Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр
(назва освітнього ступеня)

на тему: Виявлення шахрайських транзакцій з допомогою методів машинного навчання

Виконав: студент II курсу, групи СБм-61
спеціальності 125 Кібербезпека
(шифр і назва спеціальності)

(підпис) Безруков О.О. (прізвище та ініціали)

Керівник (підпис) Загородна Н.В. (прізвище та ініціали)

Нормоконтроль (підпис) Лечаченко Т.А. (прізвище та ініціали)

Завідувач кафедри (підпис) Загородна Н.В. (прізвище та ініціали)

Рецензент (підпис) (прізвище та ініціали)

Тернопіль
2023

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра кібербезпеки
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Загородна Н.В.
(підпис) (прізвище та ініціали)
«____» _____ 2023 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

Студенту Безрукову Олександру Олександровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Виявлення шахрайських транзакцій з допомогою методів машинного навчання

Керівник роботи Загородна Наталія Володимирівна, к.т.н., зав. кафедри КБ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «_16_» _листопада__ 2023 року № _4/7-1061_____

2. Термін подання студентом завершеної роботи 14 грудня 2023р.

3. Вихідні дані до роботи Відкритий набір даних, що містить набір як законних так і шахрайських транзакцій

4. Зміст роботи (перелік питань, які потрібно розробити):

1. Шахрайство як один з найпопулярніших видів загроз для фінансових установ
2. Математичні моделі виявлення шахрайства з допомогою методів машинного навчання
3. Практична реалізація
4. Охорона праці та безпека в надзвичайних ситуаціях

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1 Титульна сторінка. 2 Мета. Завдання дослідження. Наукова новизна
3. Типи шахрайств 4 Статистика в Україні 5 Типи систем виявлення шахрайських транзакцій
6 Системи виявлення аномалій 7 Алгоритми роботи з незбалансованими даними
8 Ілюстрація найпростіших способів усунення незбалансованості даних
9 Переваги Python10 Набір даних 11, 12 Оцінка точності класифікаційних моделей на незбалансованому та збалансованому наборі даних 13 Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	Осухівська Г.М., к.т.н., доцент		
Безпека в надзвичайних ситуаціях	Клепчик В.М., проректор з адміністративно-господарської роботи та будівництва		

7. Дата видачі завдання 16 листопада 2023 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	16.11.2023-17.11.2023	*Виконано*
2.	Підбір наукових джерел про виявлення шахрайських транзакцій	18.11.2023-20.11.2023	*Виконано*
3.	Переклад та опрацювання наукових джерел про		
	дослідження методів виявлення шахрайських		
	Транзакцій методами машинного навчання	21.11.2023-23.11.2023	*Виконано*
4.	Виконання дослідження щодо виявлення		
	шахрайських транзакцій на відкритому наборі		
	даних	24.11.2023-27.11.2023	*Виконано*
5.	Оформлення розділу «Шахрайство як один з найпопу-		
	лярніших видів загроз для фінансових установ»	28.11.2023-30.11.2023	*Виконано*
6.	Оформлення розділу «Математичні основи виявлення		
	Шахрайства з допомогою методів маш.навчання»	01.12.2023-04.12.2023	*Виконано*
7.	Оформлення розділу «Практична реалізація»	05.12.2023-07.12.2023	*Виконано*
8.	Виконання завдання до підрозділу «Охорона праці»	08.12.2023-09.12.2023	*Виконано*
9.	Виконання завдання до підрозділу «Безпека в надзви-		
	чай них ситуаціях»	10.12.2023-11.12.2023	*Виконано*
10.	Оформлення кваліфікаційної роботи	12.12.2023-13.12.2023	*Виконано*
11.	Нормоконтроль	14.12.2023-15.12.2023	*Виконано*
12.	Перевірка на плагіат	09.12.2023	*Виконано*
13.	Попередній захист кваліфікаційної роботи	16.12.2023	*Виконано*
14.	Захист кваліфікаційної роботи	26.12.2023	

Студент _____ Безруков О.О.
(підпис) (прізвище та ініціали)

Керівник роботи _____ Загородна Н.В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Виявлення шахрайських транзакцій з допомогою методів машинного навчання // Кваліфікаційна робота освітнього рівня «Магістр» // Безруков Олександр Олександрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБм-61 // Тернопіль, 2023 // С. 70, рис. – 12, табл. – 2, додат. – 1, бібліогр. – 18.

Ключові слова: АНОМАЛІЇ, МАШИННЕ НАВЧАННЯ, ШАХРАЙСТВО, ТРАНЗАКЦІЇ.

Кваліфікаційна робота присвячена дослідженню методів виявлення шахрайських транзакцій.

У першому розділі проаналізовано основні загрози для бізнесу, описано основні типи шахрайств, проведено огляд існуючих підходів та технічних рішень для виявлення шахрайських транзакцій.

В другому розділі наведено математичні основи виявлення шахрайства з допомогою методів машинного навчання, зокрема існуючі керовані та некеровані методи машинного навчання для виявлення аномалій. У розділі також сформульовано проблеми в системах виявлення аномалій

У третьому розділі обґрунтовано вибір Python як програмного середовища, проведені експериментальні дослідження з використанням різних підходів до виявлення аномалій, наведено оцінку якості використаних моделей.

ANNOTATION

Detection of Fraudulent Transactions Using Machine Learning Methods // Qualification work of the educational level “Master” // Oleksandr Bezrukov // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cybersecurity, SBm-61 group // Ternopil, 2023 // P. 70, fig. - 12, tables - 2, annexes - 1, references - 18.

Key words: ANOMALIES, MACHINE LEARNING, FRAUD, TRANSACTIONS.

The qualification is devoted to the research of methods of detecting fraudulent transactions.

The first chapter analyzes the main threats to business, describes the main types of fraud, and provides an overview of existing approaches and technical solutions for fraudulent transactions.

Second section presents the mathematical foundations of machine learning fraud detection, including existing supervised and unsupervised machine learning methods for anomaly detection. The chapter also formulates challenges in the anomaly detection system.

The third chapter substantiates the choice of Python as a programming environment, conducts experimental studies using various approaches to anomaly detection, and provides an assessment of the quality of the used models.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CPU - Central Processing Unit

ML – Machine Learning

PCA - Principal Component Analysis

SMOTE – Synthetic Minority Oversampling Technique

SVM – Support Vector Machine

БД – База даних

БЖД – Безпека життєдіяльності

МН – Машинне навчання

ПЗ – Програмне забезпечення

ШІ – Штучний інтелект

ЗМІСТ

ВСТУП

Актуальність теми. Не зважаючи на відносно оптимістичні новини звіту “PwC’s Global Economic Crime and Fraud Survey 2022” [1] про те, що менше половини організацій, які брали участь в опитуванні, зустрічались з шахрайством або іншими економічними злочинами протягом останніх 24 місяців, дані іншого дослідження [2] засвідчують, що кількість шахрайських платежів майже подвоїлася з 2014 по 2021 рік. Загалом у 2021 році сума шахрайських платежів за картками досягла 32 мільярдів доларів США. Очевидно така тенденція зумовлена тим, що використання платіжних карток, як способу оплати продовжує зростати в усьому світі, а, отже, зростає і вартість шахрайських операцій з кредитними картками. В Україні згідно з відкритими даними Опендатабот від початку 2023 року значно активізувались шахраї. Аналітики даної платформи стверджують з посиланням на дані Генеральної прокуратури України, що за перші чотири місяці 2023 року зареєстрували більше кримінальних проваджень про шахрайство, аніж за весь 2021 рік [3]. Отже задача своєчасного виявлення шахрайських транзакцій залишається актуальною, особливо в умовах України. Розробка заходів боротьби з шахрайством дозволить зменшити втрати клієнтів фінансових установ та зменшить репутаційні ризики.

Мета і задачі дослідження. Метою даної кваліфікаційної роботи освітнього рівня «Магістр» є побудова системи виявлення шахрайських транзакцій, як системи виявлення аномалій, з допомогою методів машинного навчання .

Для досягнення поставленої мети було потрібно виконати наступні завдання:

- розглянути основні загрози для бізнесу;
- розглянути основні види шахрайства;
- зробити огляд основних існуючих підходів для виявлення шахрайства;

- розглянути теоретичні підходи для виявлення аномалій методами машинного навчання;
- провести експериментальні дослідження на визначення шахрайських транзакцій;
- запропонувати метрики і провести оцінку точності запропонованих моделей.

Об’єкт дослідження. База даних транзакцій.

Предмет дослідження. Методи виявлення шахрайських транзакцій.

Наукова новизна одержаних результатів кваліфікаційної роботи полягає у тому, що отримано результати експериментального дослідження для різних моделей виявлення аномалій та проведено порівняльний аналіз кількох рекомендованих моделей для виявлення шахрайських транзакцій.

Практичне значення одержаних результатів. Розроблені моделі та отримані результати можуть бути впроваджені в банківську систему виявлення шахрайських транзакцій.

Апробація результатів магістерської роботи. Основні результати проведених досліджень обговорювались на XI науково-технічній конференції «Інформаційні моделі, системи та технології» (м.Тернопіль), 13-14 грудня 2023 р.

Публікації. Основні результати кваліфікаційної роботи опубліковано у тезах конференції: О.Безруков, М. Стадник Виявлення шахрайських транзакцій з допомогою методів машинного навчання // XI науково-технічна конференція «Інформаційні моделі, системи та технології» (м.Тернопіль), 13-14 грудня 2023 р. (див. Додаток А).

1 ШАХРАЙСТВО ЯК ОДИН З НАЙПОПУЛЯРНІШИХ ВИДІВ ЗАГРОЗ ДЛЯ ФІНАНСОВИХ УСТАНОВ

У 21 сторіччі насправді важко переоцінити використання інформаційних технологій у всіх сферах життя людини. Пандемія COVID у 2019-2021 роках пришвидшила темпи цифровізації послуг в десятки разів. Але поруч з тим зростають і ризики для людей. Остання кібератака на одного з найбільших мобільних операторів України «Київстар» показала наскільки вразливим стало людство до відсутності мобільних та електронних сервісів. Забезпечення безперебійної роботи онлайн-сервісів, безпечне функціонування об'єктів критичної інфраструктури, надійність офіційних електронних ресурсів ставить питання безпеки одним із пріорітетів для кожної організації, бізнесу, державної установи. В даній роботі ми сконцентруємося роботі фінансових установ. Розглянемо спершу основні кіберзагрози для бізнесу.

1.1 Основні кіберзагрози для бізнесу

Автори [4] розглянули основні загрози для бізнесу. Наведемо тут їх з визначенням основних контзаходів

У звіті Ponemon Institute підкреслюються наступні найбільші ризики безпеки для бізнесу (рис. 1.1):

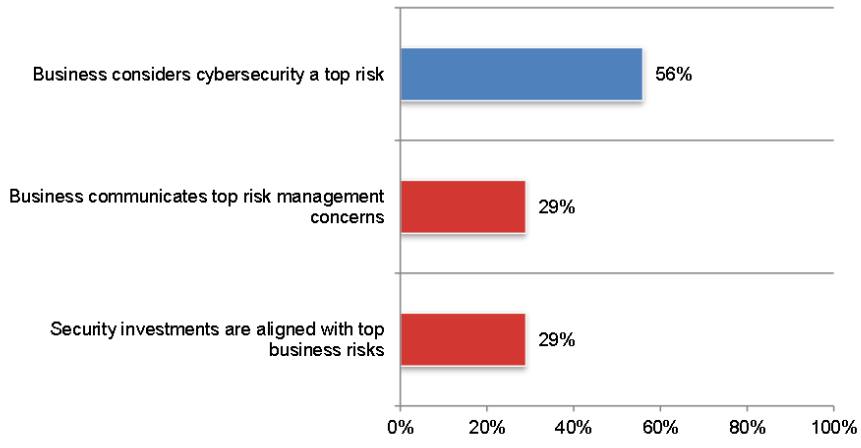

Рисунок 1. 1 - Найбільші ризики безпеки для бізнесу

Точне розуміння визначень загроз безпеці, що призводять до цих бізнес-ризиків, може допомогти компанії бути більш ефективною у розробці корпоративної політики безпеки з метою виявлення та усунення вразливостей і управління бізнес-ризиками. У минулому більшість організацій зосереджувалися на забезпеченні безпеки своїх мереж. Але сьогодні існує більше різноманітних цілей для атак на кібербезпеку. Нижче наведено найпопулярніші загрози безпеці [5] і деякі основні методи їх усунення.

Шкідливе програмне забезпечення

Шкідливе програмне забезпечення (Malware) — це програмне забезпечення, призначене для нанесення шкоди та використання вибраних ресурсів та компонентів комп'ютера. Воно включає широкий спектр різних типів: віруси (viruses), хробаки (worms), клавіатурні шпигуни (keyloggers), трояни (trojans), шпигунське програмне забезпечення (spyware), програмне забезпечення-люки (trapdoors botware) та програми-вимагачі (ransonware).

Вірус — це програма, що саморозмножується, яка, у свою чергу, може заражати інші програми, змінюючи їх. Більшість вірусів використовують особливості та слабкі сторони операційних систем і в деяких випадках можуть підлаштовуватися під певну апаратну платформу.

