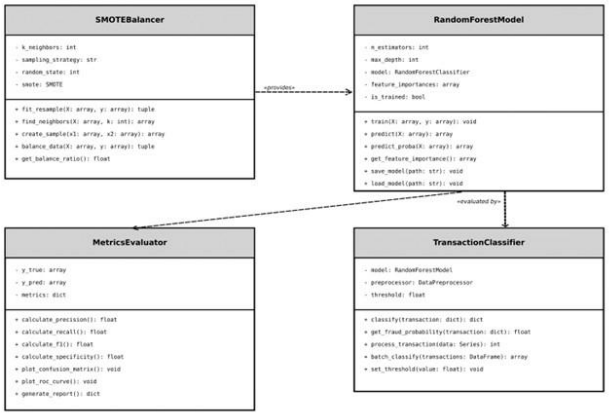


Діаграма компонентів

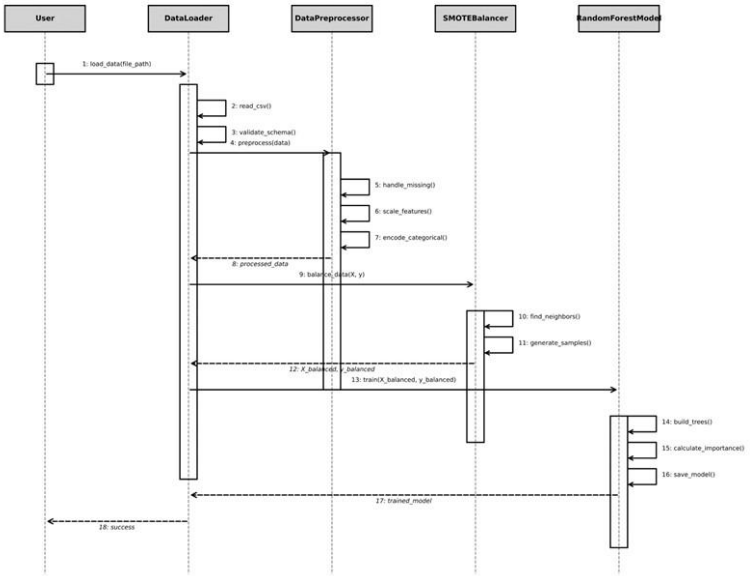
Діаграма класів модуля обробки даних



Діаграма класів модуля навчання моделі



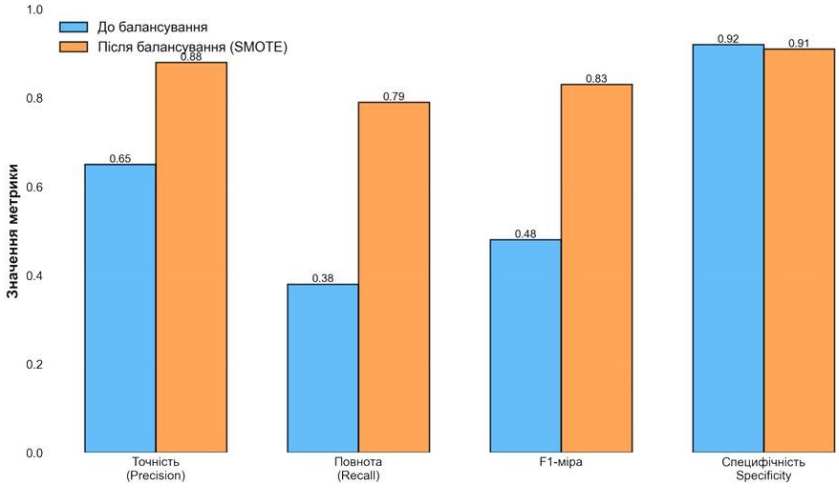
11



Діаграма послідовності процесу навчання моделі

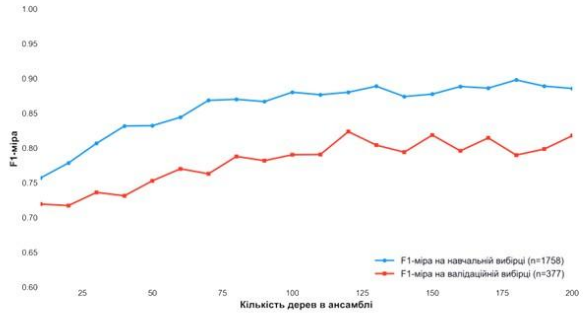
12

Порівняння метрик класифікації до та після балансування класів

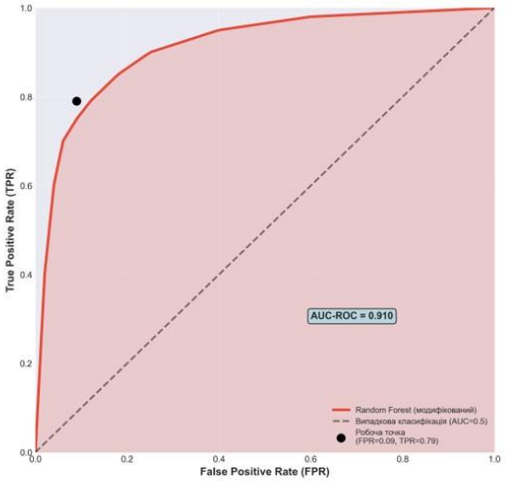


13

Залежність якості класифікації від кількості дерев

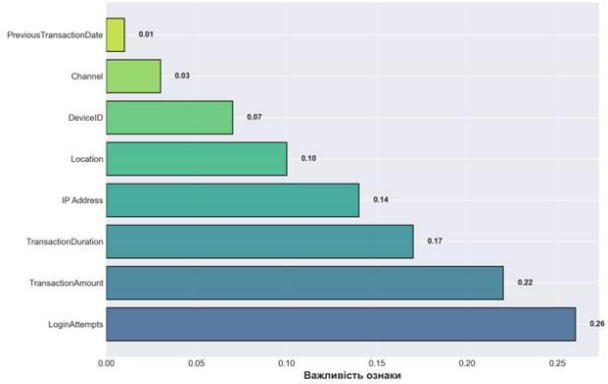


ROC-крива моделі класифікації

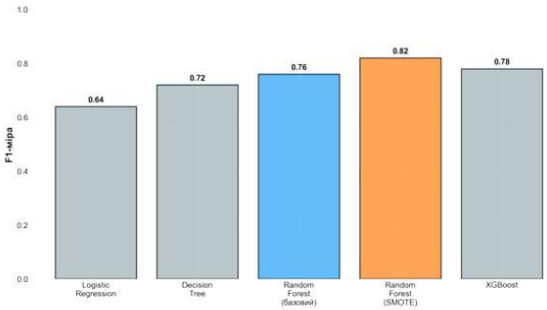


14

Важливість ознак для виявлення шахрайських операції



Порівняння якості моделей



Висновки

Основні результати дослідження полягають у наступному:

– проведено аналіз існуючих методів та підходів до виявлення фінансового шахрайства з використанням методів машинного навчання;

– розроблено метод виявлення та класифікації шахрайських транзакцій на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE для вирішення проблеми незбалансованості класів;

– створено програмну реалізацію методу класифікації банківських транзакцій з модульною архітектурою, що забезпечує можливість масштабування та адаптації;

