# ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки
Кафедра ЕОМ

**Тема: «Програмні компоненти для системи виявлення шахрайських банківських транзакцій»**

Кваліфікаційна робота
перший (бакалаврський) рівень

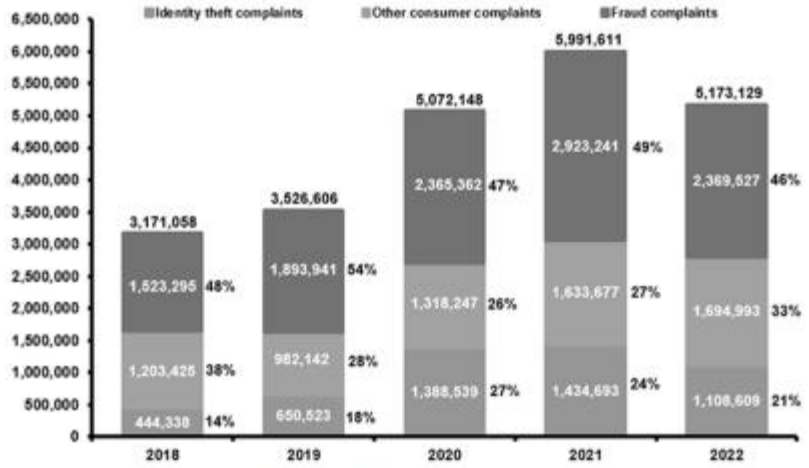
**Автор:**
**студент групи КІУКІ-21-1**
**Шевченко Станіслав Олександрович**