Черв'як — це програма, що самовідтворюється, яка використовує мережеві з'єднання для міграції з однієї системи в іншу. Після активації в системі черв'як може поводитися як комп'ютерний вірус, створювати трояни, виконувати інші руйнівні чи деструктивні дії.

Keylogger — це програма, яка записує натискання клавіш користувачами.

Троян — це програма, замаскована або імітована під легальні програми, яку часто пропонують користувачам комп'ютерів як подарунок або продаж. Небезпека трояна криється в додатковому командному блоці, якимось чином встановленому в оригінальному легальному програмному забезпеченні.

Шпигунське програмне забезпечення (spyware) — це програма, яка може контролювати діяльність комп'ютера. Trapdoor — це прихована незадокументована точка входу в програмний модуль, яка дозволяє будь-кому, хто про це знає, отримати неавторизований доступ до програми. Botware — це програма, яка підключає комп'ютер до ботнету та дозволяє дистанційно керувати ним. Програми-вимагачі (ransonware) — це криптошкідливі програми, які загрожують опублікувати дані жертви або зашифрувати доступ до інформації, якщо не буде сплачено викуп. Один із відомих троянів-вимагачів WannaCry заразив понад 300 000 комп'ютерів протягом 4 днів і завдав фінансових збитків у 4 мільярди доларів. Програмне забезпечення-вимагач було визнано однією з найбільших шкідливих загроз 2018 року, і воно продовжувало зростати протягом усього 2019 року.

Існують інші типи зловмисного програмного забезпечення, але їх можна розглядати як підтипи, згаданих вище. Потенційні наслідки зараження зловмисним програмним забезпеченням включають: скомпрометовані облікові дані користувача, несанкціонований доступ до приватних файлів, блокування авторизованого доступу користувачів, пошкодження апаратного чи програмного забезпечення. Це може призвести до фінансових, репутаційних, операційних бізнес-ризиків. Щоб захистити комп'ютерну систему компанії від більшості видів шкідливих програм, необхідно використовувати та регулярно оновлювати спеціальні антивірусні програми, здійснювати посередницький доступ до виконуваних файлів, контролювати їх цілісність і цілісність системних областей, тестувати придбане програмне забезпечення, дотримуватися політики паролів, підвищити рівень безпеки співробітників.

Вразливості бездротового Інтернету (Wi-Fi)

Широке використання бездротових мереж в комерційних цілях пояснюється тим, що вони зручні і мають відносно невисоку вартість. Зараз багато малих підприємств пропонують безкоштовне громадське бездротове

підключення до Інтернету, щоб залучити клієнтів. Мета — отримати деякі особисті дані в обмін на безкоштовний Wi-Fi. У той же час більшість підприємств використовують бездротові локальні мережі для власних потреб. Через особливості середовища передачі бездротові мережі мають як переваги, так і недоліки. Наприклад, деякі вразливості мереж Wi-Fi викликані тим, що пакети, передані клієнтом або точкою доступу, можуть бути отримані будь-яким пристроєм у зоні мережі. Основні загрози Wi-Fi включають DDoS-атаки, атаки типу "людина посередині", неправильне налаштування мережі, "випадкові асоціації", слабкі ключі шифрування або методи автентифікації. Користувачі бездротових з'єднань знаходяться під загрозою викрадення їхніх сеансів, залишаючи їхні облікові записи доступними для інших користувачів мережі без їхнього відома, їх використання для анонімного вчинення подальших правопорушень та несанкціонованого доступу до даних, наданих на серверах [6]. За словами відомого експерта з безпеки Брюса Шнайєра, неможливо мати абсолютно безпечне комп'ютерне середовище, але можна принаймні спробувати щось зробити, щоб уповільнити роботу хакерів і користувачів захисту.

Деякі загальні вимоги безпеки до конфігурації мережі Wi-Fi включають: змінити SSID і пароль за замовчуванням, вимкнути трансляцію SSID, увімкнути фільтрацію MAC, вимкнути спільний доступ. Користувачам рекомендується використовувати vpn під час використання загальнодоступного Wi-Fi.

Онлайн-шахрайство (Fraud)

У більшості випадків онлайн-шахрайство пов'язане з людьми, а не з компаніями. Хоча бізнес також міг стати жертвою інтернет-шахраїв. По-перше, скомпрометовані дані кредитної картки можуть використовуватися в Інтернеті для шахрайських транзакцій без картки, коли інформація про обліковий запис використовується без повноважень власника картки. Якщо власник картки доведе факт викрадення його персональних даних, інтернет-

магазин можуть зобов’язати повернути гроші після доставки товару. Отже, бізнес втрачає як гроші, так і речі. Наступний тип атаки пов’язаний із цією, оскільки він фактично дозволяє отримати облікові дані для входу або інформацію про картку.

Фішинг

Фішинг — це схема, яку використовують хакери, щоб змусити користувачів надати облікові дані для входу або особисту інформацію. Зазвичай із жертвами зв’язуються електронною поштою, яка, здається, надходить із надійного джерела, наприклад банку. Спам-повідомлення містить посилання на шахрайський веб-сайт, який видає себе за надійне джерело. Не підозрюючи про будь-який ризик, користувач надає особисту «конфіденційну» інформацію (ім’я користувача, пароль, дані кредитної картки тощо) на підробленому веб-сайті, вважаючи, що він перебуває на законному веб-сайті. Фішинг є прикладом методів соціальної інженерії, які використовуються для обману користувачів. За даними Data Insider, близько 91% витоків даних у мережі відбувається через фішинг.

Spear phishing — це шахрайство електронною поштою або засобами електронного зв’язку, яке здійснюється з метою отримання доступу до комп’ютерної системи компанії. Замість того, щоб націлюватися на велику кількість потенційних жертв, як-от фішингові атаки, цільові фішингові атаки персоналізуються для їхніх жертв (вибирається конкретний власник бізнесу чи працівник). Крім того, оскільки фішингові електронні листи спрямовані невеликій кількості осіб, а не багатьом одержувачам, менш імовірно, що вони будуть виявлені та заблоковані фільтрами електронної пошти [7]. Окрім фішингу електронної пошти та фішингу, існують інші типи фішингових атак: whaling, smishing and vishing, angler phishing.

Основною стратегією захисту від соціальної інженерії та фішингу є безпекова грамотність співробітників і постійне оновлення ними знань у сфері безпеки. Регулярні тренування можуть значно знизити такі ризики. Це також

має відбуватися на всіх рівнях ієрархії компанії – починаючи з керівника компанії і закінчуючи рядовим співробітником. Загальні рекомендації щодо уникнення фішингової атаки також включають такі заходи, як: мати інтелектуальні паролі, часто оновлювати програмне забезпечення, не натискати посилання, отримані в електронних листах, використовувати логіку під час відкриття електронних листів. Може здатися, що досить попросити співробітників бути уважнішими, і фішингу ніколи не станеться, хоча це не так. Відділ безпеки повинен розробити та ретельно впровадити відповідну політику безпеки в компанії.

Скомпрометовані сайти

Важко уявити як корпорації, так і SME (малі та середні підприємства) без веб-додатку. Офіційний веб-сайт для бізнесу – це нагальна і життєво необхідна потреба, оскільки це основний канал взаємодії з клієнтами. У випадку малого та середнього бізнесу власники бізнесу можуть вважати, що їхній веб-сайт занадто малий, щоб стати об’єктом кібератаки. Хоча тенденція показує, що це не так. Недостатня обізнаність про кіберризики та наслідки призводить до ситуації, коли великі та малі компанії можуть зіткнутися з реальними проблемами безпеки через злам їхніх веб-сайтів. У процесі розробки веб-сайту та його оновлення можуть бути допущені ненавмисні помилки, які можуть призвести до компрометації веб-ресурсу. Зламані веб-сервери можуть використовуватися для розміщення заборонених матеріалів неавторизованою особою, доставки зловмисного програмного забезпечення або включення оманливого вмісту. Порушення цілісності веб-сайту несе як репутаційний, так і фінансовий ризик для бізнесу. Якщо веб-сайт компанії не захищений, потенційний хакер може вставити зловмисне програмне забезпечення, щоб відстежувати відвідувачів сайту та таким чином викрадати їх особисту інформацію. Найважливішими загрозами безпеці для бізнес-сайтів є ін’єкції, несправна автентифікація, розкриття конфіденційних даних,

зовнішні об’єкти XML, контроль доступу, міжсайтовий сценарій, недостатнє ведення журналів і моніторинг.

Є кілька простих рекомендацій, наприклад використовувати надійні паролі, використовувати шифрування HTTPS, підтримувати веб-сайт в актуальному стані, регулярно контролювати веб-сайт тощо. Хоча загалом пом’якшення загроз веб-сайту не є тривіальною проблемою, саме тому рекомендується вирішити їх експертам з безпеки.

Атаки на відмову в обслуговуванні

Атаки на відмову в обслуговуванні (DoS) є одним із способів скомпрометувати веб-сайт шляхом перевантаження онлайн-сервісу. Успішна спроба DoS-атак або переповнює веб-сервіси, або призводить до їх збою. Послуги, яких це стосується, можуть включати електронну пошту, веб-сайти, онлайн-акаунти (наприклад, банківські) або інші служби, які залежать від ураженого комп’ютера чи мережі. У цьому випадку законні користувачі не можуть отримати доступ до інформаційних систем, пристроїв або інших мережевих ресурсів. У той час як DoS-атаки використовують один пристрій, підключений до Інтернету, щоб наповнити ціль зловмисним трафіком, DDoS-атака запускається з багатьох скомпрометованих пристроїв, заражених бот-програмами та відомих як ботнет. Технічно кажучи, атаки DoS і DDoS можна розділити на три типи: атаки на основі обсягу, атаки на протокол і атаки на прикладному рівні.

Крім деяких основних рекомендацій щодо захисту бізнес-служб, таких як встановлення та обслуговування антивірусного програмного забезпечення або налаштування брандмауера, компаніям потрібно використовувати більш складні рішення, як-от спеціалізовані служби для виявлення ненормальних потоків трафіку та перенаправлення трафіку зі своєї мережі.

Несанкціонований доступ

Несанкціонований доступ означає отримання доступу до веб-сайту, програми, сервера, сервісу чи іншої системи неавторизованою особою

(зовнішньою чи внутрішньою). Це може статися кількома способами, включаючи зараження зловмисним програмним забезпеченням, отримання паролів, перехоплення мережевого трафіку, використання вразливостей безпеки, соціальну інженерію тощо. Ризики, пов'язані з цими загрозами, включають: порушення даних, збій у роботі комп'ютерних послуг, втрату продуктивності, фінансові та репутаційні втрати.

Надійні паролі, шифрування, двофакторна автентифікація, брандмауер і належна політика безпеки можуть допомогти уникнути бізнес-ризиків, пов'язаних із неавторизованим доступом.

Ризики хмарних обчислень

За даними Forbes, до кінця 2020 року майже 83% компаній у всьому світі все одно будуть використовувати хмари. Хмарні обчислення стосуються доступу до мережевого сховища та додатків онлайн. Деякі переваги використання хмар для бізнесу: низькі експлуатаційні витрати (хмари не вимагають великої обчислювальної потужності обладнання та придбання програмного забезпечення), доступ до інформації з будь-якого місця, великі обчислювальні можливості, стійкість, певний рівень безпеки. Проте визначено такі загрози, які хмарні сервіси становлять для бізнесу: необхідність постійно бути онлайн, залежність від сторонньої компанії, поточні витрати, проблеми з безпекою тощо. Загрози безпеці, пов'язані з хмарними обчисленнями, можна розділити на дві широкі категорії: проблеми безпеки з якими стикаються хмарні постачальники та проблеми безпеки, з якими стикаються їхні клієнти [8]. Фішинг, неправильно налаштовані сервери, зловмисне програмне забезпечення та неавторизований доступ можуть спричинити хмарні інциденти. Ось чому хмарні провайдери повинні дотримуватися міжнародних стандартів – ISO/IEC 27017. У той же час компаніям необхідно прийняти модель нульової довіри, використовувати автоматизацію та штучний інтелект, де це можливо, для вирішення хмарних ризиків.

В даній кваліфікаційній роботі ми зупинимося більш детально на шахрайстві.

1.2 Шахрайство та його види

У світі, який стає все більш цифровим, шахрайство стало однією з найбільш критичних проблем, з якими стикаються підприємства та організації. Оскільки шахраї використовують складні методи для використання фінансових систем і особистих даних, дуже важливо застосовувати гнучкі та передові методи протидії шахрайським діям.

Автори [9] виділяють наступні види шахрайства, наведені на рис. 1.2.