– проведено експериментальне дослідження ефективності спроектованого методу шляхом порівняння з альтернативними алгоритмами класифікації та оцінки його точності на реальних транзакційних даних.

ДЯКУЮ ЗА УВАГУ!

Mon Dec 08 20:22:22 EET 2025, Петровський Сергій Степанович, Хмельницький національний університет, ХНУ

Anti-Plagiarism (UA) v-15.281 Educational

The maximum coincidence with one document 1.0%

Dictionaries check: en_US, ru_RU, ua_UA. **Errors in the documents: 9%**

	Document		Sum coincidence on the DB	
ID: 252063 Title: КВАЛІФІКАЦІЙНА РОБОТА на тему Метод виявлення шахрайських банківських операцій з використанням машинного навчання Added in a DB: 2025-12-08 Authors: Владислав ІЛЬЧИШИН Heads: Едуард МАНЗЮК Consultants: Opponents:	Symbols	Lexemes	Symbols	Lexemes
	104481	1551	2024 (2%)	31 (2%)

Plagiarism sources

ID	Description	Plagiarism presence in the document	
		Symbols	Lexemes

Протокол аналізу звіту подібності науковим керівником

Заявляю, що я ознайомився (-лась) з Повним звітом подібності, який був згенерований Системою виявлення і запобігання плагіату щодо роботи:

Автор: Владислав ІЛЬЧИШИН

Співавтор:

Назва: КВАЛІФІКАЦІЙНА РОБОТА на тему Метод виявлення шахрайських банківських операцій з використанням машинного навчання