**Керівник роботи:**
**доц. каф. ЕОМ**
**к. т. н. Шматко О.В.**

Харків 2025

1

## Актуальність






2

# Актуальність



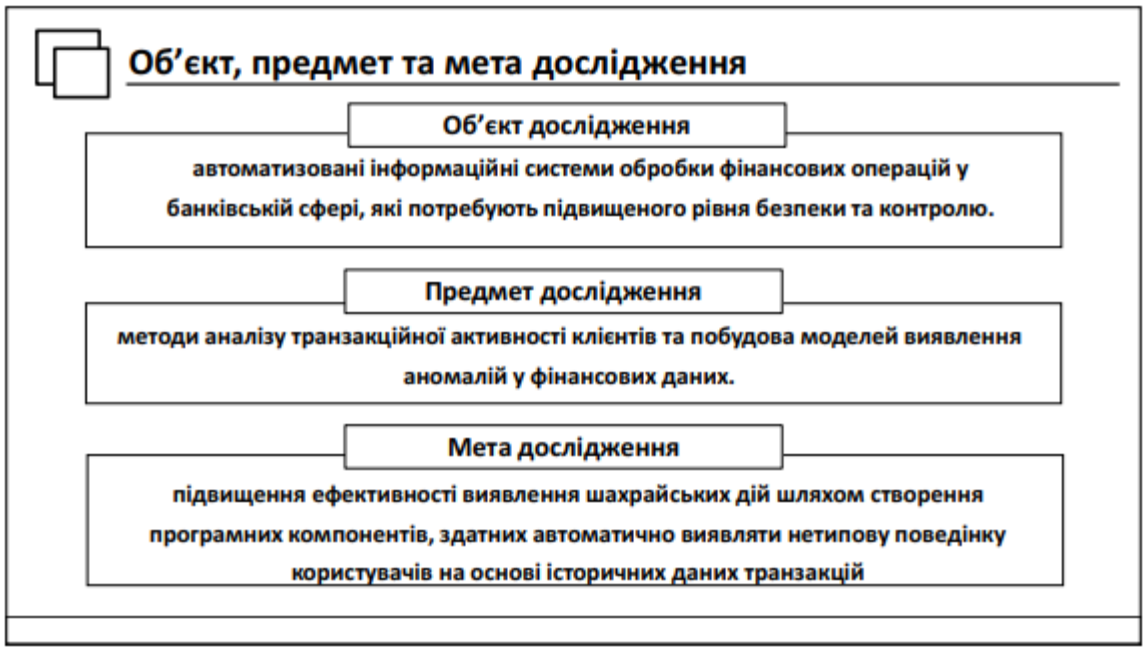

3

## Найпоширеніші способи шахрайства з платіжними картками

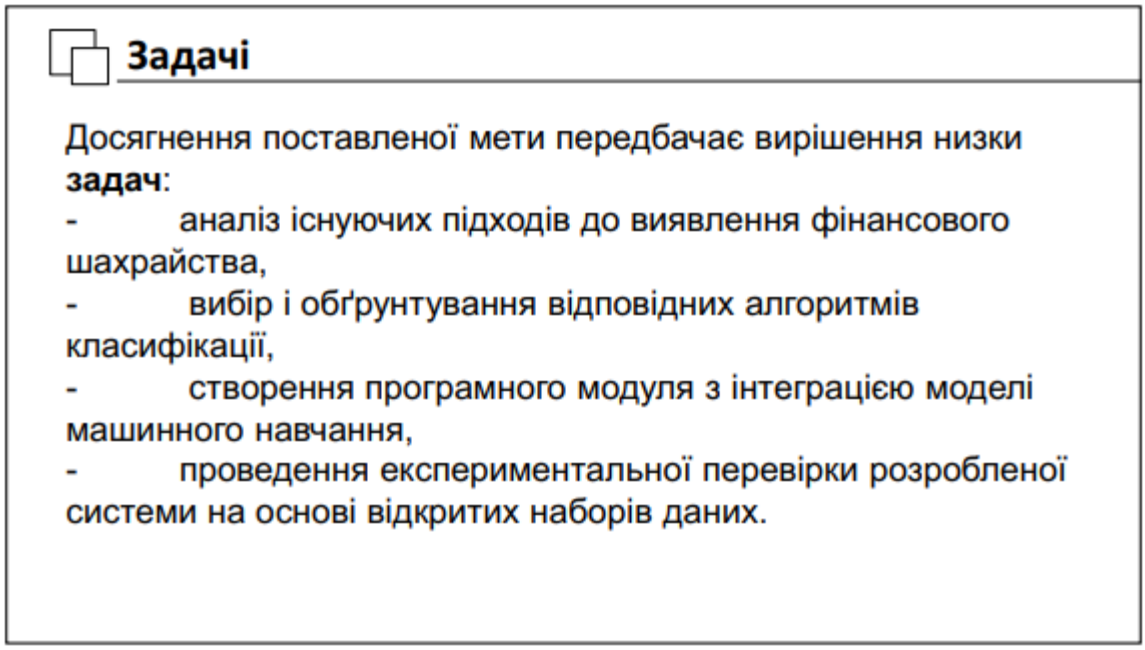
1. Заміна SIM-карти для доступу до онлайн-банкінгу
2. Підставні виграші цінних призів
3. Телефонні дзвінки з метою отримання кодів
4. Допомога НБУ під час карантину
5. Фішинг – підробка веб-сторінок та електронних листів
6. Копіювання даних картки
7. Незахищені мережі Wi-Fi
8. Скімінг
9. Несанкціоновані платежі
10. Крадіжка даних на незахищених сайтах



4


Об’єкт, предмет та мета дослідження
Об’єкт дослідження
автоматизовані інформаційні системи обробки фінансових операцій у банківській сфері, які потребують підвищеного рівня безпеки та контролю.
Предмет дослідження
методи аналізу транзакційної активності клієнтів та побудова моделей виявлення аномалій у фінансових даних.
Мета дослідження
підвищення ефективності виявлення шахрайських дій шляхом створення програмних компонентів, здатних автоматично виявляти нетипову поведінку користувачів на основі історичних даних транзакцій

5


Задачі
Досягнення поставленої мети передбачає вирішення низки **задач**:
- аналіз існуючих підходів до виявлення фінансового шахрайства,
- вибір і обґрунтування відповідних алгоритмів класифікації,
- створення програмного модуля з інтеграцією моделі машинного навчання,
- проведення експериментальної перевірки розробленої системи на основі відкритих наборів даних.

6

## Огляд методів машинного навчання

- Decision Trees
- Naive Bayes Classification
- KNN
- Logistic Regression
- SVM
- Random forest
- CART

7

## Функціональні вимоги

| Актор | Сценарій використання | Функціональні вимоги |
|---|---|---|
| **Користувач / клієнтська програма** | Надсилання даних транзакції до системи | Забезпечити інтерфейс для передачі даних транзакції до системи в реальному часі |
| **Система виявлення шахрайства** | Аналіз отриманої транзакції та визначення рівня ризику | Реалізувати алгоритм оцінки ризику (0–100); виконати класифікацію транзакції як низькоризикової або підозрілої |
| **Система перевірки транзакцій** | Перевірка підозрілих транзакцій | Забезпечити ручну або автоматизовану перевірку транзакцій з високим ризиком |
| **Система зберігання даних** | Збереження результатів аналізу до бази даних | Забезпечити надійне зберігання результатів перевірки, у тому числі для повторного використання |
| **Навчальна система (детектор)** | Перенавчання моделі на основі нових транзакцій | Забезпечити регулярне оновлення моделі машинного навчання на основі перевірених транзакцій |
| **Застосунок адміністратора** | Візуалізація результатів та управління системою | Надати адміністративний інтерфейс для перегляду підозрілих транзакцій, їх статусів, історії рішень та моделі навчання |