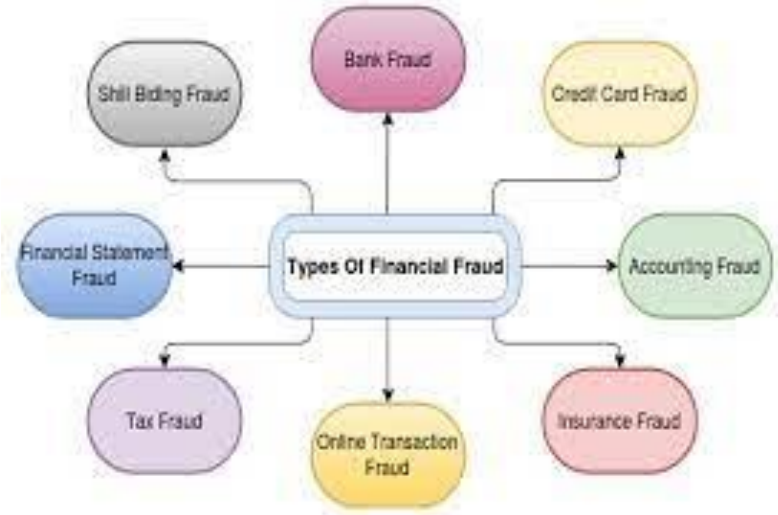

Рисунок 1.2 – Основні види фінансового шахрайства

Отже, до основних видів шахрайства відносять:

Крадіжки особистих даних (Identity Theft)

Це передбачає викрадення чиєїсь особистої інформації (наприклад, номерів соціального страхування, даних кредитної картки або інформації про банківський рахунок) з метою фінансового шахрайства.

Приклад: відкриття кредитних карток або позик на чуже ім'я.

Шахрайство з кредитною карткою (Credit Card Fraud)

Це несанкціоноване використання даних чиєїсь кредитної картки для здійснення покупок або зняття коштів. Для цього можуть використовуватись пристрої для скімінгу в банкоматах чи торгових терміналах або злом онлайн-акаунтів.

Фішинг (Phishing)

Шахраї використовують оманливі електронні листи, повідомлення або веб-сайти, щоб обманом змусити людей надати конфіденційну інформацію, як-от облікові дані для входу чи фінансові дані.

Шахрайство на онлайн-аукціоні (Online Auction Fraud)

Шахрайство в контексті онлайн-аукціонів, де продавець може спотворити товар або не доставити придбаний товар. Приклад: виграш у ставці на товар, якого не існує або який значно відрізняється від опису.

Шахрайство з передоплатою (Advance Fee Fraud)

Жертв просять сплатити наперед, щоб отримати обіцяну послугу, але ця послуга ніколи не надається. Це може буит лотерейне шахрайство, коли жертвам кажуть, що вони виграли приз, але потрібно сплатити комісію, щоб отримати його.

Інвестиційне шахрайство (Investment Fraud)

Шахрайські схеми, які переконують осіб інвестувати у фальшиві або неіснуючі фінансові можливості, наприклад схема, де прибуток попереднім інвесторам оплачується капіталом нових інвесторів.

Страхове шахрайство (Insurance Fraud)

Навмисне інсценування неправдивих претензій до страхових компаній для отримання виплат або перебільшення збитків, завданих майну, щоб вимагати страхування.

Податкове шахрайство (Tax Fraud:)

Надання неправдивої інформації в податкових деклараціях з метою ухилення від сплати податків. Для прикладу, заниження доходів, завищення відрахувань або приховування грошей на офшорних рахунках.

Шахрайство в галузі охорони здоров'я (Healthcare Fraud)

Шахрайські дії в галузі охорони здоров’я, як-от завищення рахунків, виставлення рахунків за ненадані послуги або схеми відкатів.

Благодійне шахрайство (Charity Fraud)

Використання щедрості людей шляхом створення фальшивих благодійних організацій або недобросовісне використання коштів, призначеними на законні благодійні цілі (особливо популярне під час війни). Це може бути: збір пожертв у фальшивий фонд допомоги постраждалим від стихійних лих, допомога важко-хворим людям, збір на військові потреби.

Важливо залишатися пильним і вживати заходів для захисту особистої та фінансової інформації, щоб не стати жертвою шахрайства. Крім того, отримання інформації про нові тенденції шахрайства та регулярний моніторинг фінансових рахунків можуть допомогти окремим особам і компаніям виявляти та запобігати шахрайству. Автори [10] зробили класифікацію основних видів шахрайства в залежності від виду фінансової установи. Результати їх роботи можна побачити на рисунку

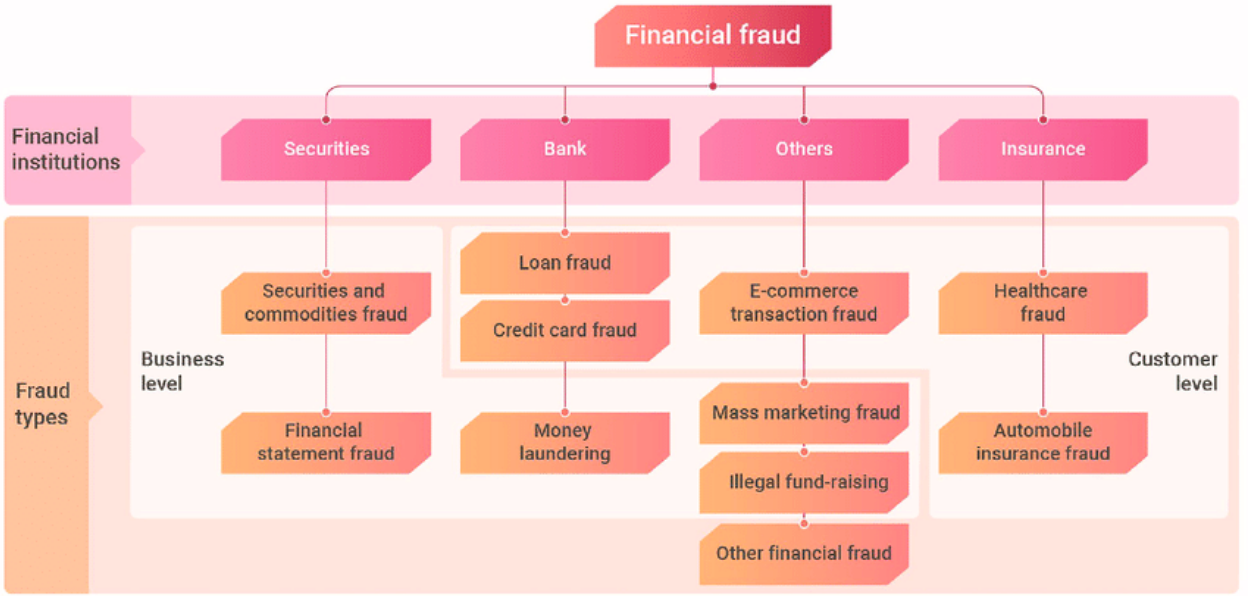

Рисунок 1.3 - Класифікація основних видів шахрайства в залежності від виду фінансової установи

Аналітика даних про шахрайство стала потужним інструментом у боротьбі з шахрайством, надаючи організаціям нове уявлення про потенційні ризики та дозволяючи їм передбачати, виявляти та запобігати шахрайським транзакціям. Використовуючи величезні обсяги даних, отриманих під час повсякденних операцій, аналітика даних про шахрайство досліджує закономірності та тенденції шахрайської поведінки, дозволяючи компаніям посилити заходи безпеки та захистити свої фінансові інтереси. У цій кваліфікаційній роботі ми дослідимо, як аналітику даних можна ефективно використовувати для виявлення та запобігання шахрайству, тим самим захищаючи організаційні ресурси та забезпечуючи довіру клієнтів.

У 2021 році Федеральна торгова комісія (FTC) надіслала майже 390 000 звітів про шахрайство з кредитними картками, що робить його одним із найпоширеніших видів шахрайства в США, але ця цифра не дає повного уявлення про проблему.

У грудні 2022 року звіт Нільсона, який займається моніторингом індустрії платежів, оприлюднив прогноз, згідно з яким збитки США від шахрайства з картками становитимуть 165,1 мільярда доларів США протягом наступних 10 років, що торкнеться кожної вікової групи в кожному штаті. Згідно з даними Insider Intelligence, лише один тип шахрайства з кредитними картками — шахрайство без використання картки, яке включає транзакції онлайн, по телефону та поштою — спричинив приблизно 5,72 мільярда доларів США втрат у 2022 році. Тож розглянемо більш детально способи виявлення шахрайських транзакцій.

1.3 Огляд існуючих підходів та рішень для виявлення шахрайських транзакцій

Рішення для виявлення шахрайства з кредитними картками – це програмні інструменти, розроблені для виявлення шахрайських дій,

пов’язаних з транзакціями кредитних і дебетових карток. Ці рішення використовують алгоритми машинного навчання та розширену аналітику для виявлення шаблонів і аномалій у даних транзакцій. Вони допомагають фінансовим установам відстежувати транзакції в реальному часі, виявляти підозрілі дії та запобігати обробці шахрайських транзакцій.

Рішення для виявлення шахрайства з кредитними картками в основному використовують три різні системи виявлення:

- Системи на основі правил (Rule-based systems). Системи на основі правил використовують попередньо визначені правила для виявлення шахрайських дій. Ці системи мають попередньо визначені правила для транзакцій, які вважаються підозрілими, і будь-які транзакції, які відповідають цим критеріям, позначаються для перевірки.
- Системи виявлення аномалій (Anomaly detection systems). Системи виявлення аномалій використовують алгоритми машинного навчання для виявлення підозрілих дій, які відрізняються від звичайних моделей транзакцій. Ці системи аналізують дані минулих транзакцій, щоб виявити закономірності та створити основу для нормальної поведінки. Будь-яке відхилення від базової лінії позначається як підозріле.
- Системи прогнозного моделювання (Predictive modeling systems). Системи прогнозного моделювання використовують передову аналітику для прогнозування шахрайських дій. Ці системи використовують історичні дані для створення моделей для виявлення потенційних шахрайських дій у режимі реального часу. Вони також можуть створювати сповіщення про певні типи шахрайських дій.

Серед найкращих технічних рішень для виявлення шахрайських транзакцій у 2023 році згідно [11] називають:

- Accertify - компанія American Express є провідним постачальником рішень із запобігання шахрайству, що пропонує наскрізні рішення для боротьби з шахрайством, включаючи автентифікацію клієнтів, оптимізацію

SCA та керування поверненням платежів. Ключовий компонент виявлення шахрайства Accertify Digital Identity використовує алгоритми машинного навчання для використання баз даних аналітики поведінки користувачів і інтелектуальних даних пристроїв. Завдяки такому підходу Accertify може точно позначати користувачів.

- CybeReady це нетрадиційне рішення, яке зосереджено на створенні культури обізнаності про безпеку серед співробітників у різних галузях, включаючи фінанси, фармацевтику та виробництво. Компанія проводить навчання з кібербезпеки та програми підвищення обізнаності, щоб допомогти організаціям запобігти шахрайству з кредитними картками та підготуватися до відповідності PCI та перевірок. Надаючи співробітникам на всіх рівнях організації знання щодо виявлення та запобігання шахрайству, CybeReady допомагає створити додатковий рівень безпеки, створюючи спільне відчуття відповідальності за стан безпеки організації.

- Cybersource - пропонує чотири основні модулі для боротьби з платіжним шахрайством, усі вони працюють на основі програмного забезпечення Visa Decision Manager. Його уніфіковане платіжне рішення дозволяє організаціям переходити між платіжними моделями, такими як BNPL, і використовувати нові форми платежів із надійним захистом від шахрайства. Cybersource також має деталізовані модулі, які допомагають торговцям розширити охоплення, з можливостями конвертації валют, дотримання глобального податкового законодавства та керування терміном служби клієнтів.

- Ekata - завдяки своїм рішенням API Ekata надає послуги з підтвердження особи та запобігання шахрайству в усьому світі, зокрема Transaction Risk API, Address Risk API та Phone Intelligence API. Крім того, постачальник засобів боротьби з шахрайством надає два інтерфейси API, спеціально розроблені для аналітиків шахрайства: Merchant і Account; та інструмент Pro Insight.

- Kount - це рішення корпоративного рівня для виявлення шахрайства, яке пропонує організаціям все необхідне для моніторингу та захисту від шахрайства CNP (Card Not Present). Його настроювані правила автоматизують запобігання відкликанню платежів і скорочують час перевірки вручну. Глобальна мережа Kount Identity Trust щорічно збирає дані про 32 мільярди взаємодій користувачів, що робить її найкращим вибором для ідентифікаційної перевірки.

Основні виклики при виявленні шахрайських транзакцій:

- Зростання витрат

Щоб йти в ногу з часом та вчасно реагувати на постійне зростання переліку методів і типів шахрайства, компанії повинні інвестувати в сучасні інструменти для виявлення шахрайства. Покладатися виключно на моніторинг шахрайських транзакцій на основі правил може бути проблемою, оскільки методи шахрайства змінюються.

- Віддалене здійснення транзакцій

Бізнес все частіше може працювати без фізичної взаємодії. Хоча це зручно та економічно ефективно, це також відкриває двері для шахраїв, які можуть видавати себе за справжніх клієнтів або перехоплювати їхні дані.

- Швидкість транзакцій

Сучасна екосистема транзакцій створена для швидкості та зручності. Навіть такий відносно складний процес, як заявка на кредит, можна здійснити за допомогою смартфона, тоді як більш рутинні покупки завершуються кількома натисканнями клавіш. Це високошвидкісне середовище з низьким рівнем тертя може полегшити шахраям завершити свої злочини та зникнути до того, як їх вдасться виявити.