Науковий керівник: Едуард МАНЗЮК, д.т.н., професор

Підрозділ: Кафедра комп'ютерних наук

Коефіцієнт подібності 1:1.7%

Коефіцієнт подібності 2:0.2%

Мікропробіли: 0

Заміна букв: 2

Інтервали: 0

Білі знаки: 0

Дата створення звіту: 2025-12-10 16:14:08.0

Після аналізу Звіту подібності констатую наступне:

☒ **Запозичення, виявлені в роботі є законними і не є плагіатом. Рівень подібності не перевищує допустимої межі. Таким чином робота незалежна і приймається.**

☐ **Запозичення не є плагіатом, але перевищено граничне значення рівня подібностей. Таким чином робота повертається на доопрацювання.**

☐ **Виявлено запозичення і плагіат або навмисні текстові спотворення (маніпуляції), як передбачувані спроби укриття плагіату, які роблять роботу невідповідною вимогам законодавства (Ст. 32. ЗУ Про вищу освіту, пункт 3.1, Ст. 42. ЗУ Про освіту) та вимог НАЗЯВО (Критерій 5), а також кодексу етики і процедурам. Таким чином робота не приймається.**

Обгрунтування:

Дата 10.12.25 експерт Пеньовський С.С.

РІШЕННЯ ЕКСПЕРТНОЇ КОМІСІЇ КАФЕДРИ КОМП'ЮТЕРНИХ НАУК

ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Назва кваліфікаційної роботи *Метод виявлення шахрайських банківських операцій з використанням машинного навчання*
Автор *студент групи КНм-24-1 Владислав ІЛЬЧИШИН*
Освітня програма *Комп'ютерні науки*
Рівень вищої освіти *другий (магістерський)*
Спеціальність *122 – Комп'ютерні науки*
Науковий керівник: *д.т.н., проф. каф. комп'ютерних наук Едуард МАНЗЮК*

На основі аналізу кваліфікаційної роботи на дотримання вимог академічної доброчесності (у т.ч. відсутності ознак академічного плагіату) з урахуванням результатів перевірки роботи спеціалізованим програмними засобами комісія зробила такий висновок:

№	Висновок	Позначка про відповідність
1	Ознаки академічного плагіату	
1.1	Запозичення, виявлені в роботі, є законними і не є академічним плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних, якщо потрібно). Робота приймається до захисту.	***відповідає***
1.2	Виявлені запозичення не є академічним плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована.	
1.3	Виявлені запозичення не є академічним плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. Робота може бути допущена до захисту після того як буде відкоригована та доопрацьована і успішно пройде повторну перевірку на академічний плагіат.	
1.4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття текстових запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
2	Інші види порушень академічної доброчесності	***відсутні***

Підтвердження:

Запозичення, виявлені в роботі Владислава ІЛЬЧИШИНА, не є плагіатом, оскільки: запозичення розміщені в розділі огляду існуючих підходів, не описують безпосередньо авторську роботу і не стосуються її результатів; усі запозичення фрагментарні; до запозичень входять фрагменти, які не мають авторства і містять поширені конструкції та загальновідомі терміни, скорочення. Рівень подібності не перевищує допустимої межі. Таким чином, робота є законною та приймається до захисту.

Обсяг запозичень, визначений системами виявлення збігів/ідентичності/схожості:
- за системою Anti-Plagiarism: 1%;
- за системою StrikePlagiarism КП1: 1,7%, КП2: 0,2%.

10.12.2025

Завідувач кафедри _____ *Олександр БАРМАК*

Гарант освітньої програми _____ *Руслан БАГРІЙ*

Керівник кваліфікаційної роботи _____ *Едуард МАНЗЮК*

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МОН УКРАЇНИ

Кафедра комп'ютерних наук

ВІДГУК ОПОНЕНТА

на кваліфікаційну роботу магістра

студента *гр. КНм-24-1 Владислава ІЛЬЧИШИНА*

за темою *Метод виявлення шахрайських банківських операцій з використанням машинного навчання*

1. Актуальність обраної теми

Актуальність обраної теми зумовлена критичною необхідністю фінансових установ у надійних засобах виявлення шахрайських операцій в умовах зростаючих обсягів цифрових транзакцій та еволюції схем фінансового шахрайства. Застосування методів машинного навчання, зокрема ансамблевих алгоритмів та технік балансування незбалансованих даних, дозволяє значно підвищити ефективність виявлення шахрайських транзакцій, скоротити час реагування на підозрілі операції та зменшити фінансові втрати банків від шахрайства.