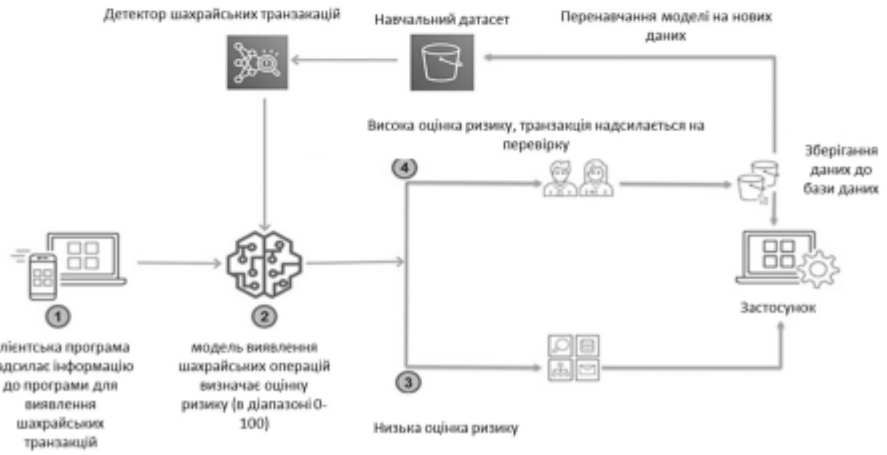
8

## Нефункціональні вимоги

| № | Нефункціональна вимога | Критерій вимірювання | Цільове значення |
|---|---|---|---|
| 1 | Час обробки транзакції | Середній час від моменту надсилання до відповіді | Не більше 1 секунди |
| 2 | Доступність системи | Частка часу, коли система доступна | Не менше 99,9% |
| 3 | Надійність | Частота відмов у роботі системи | Не більше 1 відмова на 10 000 транзакцій |
| 4 | Масштабованість | Можливість обробки зростаючого навантаження | Лінійне масштабування до 10 млн транзакцій/добу |
| 5 | Конфіденційність даних | Наявність механізмів шифрування | Використання SSL/TLS та AES-256 |
| 6 | Безперервність обслуговування | Максимальний час відновлення після збою (RTO) | До 5 хвилин |
| 7 | Зручність інтерфейсу для аналітика | Суб'єктивна оцінка зручності (за шкалою SUS) | Не менше 80 балів |
| 8 | Узгодженість інтерфейсу | Відсутність критичних помилок у навігації | 0 критичних помилок при тестуванні |
| 9 | Логування подій та аудит | Повнота журналу аудиту | 100% транзакцій логуються |
| 10 | Можливість перенавчання моделі | Інтервал між циклами оновлення моделі | Не більше 24 годин |