- Помилкові спрацьовування

Надмірно старанна система виявлення шахрайства може призвести до більшої кількості помилкових спрацьовувань. Це незручно для клієнтів, які в

результаті можуть стати менш лояльними, і дорого для компаній, які повинні витрачати час і ресурси на обробку попередження.

- Еволюція методів шахрайства.

Шахраї постійно розробляють нові методи, щоб уникнути систем виявлення. Традиційним системам, заснованим на правилах, може бути важко пристосуватися до шахрайства, що розвивається, і вимагають постійних оновлень, щоб залишатися ефективними. Впровадження алгоритмів виявлення шахрайства в реальному часі вносить додаткову складність, оскільки рішення повинні прийматися швидко без шкоди для точності.

- Діапазон типів транзакцій

Зараз для переміщення грошей використовується величезна кількість інструментів і сервісів, від платіжних програм і платформ для торгівлі криптовалютою до традиційних позик, кредитних карток і ощадних рахунків. Поширення сфери цифрових фінансових послуг, зокрема, створює численні потенційні точки доступу для шахраїв.

- Змагальні атаки

Шахраї можуть навмисно маніпулювати даними або використовувати змагальні методи, щоб ввести в оману моделі машинного навчання. Це може призвести до погіршення якості моделі або неправильної класифікації, якщо система виявлення недостатньо надійна.

- Відповідність нормативним вимогам безпеки

Дотримання нормативних вимог, таких як GDPR або інших законів про конфіденційність, під час впровадження заходів виявлення шахрайства додає складності. Збалансувати потребу в надійній безпеці з проблемами конфіденційності є делікатним завданням.

- Інтеграція з існуючими системами.

Інтеграція систем виявлення шахрайства з існуючою інфраструктурою та програмним забезпеченням може бути складною. Можуть виникати проблеми сумісності програмного забезпечення, даних.

1.4 Висновок до першого розділу

В даному розділі розглянуто основні кіберзагрози для бізнесу та проведено огляд найбільш типових рекомендацій для їх попередження. Виявлено, що шахрайство є одним з найбільших векторів роботи кіберзлочинців. В розділі також розглянуто типи та приклади шахрайства та обґрунтовано необхідність створення нових та удосконалення існуючих методів та засобів боротьби з шахрайством. Шахрайські транзакції належать до топ-трьох за розміром нанесених збитків видів шахрайства. В розділі наведено основні підходи та виклики в роботі систем для виявлення шахрайських транзакцій.

Ця тематика стала продовженням кваліфікаційної роботи бакалавра, де, проте, більше уваги було приділено особливостям функціонування електронного банкінгу та проєктуванню архітектури програмного забезпечення (ПЗ). В даній кваліфікаційній роботі ми зосередимося більше на виявленні шахрайських транзакцій методами Anomaly Detection

2 МАТЕМАТИЧНІ ОСНОВИ ВИЯВЛЕННЯ ШАХРАЙСТВА З ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

2.1 Машинне навчання, його типи

Артур Самуел ще в 1959 році визначив машинне навчання як область дослідження, яка дає комп'ютеру здатність навчатися з даних, не будучи явно запрограмованим. Зараз IBM дає наступне означення машинного навчання: Машинне навчання – це галузь штучного інтелекту (ШІ) та інформатики, яка зосереджується на використанні даних і алгоритмів для імітації способу навчання людей, поступово покращуючи його точність. А Amazon наводить твердження, що моделі машинного навчання генерують прогноз, знаходячи закономірності та шаблони в даних.

Загалом виділяють класичне навчання, ансамблеві методи, глибоке навчання та навчання з підкріпленням (див.рисунок 2.1)

Класичне навчання в свою чергу поділяється на кероване (Supervised), некероване (Unsupervised). Як ілюстративний приклад, розглянемо завдання навчитися виявляти спам електронної пошти проти завдання виявлення аномалій. Для завдання виявлення спаму ми розглядаємо налаштування, за яких «учень» отримує навчальні електронні листи з міткою «спам/не спам». На основі такого навчання «учень» повинен визначити правило позначення нових отриманих електронних листів. Навпаки, для завдання виявлення аномалії все, що «учень» отримує під час навчання, — це велика кількість повідомлень електронної пошти (без міток), а завдання «учня» — виявити «незвичайне» повідомлення.

Більш формалізовано, якщо розглядати навчання як процес «використання досвіду, щоб отримати експертизу», то кероване навчання описує сценарій, в якому досвід описується навчальним датасетом, що містить важливу інформацію (бажаний результат, мітку), якої, натомість немає в

майбутніх тестових прикладах, до яких планується застосовувати «навчений досвід» В такому випадку може розглядати середовище як вчителя, який керує навчанням, надаючи додаткову інформацію (мітки). В некерованому навчанні відсутня різниця між навчальним та тестовим датасетом. В такому випадку навчання полягає в тому, щоб отримати коротшу (стиснену) версію даних або ж прийти до певного висновку на основі даних.

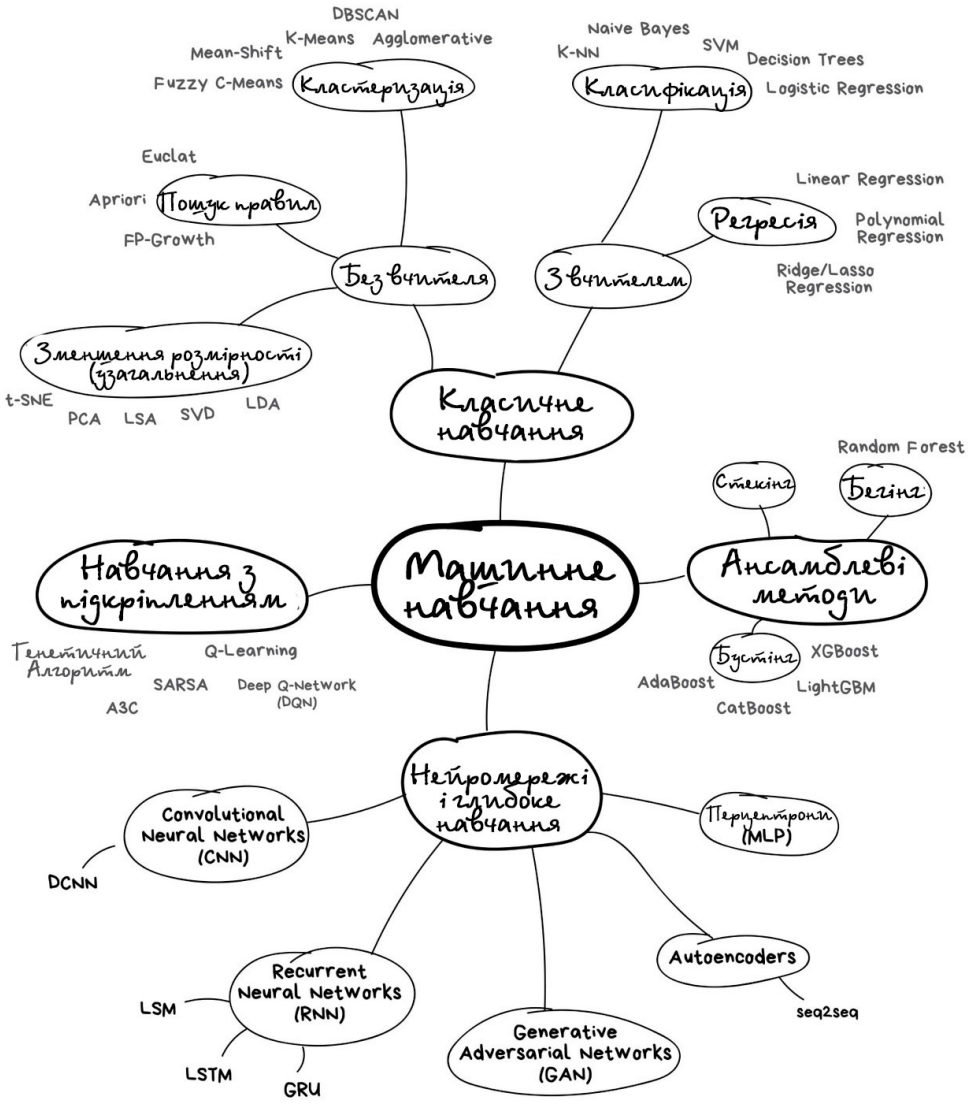

Рисунок 2.1 – Класифікація методів машинного навчання

Навчання з підкріпленням розклядає навчальну систему як агента, що може спостерігати за середовищем, вибирати і виконувати дії та отримати винагороду в підсумку.

Ансамблеві методи – це техніка машинного навчання, яка поєднує кілька базових моделей для створення однієї оптимальної прогнозної моделі.

Глибоке навчання — це галузь машинного навчання, яка базується на штучних нейронних мережах. Він здатний вивчати складні моделі та зв'язки в даних. У глибокому навчанні нам не потрібно явно програмувати все. Останніми роками він стає все більш популярним завдяки прогресу в обчислювальній потужності та доступності великих наборів даних. Оскільки він заснований на штучних нейронних мережах (ANN), також відомих як глибокі нейронні мережі (DNN).

В [12] наведено основні етапи керованого навчання (рис.2.2).

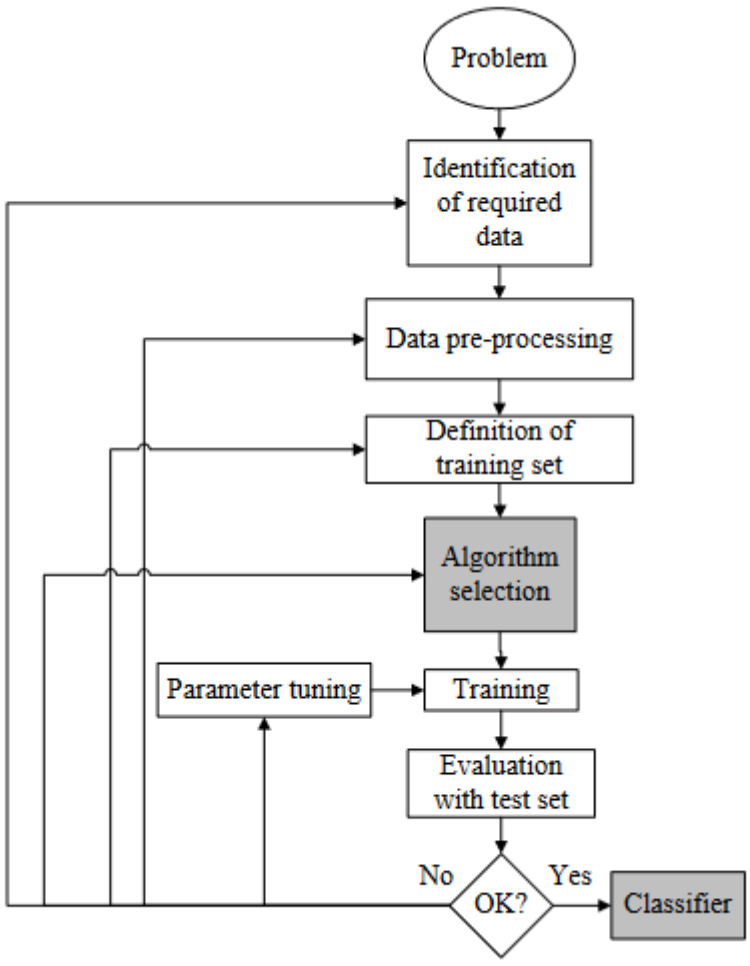

Рисунок 2.2 – Ілюстрація алгоритму роботи керованого навчання

Першим кроком є збір набору даних. Вдалою ідеєю є залучити експерта, який зможе підказати, які поля (атрибути, ознаки) є найбільш інформативними. Якщо такої можливості немає, то найпростішим методом є метод «грубої сили», що означає вимірювання всіх доступних параметрів в надії, що правильні (інформативні, релевантні) ознаки можна виділити. Однак набір даних, зібраний методом грубої сили, не підходить безпосередньо для індукції. У більшості випадків він містить шум і відсутні значення ознак, тому вимагає значної попередньої обробки.

Іншим способом зменшення розміру даних є виділення значущих ознак або їх перетворення з метою зменшення розмірності вхідного простору. Вибір підмножини ознак є процесом ідентифікації і видалення максимальної кількості нерелевантних і зайвих наскільки це можливо. Це зменшує розмірність даних і дозволяє алгоритмам аналізу даних працювати швидше та ефективніше. Той факт, що багато ознак залежать одна від одної, часто створює надмірність, що негативно впливає на точність та продуктивність контрольованих моделей класифікації ML. Цю проблему можна вирішити шляхом створення нових функцій із базового набору ознак. Ця техніка називається конструювання/перетворення ознак. Ці нові згенеровані ознак можуть призвести до створення нових лаконічних та точніих класифікаторів.

Розглянемо більш детальніше напрямки класичного навчання. Кероване навчання включає в себе два основні напрямки – класифікацію та регресію.

Другим кроком є підготовка даних і попередня обробка даних. Залежно від обставин дослідники мають на вибір кілька методів обробки відсутніх даних, викидів, шумів, помилкових даних. Вибір екземплярів з датасету є фактично задачею мінізації обсягу даних, що дозволяє ефективно працювати з великими обсягами даних, при цьому забезпечуючи необхідний рівень точності.