2. Відповідність роботи предметній області спеціальності 122 Комп'ютерні науки та загальним вимогам до наукових робіт

Магістерська робота повністю відповідає предметній області спеціальності 122 Комп'ютерні науки, оскільки ґрунтується на застосуванні методів машинного навчання, обробки даних та створенні інтелектуальних систем для вирішення задачі класифікації банківських транзакцій. Дослідження спирається на використання фундаментальних знань у галузі комп'ютерних наук, таких як алгоритми машинного навчання, структури даних, програмування та статистичний аналіз.

3. Повнота розкриття мети та завдань дослідження

Мета та завдання дослідження розкриті повністю. Автор чітко формулює мету роботи – підвищення точності виявлення та класифікації шахрайських банківських операцій шляхом розробки методу на основі ансамблевого навчання з балансуванням класів. Для досягнення мети послідовно вирішуються поставлені завдання, що включають комплексний аналіз наявних підходів до виявлення фінансового шахрайства, розробку методу класифікації банківських транзакцій.

4. Наявність наукової новизни

Наукова новизна роботи полягає в удосконаленні методу виявлення шахрайських банківських операцій, який відрізняється від існуючих інтегрованим застосуванням

алгоритму випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу та використанням комплексної системи генерації міток на основі множини критеріїв підозрілості, що дозволило підвищити повноту виявлення шахрайських транзакцій при збереженні високої точності класифікації.

5. Зміст кожного розділу роботи

Робота містить чотири розділи. В першому розділі представлено аналіз методів виявлення шахрайських банківських операцій, характеристику задачі, огляд існуючих публікацій та наукових підходів, а також архітектур, методів та моделей машинного навчання для класифікації шахрайських транзакцій. Другий розділ містить розробку методу виявлення шахрайських операцій з описом концепції, архітектури моделі класифікації, модифікації моделі. Третій розділ присвячено програмній реалізації методу з описом технологій. Розділ чотири містить експериментальне дослідження ефективності методу з аналізом результатів та порівнянням з альтернативними підходами.

6. Ступінь розкриття теми роботи

Тема роботи розкрита повністю. Автор аналізує проблематику виявлення банківського шахрайства, розглядає існуючі методи машинного навчання для класифікації транзакцій, обґрунтовує необхідність розробки удосконаленого методу. Детально описано запропонований метод на основі алгоритму випадкового лісу з інтегрованою технікою SMOTE, наведено архітектуру моделі з механізмами генерації міток на основі критеріїв підозрілості. Експериментальні дослідження на реальних банківських даних підтверджують ефективність методу.

7. Якість оформлення кваліфікаційної роботи

Якість оформлення кваліфікаційної роботи відповідає встановленим академічним стандартам, демонструючи чіткість, послідовність та професійність у структурі, форматуванні та презентації матеріалу.

8. Недоліки кваліфікаційної роботи

Недоліки кваліфікаційної роботи включають відсутність порівняння з більш широким спектром альтернативних алгоритмів машинного навчання, таких як градієнтний бустинг або глибокі нейронні мережі.

9. Загальний висновок (допускається чи не допускається до захисту), якої оцінки заслуговує кваліфікаційна робота

Враховуючи рівень виконання та забезпечення усіх необхідних вимог, наукову новизну отриманих результатів, якість проведених експериментальних досліджень робота може бути допущена до захисту. Рекомендована оцінка – відмінно.

Опонент _____ Світлана ГОВОРУЩЕНКО
д.т.н., професор

ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МОН УКРАЇНИ

Кафедра комп'ютерних наук

ВІДГУК НАУКОВОГО КЕРІВНИКА

на кваліфікаційну роботу магістра

студента *КНм-24-1 Владислава ІЛЬЧИШИНА*

за темою *Метод виявлення шахрайських банківських операцій з використанням машинного навчання*

1. Актуальність теми

Актуальність теми базується на необхідності фінансових установ у надійних засобах виявлення шахрайських операцій в умовах зростаючих обсягів цифрових транзакцій. Традиційні підходи до моніторингу, що базуються на ручному аналізі та простих правилах, виявляються неефективними через високу трудомісткість, значну кількість помилкових спрацювань і обмежену здатність виявляти нові схеми шахрайства. Застосування ансамблевих методів машинного навчання та технік обробки незбалансованих даних дозволяє суттєво підвищити якість детектування шахрайських операцій при одночасному зменшенні навантаження на служби безпеки, що робить цю тему надзвичайно актуальною для банківської індустрії та наукових досліджень у галузі фінансової безпеки.