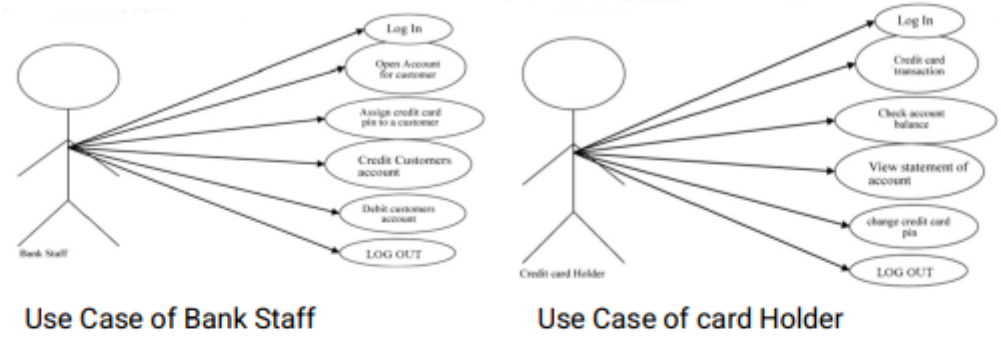
9

## Схема процесу виявлення шахрайських транзакцій




10

# Діаграма варіантів використання

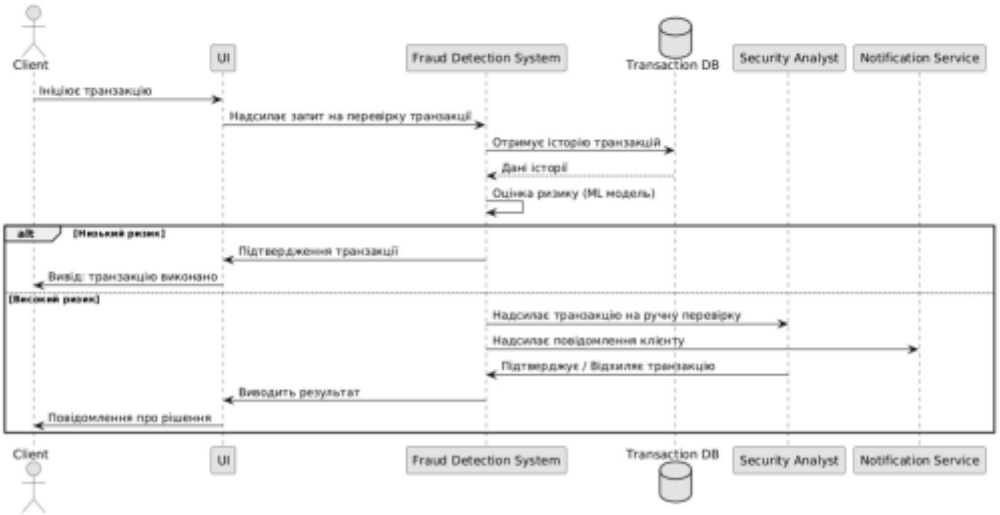



Use Case of Bank Staff

Use Case of card Holder

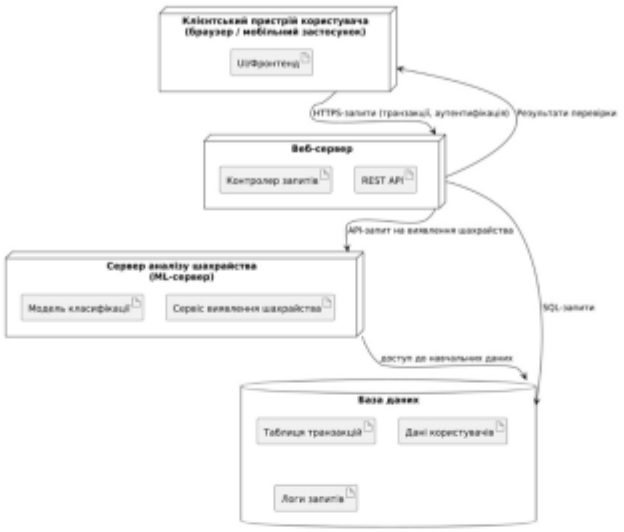
11

# Діаграма послідовності




12

## Діаграма розгортання




13

## Набір даних для моделей

| № | Назва набору даних | Обсяг | Джерело | Кількість записів | Кількість шахрайських транзакцій | Примітки |
|---|---|---|---|---|---|---|
| 1 | Credit Card Fraud Detection | 150.83 МБ | Kaggle, mlg-ulb | 284 807 | 492 | Справжні транзакції; набір є сильно дисбалансованим |
| 2 | Credit Card Fraud (simulated) | 76.28 МБ | Kaggle, dhanushnarayanan r | Не вказано | Не вказано | Симульований набір даних; можливість отримання штучно високої точності |
| 3 | Fraud Detection — Credit Card | 102.92 МБ | Kaggle, yashpaloswal | Менше 284 807 | ~492 | Варіант першого набору без пропущених значень; структура та класовий розподіл збережено |

Для оцінювання здатності моделей до узагальнення, вхідний набір даних було розділено на три підмножини: **навчальну (70%), валідаційну (15%) та тестову (15%).**