Етап вибору алгоритму в машинному навчанні є ключовим кроком у створенні моделі, яка ефективно вирішує конкретну задачу. Він передбачає

тестування алгоритмами класичного машинного навчання та глибинного навчання в залежності від характеристик задачі та обсягу даних. Попередньо потрібно обрати метрики для тестування, наприклад, точність, F1-мера для класифікації, середньоквадратична помилка для регресії.

На наступному етапі відбувається запуск базових алгоритмів на невеликій частині даних та оцінка їхньої ефективності за допомогою визначених метрик.

Підбір гіперпараметрів моделі проводиться для отримання оптимальної продуктивності.

2.2 Методи для виявлення аномалій

Одним з підходів з підходів для створення систем для боротьби з шахрайскими транзакціями є системи, побудовані на принципах виявлення аномалій серед пулу транзакцій.

2.2.1 Статистичні методи виявлення аномалій

У статистичних методах виявлення аномалій система спостерігає за діяльністю суб’єктів і створює профілі для представлення їх поведінки. Профіль зазвичай включає такі показники, як показник інтенсивності діяльності, показник розподілу записів аудиту, категоріальний показник (розподіл діяльності за категоріями) і порядковий показник (наприклад, використання CPU). Як правило, для кожного суб’єкта підтримуються два профілі: поточний і збережений профіль. У міру обробки системних/мережевих подій (зокрема, записів журналу аудиту, вхідних пакетів тощо) система виявлення вторгнень оновлює поточний профіль і періодично обчислює показник аномалії (що вказує на ступінь нерегулярності для конкретної події) шляхом порівняння поточного профілю зі збереженим профілем за допомогою функції ненормальності всіх показників у профілі.

Якщо оцінка аномалії перевищує певний поріг, система виявлення вторгнення генерує сповіщення. Статистичні підходи до виявлення аномалій мають ряд переваг. По-перше, ці системи, як і більшість систем виявлення аномалій, не вимагають попереднього знання недоліків безпеки та/або самих атак. У результаті такі системи мають можливість виявляти «нульовий день» або найновіші атаки. Крім того, статистичні підходи можуть забезпечити точне сповіщення про зловмисну діяльність, яка зазвичай відбувається протягом тривалих періодів часу та є хорошими показниками загрозливих атак типу «відмова в обслуговуванні» (DoS).

Однак статистичні схеми виявлення аномалій мають і недоліки. Досвідчені зловмисники можуть навчити статистичне виявлення аномалій сприймати ненормальну поведінку як нормальну. Також може бути важко визначити порогові значення, які врівноважують ймовірність помилки позитивні з ймовірністю помилкових негативів. Крім того, статистичні методи потребують точних статистичних розподілів, але не всю поведінку можна моделювати за допомогою суто статистичних методів. Насправді більшість запропонованих статистичних методів виявлення аномалій вимагають припущення про квазістаціонарний процес, що неможливо припустити для більшості даних, оброблених системами виявлення аномалій.

Більшість команд мають набори даних, які вони використовують для навчання алгоритму машинного навчання для виявлення аномальних даних. Від того, чи позначено дані відповідними класами в цих наборах зразків, визначається, який із двох основних типів систем виявлення аномалій можна застосовувати — контрольоване чи неконтрольоване.

2.2.2 Виявлення аномалій методами методами керованого навчання

Контрольоване виявлення аномалій — це підмножина виявлення аномалій, яка передбачає використання позначених даних для навчання моделі, яка може точно виявляти аномалії. У контрольованих умовах алгоритм

має доступ як до нормальних, так і до ненормальних даних під час фази навчання. Набори даних містять попередньо визначені нормальні дані та чітко позначені приклади аномалій. Алгоритм вчиться ідентифікувати закономірності в нормальних даних, а потім може розпізнавати відхилення від цієї моделі в аномальних даних.

2.2.1.1 Незбалансованість даних

Однією з головних проблем у контрольованому виявленні аномалій є робота з незбалансованими даними. У багатьох випадках кількість нормальних зразків значно перевищує кількість аномальних зразків, що призводить до незбалансованого розподілу даних. Це може створити проблеми для алгоритму навчання, оскільки він може стати упередженим до більшості класів (нормальні зразки) і не зможе точно визначити менший клас (ненормальні зразки). Існують різні методи, які можна використовувати для вирішення проблеми незбалансованих даних. Один із підходів полягає в повторній вибірці даних, щоб збалансувати класи. Це може включати або надмірну вибірку для класу меншості, або недостатню вибірку для класу більшості.

Недостатню вибірку можна визначити як видалення деяких спостережень основного класу. Це робиться до тих пір, поки більшість і меншість не будуть збалансовані. Недостатня вибірка може бути хорошим вибором, якщо у вас є тонна даних — подумайте про мільйони рядків. Але недоліком недостатньої вибірки є те, що ми видаляємо інформацію, яка може бути цінною. Надмірну вибірку можна визначити як додавання більшої кількості копій до класу меншості. Надмірна вибірка може бути хорошим вибором, коли у вас немає тони даних для роботи. Мінус, який слід враховувати під час недостатньої вибірки, полягає в тому, що це може спричинити перенавчання та погане узагальнення вашого тестового набору.

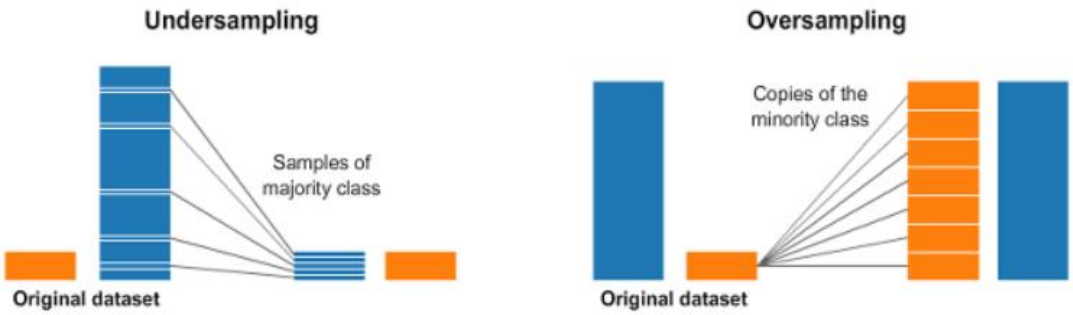

Рисунок 2.3 – Ілюстрація надмірної та недостатньої вибірки

Існують інші методи створення надмірної чи недостатньої вибірки. Наприклад, зв’язки Томека (Tomek links) — це пари дуже близьких екземплярів, але протилежних класів. Видалення екземплярів основного класу кожної пари збільшує простір між двома класами, полегшуючи процес класифікації. Зв’язок Томека існує, якщо два зразки є найближчими сусідами один одного (рис.2.4).

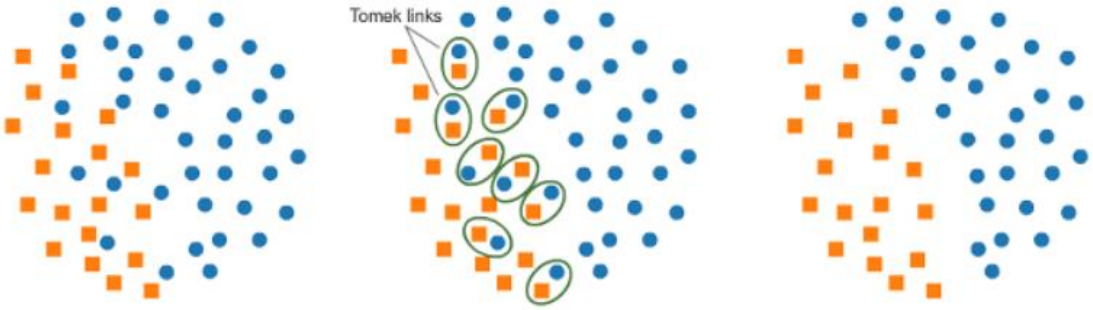

Рисунок 2.4 – Створення недостатньої вибірки через зв’язки Томека

Техніка синтетичного створення надмірної вибірки з меншості (Synthetic Minority Oversampling Technique -SMOTE) .Цей метод генерує синтетичні дані для класу меншості. SMOTE працює шляхом випадкового вибору точки з класу меншості та обчислення k-найближчих сусідів для цієї точки. Синтетичні точки додаються між вибраною точкою та її сусідами (рис. 2.5).

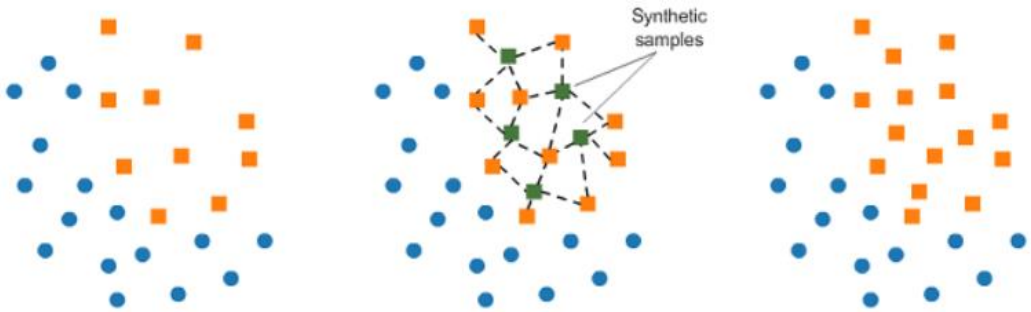

Рисунок 2.5 – Створення надмірної вибірки технікою SMOTE

Інший підхід полягає у використанні алгоритмів, спеціально розроблених для обробки незбалансованих даних, таких як економічне навчання, яке полягає у використанні штрафних алгоритмів навчання, які збільшують вартість класифікаційних помилок у класі меншості. (Penalized-SVM) або ансамблеві методи.

2.2.1.2 Методи керованого класичного навчання для роботи з незбалансованими даними

В [13] було проведено огляд та порівняння наступних методів керованого навчання для виявлення шахрайських транзакцій, а саме: логістичної регресії, методу опорних векторів, випадкових лісів та моделі наївного Баєса. Проте в даній публікації не було проведено узагальненого аналізу згаданих методів в ракурсі однакових метрик, що ускладнює порівняння.

Розглянемо коротко кожен з методів.

Метод опорних векторів (SVM) працює подібно до лінійного дискримінантного аналізу. Створюється гіперплощина або набір гіперплощин, щоб розділити вектори ознак на кілька класів, як LDA, але вибирається гіперплощина, яка знаходиться на максимальній відстані від найближчих навчальних зразків. SVM знаходить гіперплощину з максимальним запасом згідно з теоремою Ковера, відображаючи вхідні дані у багатовимірному просторі (рис.2.6). Теорема Ковера стверджує, що коли складну задачу

класифікації викладено у високовимірному нелінійному просторі, вона, швидше за все, буде лінійно роздільною, ніж при відведенні в низьковимірному нелінійному просторі. SVM також використовує регуляризацію для запобігання артефактам. Нелінійна SVM також може існувати з нелінійною межею прийняття рішень за допомогою функції ядра, що забезпечує більшу гнучкість. SVM широко використовується в системах BCI, оскільки він швидкий, надійніший, а також простий.

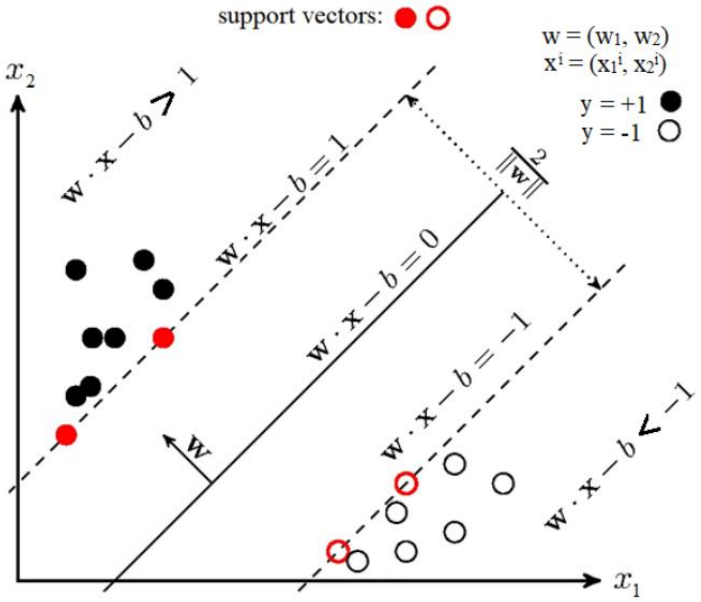

Рисунок 2.6 – Ілюстрація методу опорних векторів

Наївний Байєс — це ймовірнісний алгоритм машинного навчання на основі теореми Байєса, який використовується в багатьох класифікаційних завданнях.