2. Відповідність роботи предметній області Стандарту спеціальності 122 Комп'ютерні науки

Робота повністю відповідає предметній області спеціальності 122 "Комп'ютерні науки", оскільки вона ґрунтується на застосуванні методів машинного навчання, зокрема ансамблевого алгоритму випадкового лісу, та методів обробки даних. Дослідження передбачає використання алгоритмів класифікації, технік синтетичної генерації даних, методів нормалізації та трансформації ознак, статистичного аналізу та експериментального тестування, які є фундаментальними для комп'ютерних наук, та має практичне значення в галузі фінансової безпеки.

3. Професійні та особистісні якості

Владислав ІЛЬЧИШИН продемонстрував високий рівень професійної компетентності в галузі комп'ютерних наук та машинного навчання, відповідально та вчасно вирішуючи завдання з розробки методу виявлення шахрайських банківських операцій. Студент проявив глибоке розуміння алгоритмів класифікації, принципів роботи з незбалансованими даними, наполегливість у проведенні експериментальних досліджень та здатність до критичного аналізу результатів.

4. Ступінь самостійності під час виконання кваліфікаційної роботи

При виконанні магістерської роботи студент виявив високий рівень самостійності, запропонувавши удосконалення методу виявлення шахрайства шляхом інтегрованого застосування алгоритму випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу та використанням комплексної системи генерації міток на основі множини критеріїв підозрілості. Студент самостійно провів аналіз літератури, розробив архітектуру системи класифікації, реалізував програмне забезпечення з модульною структурою та провів всебічне експериментальне дослідження на реальних банківських даних.

5. Наукова новизна та оригінальність запропонованих підходів

Удосконалено метод виявлення шахрайських банківських операцій, який відрізняється від існуючих інтегрованим застосуванням алгоритму випадкового лісу з технікою SMOTE для синтетичної генерації зразків меншого класу та використанням комплексної системи генерації міток на основі множини критеріїв підозрілості, що дозволило підвищити повноту виявлення шахрайських транзакцій при збереженні високої точності класифікації.

6. Ступінь оволодіння методами дослідження

Студент продемонстрував глибоке розуміння та вміле застосування методів комп'ютерних наук та машинного навчання, зокрема ансамблевих методів класифікації, технік синтетичної генерації даних, методів нормалізації та трансформації ознак, статистичного аналізу та експериментального тестування на реальних транзакційних даних, для вирішення задачі виявлення фінансового шахрайства, що свідчить про його високий рівень оволодіння сучасними методами дослідження в галузі машинного навчання та фінансової безпеки.

7. Повнота та якість розкриття теми роботи

У магістерській роботі тема розкрита в повній мірі. Робота відзначається логічною структурою, глибиною аналізу існуючих підходів до виявлення фінансового шахрайства, детальним описом розробленого методу на основі випадкового лісу з балансуванням класів, його програмною реалізацією з модульною архітектурою та всебічним експериментальним дослідженням на реальних банківських даних, що свідчить про високий рівень розуміння предметної області та здатність до самостійного наукового дослідження.

8. Логічність, послідовність, аргументованість, літературна грамотність викладення матеріалу

Магістерська робота характеризується чіткою логічною структурою, послідовним викладенням матеріалу від аналізу проблематики фінансового шахрайства та існуючих підходів до розробки та експериментальної перевірки власного методу,

аргументованістю висновків, підкріплених детальним статистичним аналізом з використанням метрик точності, повноти, F1-міри та AUC-ROC. Автор демонструє високий рівень літературної грамотності, дотримуючись наукового стилю викладення та забезпечуючи легкість сприйняття матеріалу. Думки та ідеї подано в логічній послідовності, з належним обґрунтуванням та посиланнями.

9. Можливість практичного застосування кваліфікаційної роботи, окремих її частин

Розроблений у магістерській роботі метод виявлення шахрайських банківських операцій з використанням машинного навчання має широкі можливості практичного застосування в галузях автоматизованого моніторингу транзакцій, банківської безпеки, фінансової аналітики та управління ризиками. Метод може бути інтегрований у сучасні системи банківського моніторингу та платформи фінансової безпеки, що дозволить значно підвищити ефективність виявлення шахрайства при одночасному зменшенні кількості помилкових спрацювань та навантаження на служби безпеки фінансових установ.

10. Висновок про можливість допуску кваліфікаційної роботи до захисту, на яку оцінку заслуговує робота

Враховуючи належний рівень виконання та забезпечення усіх необхідних вимог, робота може бути допущена до захисту. Рекомендована оцінка «відмінно».

Керівник _____ *д.т.н., професор каф. КН Едуард МАНЗЮК*