14

# Метрики оцінювання моделей

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

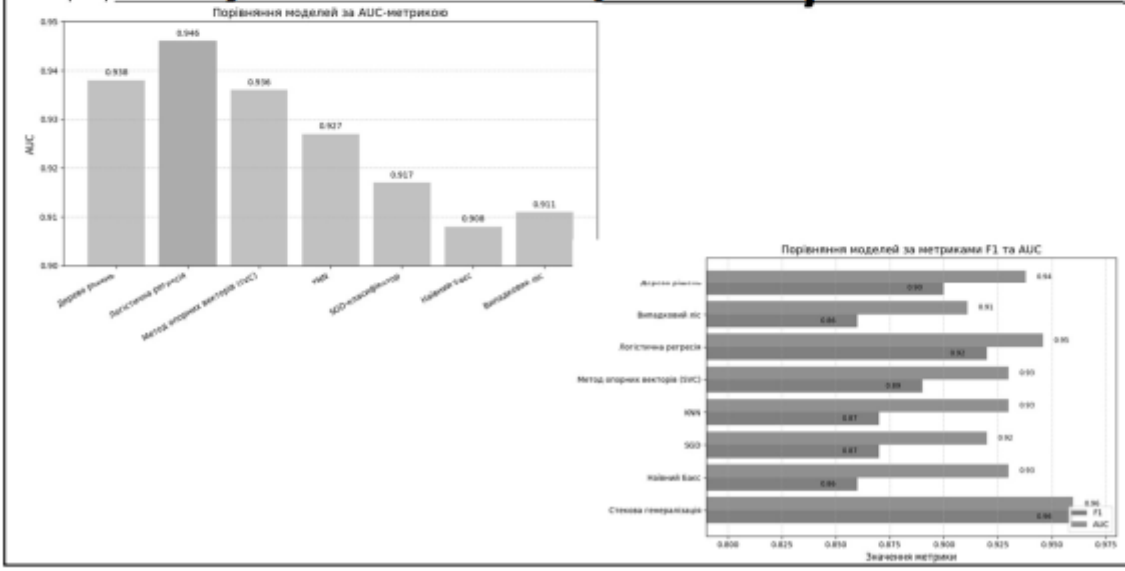
$$Precision = \frac{TP}{TP + FP}$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

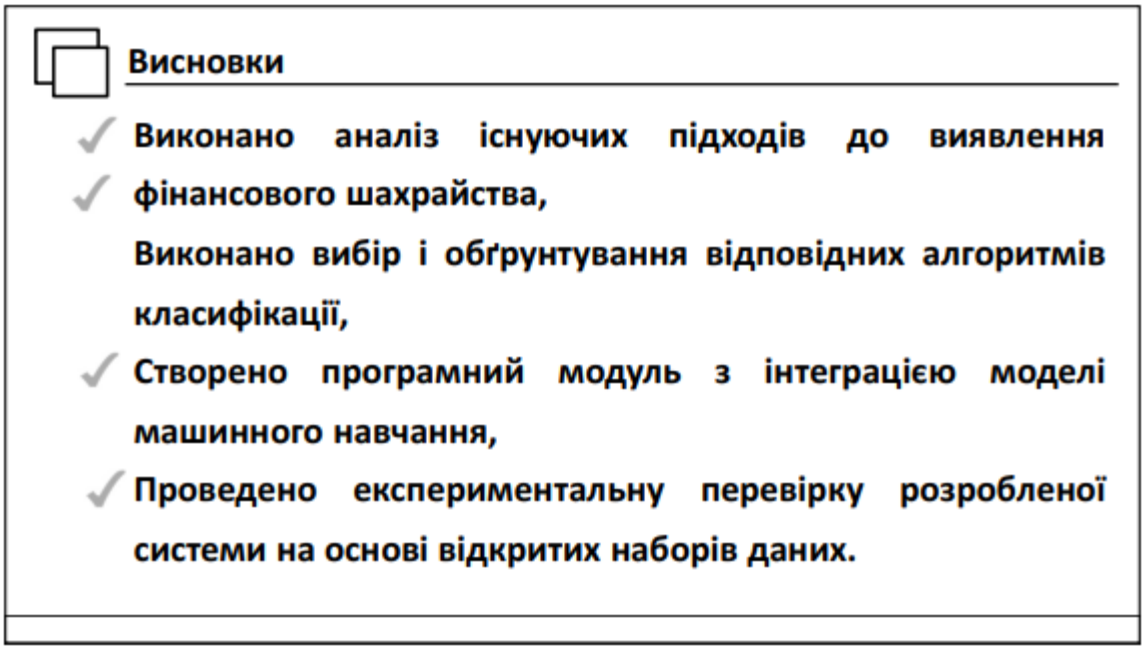
TP- True Positive
TN- True Negative
FP- False Positive
FN- False Negative

15

# Результати експерименту




16


Висновки
Виконано аналіз існуючих підходів до виявлення фінансового шахрайства,
Виконано вибір і обґрунтування відповідних алгоритмів класифікації,
Створено програмний модуль з інтеграцією моделі машинного навчання,
Проведено експериментальну перевірку розробленої системи на основі відкритих наборів даних.
17