Оскільки наївний класифікатор Байєса передбачає, що всі змінні є незалежними, для оцінки параметрів, необхідних для класифікації, необхідна лише невелика кількість навчальних даних. У контексті діагностики несправностей класи представлятимуть несправності або набір несправностей, які може розвинути система, а предиктори представлятимуть симптоми, які представляє система. Незважаючи на те, що наївні класифікатори Байєса легко і швидко реалізувати, розгляд предикторів як незалежних змінних можна

розглядати як недолік методу, оскільки в більшості випадків реальної діагностики несправності симптоми можуть залежати один від одного.

Логістична регресія, незважаючи на свою назву, є моделлю класифікації, а не моделлю регресії. Логістична регресія є простим і більш ефективним методом для задач бінарної та лінійної класифікації. Це модель класифікації, яку дуже легко реалізувати та досягає дуже високої продуктивності з лінійно роздільними класами. Це широко використовуваний алгоритм для класифікації в промисловості.

Випадковий ліс — це контрольований алгоритм навчання. «Ліс», який він будує, являє собою ансамбль дерев рішень, зазвичай навчених методом пакетування. Загальна ідея методу bagging полягає в тому, що поєднання моделей навчання підвищує загальний результат. На рисунку 2.7 наведено схематична ілюстрацію роботи методу випадкових лісів

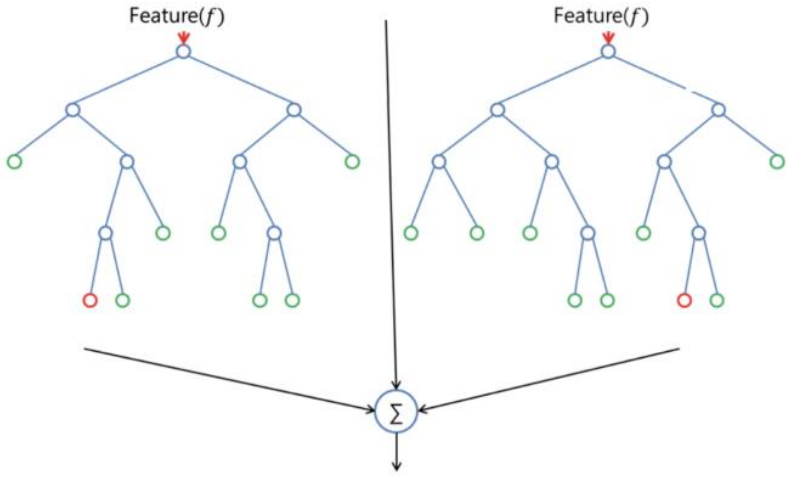

Рисунок 2.7 – Ілюстрація методу випадкових лісів

Хоча це може зробити платформу виявлення аномалій кращою для виявлення очікуваних аномалій у даних, вона не враховуватиме аномалії, яких не бачили раніше. Крім того, багато позначених наборів даних не містять достатньо аномальних даних для ефективного навчання алгоритму.

2.3 Некеровані методи навчання для виявлення та роботи з аномаліями

Цей метод вимагає будь-яких навчальних даних і натомість передбачає дві речі щодо даних, тобто лише невеликий відсоток даних є аномальним і будь-яка аномалія статистично відрізняється від звичайних зразків. Виходячи з наведених вище припущень, дані потім кластеризуються за допомогою міри подібності, а точки даних, які знаходяться далеко від кластера, вважаються аномаліями.

Програми для неконтрольованого навчання можна розділити на дві основні групи: кластеризація та асоціація. Асоціація: метою є виявлення та вивчення репрезентативних правил у наборі даних; наприклад, що клієнти, які купують продукт A, також схильні купувати продукт B. Кластеризація. Програми кластеризації прагнуть вивчити властиві групування даних, наприклад диференціацію сегментів клієнтів на основі їх купівельної поведінки. Алгоритми неконтрольованого навчання можна застосовувати в багатьох областях і особливо корисні для виявлення аномалій. Одним із найпоширеніших алгоритмів є кластеризація сканування на основі щільності (DBSCAN).

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) — дуже корисний алгоритм для виявлення шуму в даних. Логіка цього алгоритму базується на ідентифікації серії центральних точок, визначених сусідніми точками кожного з даних у визначеному радіусі. Ці «околиці» даних утворять кластери, так що дані за їх межами ідентифікуватимуться як шум. Основна перевага цього алгоритму полягає в тому, що він сам визначає кількість кластерів і може групувати їх у різні форми та розміри. Це робить його дуже корисним алгоритмом для роботи з шумними даними та викидами.

Іншим типовим алгоритмом для виявлення аномалії є Isolation Forest. Логіка, за якою він дотримується, відрізняється від інших відомих методів і

обертається навколо ідеї, що аномальні точки в наборах даних легше виділити, ніж нормальні точки. Щоб досягти цього, алгоритм генерує розділи набору даних шляхом випадкового вибору атрибута, потім бере випадкове значення цього атрибута та ділить вибірку на дві частини, групуючи ті, що вище та нижче цього значення. Ці операції повторюються, поки не будуть ізольовані всі спостереження. Основна відмінність цього алгоритму полягає в тому, що він вимагає меншої потужності обробки, ніж інші методи, що робить його особливо придатним для великих наборів даних (дані великої розмірності).

Одним із популярних методів виявлення аномалій на основі кластеризації є алгоритм кластеризації k-середніх. Алгоритм k-means групує точки даних у «k» кластерів на основі їх подібності, яку можна обчислити за допомогою евклідової відстані в просторі ознак або іншими подібними методами. Точки даних, які погано вписуються в жоден із кластерів або належать до кластера із значно іншим розподілом, вважаються аномаліями.

kNN — це простий алгоритм, який використовує k сусідів точки даних як основу для класифікації. Потрібен лише параметр k - кількість сусідів, які слід враховувати. Переглядаючи класи k сусідів, клас більшості вибирається як клас p. У виявленні аномалій kNN можна використовувати для виявлення глобальних аномалій шляхом пошуку k-найближчих сусідів. Оцінку аномалії точки даних можна обчислити кількома різними способами. Одним із способів є використання відстані до k-го найближчого сусіда [14]. Роблячи це, точки даних, які розріджені та віддалені порівняно з іншими точками даних, мають більшу відстань до свого k-го найближчого сусіда, тому ймовірніше, що вони є аномалією. Іншим варіантом є використання середніх відстаней до k-найближчих сусідів. Цей підхід також має переваги з урахуванням локальної щільності точок .

Коефіцієнт локального викиду на основі кластера (Cluster-based Local Outlier Factor) – це виявлення аномалії на основі відстані, яке тісно пов’язане з методами кластеризації. Цей метод має одну функцію для кластеризації та

виявлення викидів за допомогою того самого процесу. В оригінальному алгоритмі він працює шляхом кластеризації даних у кластери з використанням певного алгоритму кластеризації [15]. Ці кластери пізніше визначені як великі кластери або невеликі кластери. Щоб визначити віддаленість, CBLOF використовує комбінацію відстані та розміру. Припустімо, що p є точкою даних, щоб визначити її оцінку аномалії: спочатку обчисліть відстань від p до найближчого великого центру кластера. По-друге, обчисліть розмір кластера p, до якого належить точка, щоб в подальшому використовувати для зважування.

2.4 Виклики та проблеми в системах виявлення аномалій

Щоб забезпечити точне та надійне виявлення, необхідно вирішити низку проблем, пов'язаних із виявленням аномалій. Нижче наведено деякі типові труднощі з виявленням аномалій:

- Відсутність позначених даних про аномалії ускладнює отримання даних для навчання та оцінки.
- Дисбаланс даних: проблема з дисбалансом класів, де аномалії є рідкісними порівняно з типовими випадками, що впливає на продуктивність моделі.
- Робота з різними та складними типами даних, структурами та даними великої розмірності відома як неоднорідність і складність даних.
- Динамічні та еволюційні аномалії: оскільки аномалії можуть змінюватися з часом, моделі повинні розвиватися, щоб розпізнавати нові закономірності.
- Розрізнення реальних аномалій від шумних або невизначених точок даних передбачає розгляд шуму та невизначеності.
- Інтерпретація та пояснення аномалій: надання логічних обґрунтувань виявлених аномалій.

- Масштабованість і обчислювальна ефективність: керування величезними наборами даних і потреба в обробці в реальному часі.

2.5 Висновок до другого розділу

Отже, існує ряд моментів, які потрібно враховувати для виявлення шахрайських транзакцій, зокрема вибір відповідних ознак для виявлення шахрайських транзакцій має вирішальне значення. Витяг значущої інформації з доступних даних з відкиданням шуму є значною проблемою. Це вимагає глибокого розуміння сфери діяльності та моделей шахрайства.

Неповні або суперечливі дані можуть перешкоджати роботі систем виявлення шахрайства. Забезпечення якості та узгодженості даних у різних джерелах має важливе значення для точних прогнозів.

Незбалансовані дані: шахрайські транзакції зазвичай рідкісні порівняно з законними, що призводить до незбалансованих наборів даних. Цей дисбаланс може вплинути на продуктивність моделей машинного навчання, оскільки вони можуть стати упередженими щодо більшості.

В другому розділі описано основні підходи до виявлення шахрайських транзакцій з допомогою методів машинного навчання, зокрема наведено шляхи подолання проблем, які існують при виявленні аномалій як в керованому, так і в некерованому підходах.

3 ПРАКТИЧНА РЕАЛІЗАЦІЯ

3.1 Вибір середовища програмування

Python заслужив своє місце як одна з найпопулярніших мов програмування серед професіоналів машинного навчання завдяки легкому для читання синтаксису, великим бібліотекам і крос-платформній сумісності. Як мова програмування високого рівня з відкритим вихідним кодом, Python став найкращим вибором для широкого спектру завдань машинного навчання, від аналізу даних до глибокого навчання. Зростаюча популярність Python у проектах зі штучного інтелекту (ШІ) і машинного навчання (МН) не випадкова, оскільки вона забезпечує чудове середовище для розробників, щоб вирішувати навіть найскладніші завдання машинного навчання.

Розгалужена бібліотечна екосистема Python, надійні можливості візуалізації, низький бар’єр входу, потужна підтримка спільноти, гнучкість, читабельність і незалежність від платформи роблять його ідеальним вибором для цілей машинного навчання. Як наслідок, Python спостерігав сплеск використання в програмах ШІ та МН, включаючи розпізнавання зображень і мови, прогнозну аналітику.

Зростаюча популярність Python у проєктах штучного інтелекту сьогодні не є простою випадковістю. Його всеосяжна бібліотечна екосистема та активна спільнота розробників полегшили, ніж будь-коли, роботу професіоналам машинного навчання. Завдяки зручному для читання синтаксису, широким бібліотекам і крос-платформній сумісності Python став важливим інструментом для розробників штучного інтелекту та машинного навчання в усьому світі.

Python пропонує безліч переваг як для професіоналів машинного навчання, так і для ентузіастів, особливо при роботі з моделями машинного навчання за допомогою мови Python, зокрема:

Інтуїтивно зрозумілий синтаксис.

Це робить Python популярною мовою програмування, яку легко читати. Об’єктно-орієнтоване програмування надає розробникам логічний метод організації, обробки та планування коду відповідно. Це полегшує розробку чистого та лаконічного коду для проектів будь-якої складності. У результаті Python став популярною початковою мовою для розробників-початківців і вибором для досвідчених програмістів. Легкий для читання синтаксис Python не тільки робить його доступним для початківців, але також дозволяє швидше розробляти та налагоджувати. З Python код стає більш розбірливим і легшим для налагодження, що полегшує виявлення та виправлення помилок і оперативну розробку нових функцій. Цей зручний характер Python значно сприяв його широкому впровадженню в спільноті машинного навчання.

Великі бібліотеки та фреймворки.

Одним із ключових факторів, які відрізняють Python від інших мов програмування, є його комплексна бібліотечна екосистема. Python пропонує широкий спектр бібліотек і фреймворків, спеціально розроблених для машинного навчання, що полегшує розробникам впровадження алгоритмів МН. Нижче наведено основні популярні бібліотеки Python для машинного навчання:

- NumPy — це фундаментальна бібліотека Python для ефективних числових обчислень і операцій з масивами.
- Scikit-learn — це комплексна бібліотека машинного навчання, яка пропонує широкий спектр інструментів для різних завдань, зокрема класифікації, регресії, кластеризації тощо.
- Pandas — це потужна бібліотека для аналізу та обробки даних, що забезпечує інтуїтивно зрозумілі структури даних, такі як DataFrames і Series.
- TensorFlow — це передова бібліотека глибокого навчання, відома своїми можливостями розподіленого обчислення та надійною екосистемою.

- Theano — це бібліотека Python, розроблена для швидких числових обчислень, особливо корисна для навчання моделей глибокого навчання.
- Keras — це простий у використанні API глибокого навчання, який діє як інтерфейс для TensorFlow, Theano або Microsoft Cognitive Toolkit (CNTK), спрощуючи створення та навчання нейронних мереж.
- PyTorch — це динамічна бібліотека глибокого навчання з гнучким графіком обчислень, що робить її ідеальною для розробки та навчання складних нейронних мереж.

Ці бібліотеки та фреймворки Python надають потужні можливості для аналізу даних, машинного та глибокого навчання, дозволяючи розробникам зосередитися на вирішенні складних завдань без необхідності винаходити колесо. Завдяки цій чудовій бібліотечній екосистемі Python став незамінним інструментом для інженерів машинного навчання, науковців із обробки даних і дослідників.

Кросплатформна сумісність

Кросплатформна сумісність Python дозволяє розробникам створювати код, який можна використовувати на різних платформах, таких як Windows, Mac і Linux. Ця гнучкість полегшує розробку програм, які можна використовувати в різних операційних системах без необхідності переписувати вихідний код. Таким чином, це дозволяє розробникам використовувати той самий код для різних платформ, заощаджуючи час і зусилля. Однак кросплатформна сумісність пов’язана з певними труднощами. На різних платформах можуть бути встановлені різні версії Python, що може призвести до проблем із сумісністю під час виконання коду на різних платформах. Щоб подолати ці проблеми, важливо переконатися, що код написаний у спосіб, сумісний з усіма підтримуваними версіями, і що він протестований на всіх платформах, щоб гарантувати, що він функціонує належним чином.

Масштабування та продуктивність

Python широко відомий своєю масштабованістю та винятковою продуктивністю в машинному навчанні. Його універсальність, зручний характер і великі бібліотеки роблять його ідеальним вибором для масштабування операцій машинного навчання. Завдяки таким бібліотекам, як NumPy, Pandas і TensorFlow, Python дає змогу виконувати складні операції з масивними наборами даних, демонструючи свою високу масштабованість. Його вміння працювати з великими даними сприяє його широкому застосуванню. Простота та читабельність Python ще більше сприяють швидкому створенню прототипів, прискорюючи ітераційний процес розробки та тонкого налаштування моделей МН. Однак продуктивність Python створює проблеми. Як інтерпретована мова, Python є відносно повільнішим порівняно з такими мовами, як C++ або Java. Тим не менш, такі бібліотеки, як NumPy і Cython, вирішують цю проблему, виконуючи обчислення зі швидкістю, близькою до C. Крім того, інфраструктури розподілених обчислень, такі як Apache Spark і Dask, значно підвищують продуктивність Python у програмах МН. Загалом, багатий набір бібліотек, простота використання та масштабованість Python роблять його надійним вибором для машинного навчання.

3.2 Опис датасету

Для проведення власних досліджень ми використали датасет, доступний на сайті https://github.com

Набори даних містять транзакції, здійснені за допомогою кредитних карток у вересні 2013 року європейськими власниками карток. Цей набір даних представляє транзакції, які відбулися за два дні, де ми маємо 492 шахрайства з 144807 транзакцій. Набір даних дуже незбалансований, позитивний клас (шахрайство) становить 0,3397 % усіх транзакцій.

Набір даних містить лише числові вхідні змінні, які є результатом перетворення PCA. Метод головних компонент (Principal Component Analysis – PCA) є методом некерованого машинного навчання, що використовується для скорочення простору ознак. Цей метод перетворює змінні в такий спосіб, що дозволяє виокремити нові ознаки з максимальним рівнем дисперсії та усунути кореляцію між ознаками. На жаль, через питання з конфіденційністю оригінальні ознаки та додаткова довідкова інформація про дані не могли бути подана. Тому метод головних компонент був радше застосований для перетворення початкових ознак з метою приховування чутливої інформації. Функції V1, V2, ... V28 є основними компонентами, отриманими за допомогою PCA, єдиними ознаками, які не були перетворені за допомогою PCA, є «Час» і «Сума». Ознака «Час» містить секунди, що минули між кожною транзакцією та першою транзакцією в наборі даних. Функція «Клас» — це змінна відповіді, яка приймає значення 1 у разі шахрайства та 0 в іншому випадку. Оскільки дані промітковані, то задача визначення шахрайських транзакцій фактично зводиться до задачі класифікації.

3.3 Попередня обробка даних

Попередня обробка даних включає в себе роботу з пропущеними даними, нормалізацію даних, перевірку датасету на збалансованість. За звичай працюють ще й з викидами. Але, оскільки перед нами стоїть задача фактично визначення аномальних значень, яку можна трактувати як задачу знаходження викидів.

Для зчитування датасету та виведення його перший n рядків спостережень можна використати команди:

```
credit_card_df = pd.read_csv("creditcard.csv")
credit_card_df.head(10)
```

Для перевірки, чи містить датасет пропущені дані можна скористатись наступним кодом:

```
number_missing_values = credit_card_df.isnull().sum()
```

Результати виконання коду засвідчили, що пропущені дані в датасеті відсутні.

Перевіримо датасет на збалансованість. Набір даних містить дві можливі мітки 1 -відповідає шахрайським транзакціям, а 0 – звичайним.

Командою

```
fraud_count = credit_card_df[credit_card_df["Class"] == 1].shape[0]
```

підтвердили попередню інформацію про те, що шахрайських транзакцій – 492.

Натомість кількість звичайних транзакцій встановлена командою

```
norm_count = credit_card_df[credit_card_df["Class"] == 0].shape[0]
```

дорівнює 144315, що підтвердило попередню інформацію про те, що відсоток шахрайських транзакцій в датасеті рівний 0,33%.

Побудуємо далі гістограми кожної ознак для проведення подальшого аналізу. Для цього скористаємося командою

```
credit_card_df.hist(figsize=(20,20));
```

Результат виконання команди зображено на рис. 3.1

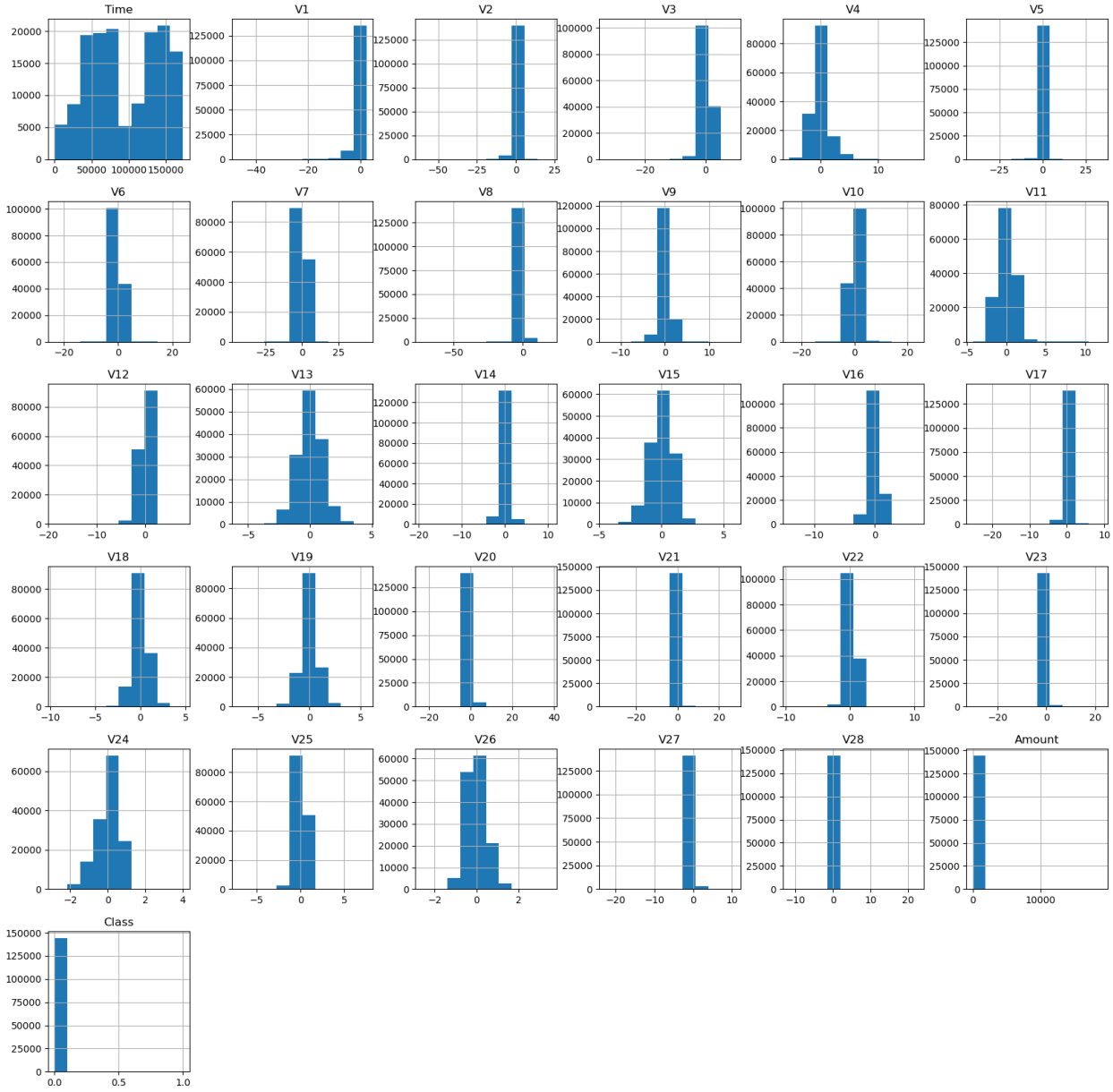

Рисунок 3.1 – Гістограмний аналіз кожної з ознак датасету

Можемо зробити наступні висновки з наведених на рисунку гістограм:

- По-перше, діапазон значень характеристик V1-V28 є не великим, а отже дуже ймовірно, що вони вже нормалізовані (що і є очікувано, оскільки це необхідний елемент PCA).
- Необхідно нормалізувати стовпчик `Amount`. Для цього можна використати `StandardScaler`.

- Стовпчик `Time` може створювати проблеми з моделлю і тому з його використанням варто бути обережним. Проте проведемо наразі нормалізацію цих даних з використанням `RobustScaler`.

В лістингу 3.1 наведено код для нормалізації значень сум та дат транзакцій:

Лістинг 3.1 – Нормалізація значень сум та дат транзакцій

```
std_scaler = StandardScaler()
rob_scaler = RobustScaler()
credit_card_df["Amount"] =
std_scaler.fit_transform(credit_card_df[["Amount"]])
credit_card_df["Time"] =
rob_scaler.fit_transform(credit_card_df[["Time"]])
credit_card_df.rename(inplace=True,columns={"Amount":"Scaled
Amount","Time":"ScaledTime"})
credit_card_df.head(10)
```

У чистому вигляді час напевне не матиме значення для побудови моделі, або й навпаки шкодитиме їй.

Якщо б у нас була ще інформація про клієнта (умовно ClientID), то можна було б побудувати ряд статистики на основі часу та суми покупки (умовно статистику за тиждень, місяць, рік).

Щоб перевірити, чи буде корисним час у даному випадку, достатньо візуалізувати дані. Для цього створимо два датасети (один з часом, та без). Ми зробимо датасети значно меншими ніж оригінальний, щоб час виконання візуалізації був невеликим (лістинг 3.2)

Лістинг 3.2 – Вибірка з датасету для кращої візуалізації

```
reducer_data = pd.concat([
credit_card_df[credit_card_df["Class"]==0].sample(frac=1)[:2
000],
    credit_card_df[credit_card_df["Class"]==1]
```

```
])
```

Даний код відбирає 2000 нешахрайських транзакцій та всі 492 шахрайські транзакції.

Будуємо 2D проєкцію з та без часової мітки в даних (лістинг 3.3)

Лістинг 3.3 – Код для створення 2D проєкції даних з та без часової мітки

```
reducer = TSNE(n_components=2, random_state=42)
proj_2d_with_time =
reducer.fit_transform(reducer_data.drop(["Class"], axis=1))
reducer = TSNE(n_components=2, random_state=42)
proj_2d_without_time =
reducer.fit_transform(reducer_data.drop(["Class","ScaledTime
"], axis=1))
```

На рисунку 3.2 візуалізовано дві проєкції

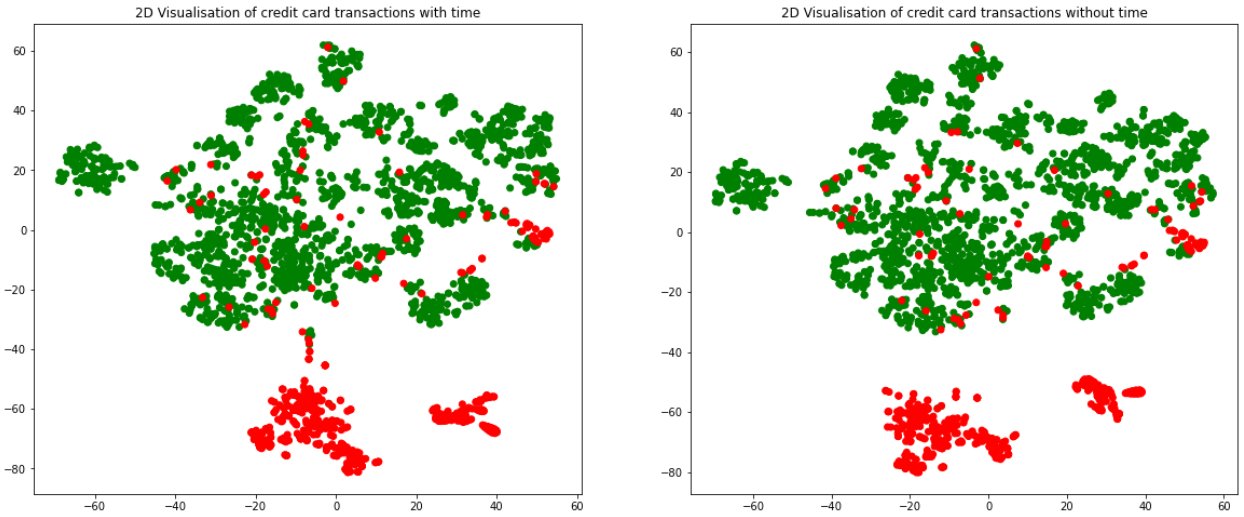

Рисунок 3.2 – Візуалізація даних з та без часової мітки

Як бачимо, в даному випадку ScaledTime явно не шкодить передбаченням, тому можемо його використовувати.

3.4 Реалізація моделей та оцінка точності

3.4.1 One Class SVM

Першим підходом до виявлення аномалій є побудова однокласового SVM. Суть полягає в тому, що ми навчаємо модель виключно на нормальних даних і вона з них будує один кластер. Далі ми закидаємо туди тестові дані і дивимось чи вони попадають у вказаний кластер (нормальні дані) чи ні (аномалії).

У цьому випадку зазвичай використовують не лінійні ядра (kernel), а ядра що будуть певну замкнуту фігуру навколо даних.

SVM є доволі повільною моделлю, тому ми зробили рандомно зменшену навчальну вибірку з нормальних даних у розмірі 10000 елементів. Таку ж вибірку ми зробили для тестування (лістинг 3.4)

Лістинг 3.4 – Побудова меншої вибірки для проведення класифікації:

```
svmXnormal =
credit_card_df[credit_card_df["Class"]==0].drop("Class",
axis=1).sample(frac=1)
svmXfraud =
credit_card_df[credit_card_df["Class"]==1].drop("Class",
axis=1)
subsample_size=10000
# навчальні нормальні дані
svmXtrain = svmXnormal[:subsample_size]
# тестові нормальні дані
svmXtest = svmXnormal[subsample_size:2*subsample_size]
svmYtest = np.ones([svmXtest.shape[0], 1])
# з'єднуємо нормальні дані з аномальними у тестовому датасеті
svmXtest = pd.concat([svmXtest, svmXfraud])
svmYtest = np.concatenate((
    svmYtest,
    np.ones([svmXfraud.shape[0], 1]) * -1
```

))

Після цього будуємо модель OneClassSVM з ядром rbf та значенням gamma=scale та навчаємо її на навчальних даних.

Дана модель не отримала високої оцінки. Насамперед це через те, що ми їй дозволяємо робити багато помилок. Це можна змінити поставивши значно менше значення параметра nu конструктора моделі. Стандартне значення цього параметра - 0.5. Встановимо його значенням наближеним до відсотка аномалій, наприклад 0.02.

В таблиці 3.1 наведено різні метрики для оцінки точності моделі

Таблиця 3.1 – Оцінка точності класифікаційної моделі

	Незбалансований датасет	
Метрики	OneClassSVM з ядром rbf	OneClassSVM з ядром rbf, nu=0.02
recall_score	0,49	0,93
precision_score	0,91	0,97
f1_score	0,62	0,96

Як бачимо налаштування параметрів моделі дозволило суттєво підняти точність. Слід зазначити, що у випадку виявлення аномалій метрику accuracy недоцільно застосовувати для оцінки моделей, бо високе значення точності не вказує на високу оцінку моделі виявлення аномальних транзакцій. На незбалансованому датасеті висока точність може свідчити лише про високу точність визначення нормальних транзакцій.

3.4.2 Модель випадкового лісу та XGB на збалансованому датасеті

Для виявлення аномальних транзакцій часто рекомендують використовувати методи для балансування датасету. Збалансуємо датасет наступним чином.

З оригінального датасету створимо нові навчальний та тестовий датасети. У навчальний датасет візьмемо 80% даних з шахрайських транзакцій (рандомно) та ідентичну кількість даних з нормальних транзакцій (рандомно). Таким чином, у нас вийде збалансований навчальний датасет. У тестовий датасет візьемо решту 20% даних з шахрайських транзакцій та додамо 1000 записів з нормальних транзакцій. Потрібно враховувати, щоб у тестовий датасет не потрапили ті ж дані з непідозрілих транзакцій, що містяться у навчальному датасеті. Результати записуємо у train_df та test_df відповідно. Виведемо перші 10 записів з навчального датасету (лістинг 3.5).

Лістинг 3.5 – Формування збалансованого навчального та тестового наборів даних

```
fraud =
credit_card_df[credit_card_df["Class"]==1].sample(frac=1,ran
dom_state=np_seed)
normal =
credit_card_df[credit_card_df["Class"]==0].sample(frac=1,ran
dom_state=np_seed)
fr_c = int(fraud.shape[0]*0.8)
train_df = pd.concat([fraud[:fr_c],normal[:fr_c]])
test_df = pd.concat([normal[fr_c:fr_c+1000],fraud[fr_c:]])
test_df.head(10)
```

Побудуємо дві моделі для класифікації для збалансованого датасету RandomForestClassifier та XGBClassifier. Приклад коду для класифікації та оцінки точності моделі наведено в лістингу 3.6

Лістинг 3.6 – Код для проведення класифікації та оцінки точності моделі RandomForestClassifier

```
forest = RandomForestClassifier(n_estimators=1000,
random_state=17)
forest.fit(trainX,trainY)
predicted_f = forest.predict(testX)
precision_score(testY,predicted_f)
recall_score(testY,predicted_f)
f1_score(testY,predicted_f)
```

Оцінка точності побудованих моделей наведена в таблиці 3.2.

Таблиця 3.2 - Оцінка точності класифікаційної моделі

	Збалансований датасет	
Метрики	RandomForestClassifier	XGBClassifier
recall_score	0,75	0,77
precision_score	0,9	0,88
f1_score	0,81	0,82

Як бачимо, результат вийшов протилежний до очікуваного. Використання збалансованого датасету не дало приросту точності, а радше навпаки призвело до її зменшення. Потрібно застосувати інші методи балансування датасету і оцінити точність.

3.5 Висновок до третього розділу

В даному розділі обгрунтовано вибір програмного середовища для побудови моделей машинного навчання. Наведено опис датасету, описано основні етапи попередньої обробки даних, застосовано рекомендовані дослідниками моделі для виявлення шахрайських транзакцій.

В даному дослідженні отримали доволі несподівані результати, адже модель класифікації на збалансованих даних показала дещо нижчі результати точності, ніж модель однокласового SVM, де модель виключно на нормальних даних і вона з них будує один кластер. Оцінку моделі проводимо на основі тестових даних, які містять нормальні транзакції та шахрайські та перевіряємо, котрі з инх належать до вказаного кластеру.

Дана робота може бути основою для побудови системи виявлення шахрайських транзакцій, оскільки має порівнювану точність з існуючими, проте існуючі рішення часто коштують від кількасот доларів в місяць. Тому розроблене ПЗ може стати доволі непоганою безкоштовною альтернативою комерційним рішенням.

4 ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Охорона праці

Метою кваліфікаційної роботи магістра є дослідження методів виявлення шахрайських транзакцій. Оскільки, проведення робіт з розробки та використання системи передбачає використання комп'ютерної техніки, зокрема ПК та периферійних пристроїв, то обов'язковим є дотримання вимог з охорони праці і техніки безпеки.

Для ефективної і безпечної роботи колективу працівників, в тому числі і фахівців з підвищення ефективності контролю доступу в приміщення, необхідно організувати безпечні умови праці. При цьому керівник організації несе безпосередню відповідальність за порушення нормативно-правових актів з охорони праці [17]. Окрім цього, на робочих місцях працівників необхідно забезпечити дотримання вимог, затверджених Наказом Мінсоцполітики від 14.02.2018 за № 207 «Про затвердження вимог щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Згідно вимог приміщення, де розміщені робочі місця операторів, крім приміщень, у яких розміщені робочі місця операторів великих ЕОМ загального призначення (сервер), мають бути оснащені системою автоматичної пожежної сигналізації відповідно до цих вимог:

- переліку однотипних за призначенням об'єктів, які підлягають обладнанню автоматичними установками пожежогасіння та пожежної сигналізації, затвердженого наказом Міністерства України з питань надзвичайних ситуацій та у справах захисту населення від наслідків Чорнобильської катастрофи від 22.08.2005 N 161, зареєстрованого в Міністерстві юстиції України 05.09.2005 за N 990/11270 (НАПБ Б.06.004-2005);

- Державних будівельних норм "Інженерне обладнання будинків і споруд. Пожежна автоматика будинків і споруд", затверджених наказом Держбуду України від 28.10.98 N 247 (далі - ДБН В.2.5-56:2014, з димовими пожежними сповіщувачами та переносними вуглекислотними вогнегасниками.

В інших приміщеннях допускається встановлювати теплові пожежні сповіщувачі. Приміщення, де розміщені робочі місця операторів, мають бути оснащені вогнегасниками, кількість яких визначається згідно з вимогами ДСТУ 4297:2004 «Пожежна техніка. Технічне обслуговування вогнегасників». Загальні технічні вимоги і з урахуванням граничнодопустимих концентрацій вогнегасної рідини відповідно до вимог НАПБ А.01.001-2014. Приміщення, в яких розміщуються робочі місця операторів сервера загального призначення, обладнуються системою автоматичної пожежної сигналізації та засобами пожежогасіння відповідно до вимог ДБН В.2.5-56:2014, ДБН В.2.5-56:2010, НАПБ А.01.001-2014 і вимог нормативно-технічної та експлуатаційної документації виробника. Проходи до засобів пожежогасіння мають бути вільними.

Лінія електромережі для живлення комп’ютера та периферійних пристроїв повинні бути виконаними як окрема групова трипровідна мережа шляхом прокладання фазового, нульового робочого та нульового захисного провідників. Нульовий захисний провідник використовується для заземлення (занулення) електроприймачів. Не допускається використовувати нульовий робочий провідник як нульовий захисний провідник. Нульовий захисний провідник прокладається від стійки групового розподільного щита, розподільного пункту до розеток електроживлення. Не допускається підключати на щиті до одного контактного затискача нульовий робочий та нульовий захисний провідники.

Площа перерізу нульового робочого та нульового захисного провідника в груповій трипровідній мережі має бути не менше площі перерізу фазового

провідника. Усі провідники мають відповідати номінальним параметрам мережі та навантаження, умовам навколишнього середовища, умовам розподілу провідників, температурному режиму та типам апаратури захисту, вимогам НПАОП 40.1-1.01-97.

У приміщенні, де одночасно експлуатуються понад п’ять комп’ютерів, на помітному, доступному місці встановлюється аварійний резервний вимикач, який може повністю вимкнути електричне живлення приміщення, крім освітлення. Комп’ютери повинні підключатися до електромережі тільки за допомогою справних штепсельних з'єднань і електророзеток заводського виготовлення.

У штепсельних з'єднаннях та електророзетках, крім контактів фазового та нульового робочого провідників, мають бути спеціальні контакти для підключення нульового захисного провідника. Їхня конструкція має бути такою, щоб приєднання нульового захисного провідника відбувалося раніше, ніж приєднання фазового та нульового робочого провідників. Порядок роз’єднання при відключенні має бути зворотним. Не допускається підключати комп’ютери до звичайної двопровідної електромережі, в тому числі – з використанням перехідних пристроїв. Електромережі штепсельних з'єднань та електророзеток для живлення комп’ютерної техніки повинні бути виконаними за магістральною схемою, по 3-6 з'єднань або електророзеток в одному колі. Штепсельні з'єднання та електророзетки для напруги 12 В та 42 В за своєю конструкцією мають відрізнятися від штепсельних з'єднань для напруги 127 В та 220 В. Штепсельні з'єднання та електророзетки, розраховані на напругу 12 В та 42 В, мають візуально (за кольором) відрізнятися від кольору штепсельних з'єднань, розрахованих на напругу 127 В та 220 В.

При підвищенні ефективності контролю доступу в приміщення, де для забезпечення безпеки мешканців, співробітників і збереження майна використовуються ДС, важливим, з точки зору охорони праці, є забезпечення достатньої величини природного та штучного освітлення, які визначені у

НПАОП 0.00-7.15-18. Організація робочого місця фахівця із дослідження методів та програмно-апаратних засобів оптимізаційних процесів на основі ГА повинна забезпечувати відповідність усіх елементів робочого місця та їх розташування ергономічним вимогам ДСТУ 8604:2015 «Дизайн і ергономіка. Робоче місце для виконання робіт у положенні сидячи. Загальні ергономічні вимоги». Відстань від екрана до ока фахівців, які працюють за комп'ютером визначається згідно з вимогами ДСанПіН 3.3.2.007-98.

Розміщення принтера або іншого пристрою введення-виведення інформації на робочому місці має забезпечувати добру видимість екрана комп'ютера, зручність ручного керування пристроєм введення-виведення інформації в зоні досяжності моторного поля згідно з вимогами ДСанПіН 3.3.2.007-98.

4.2 Безпека в надзвичайних ситуаціях

4.2.1 Міжнародний тероризм

Терор (лат. terror – страх, жах) – має ознаку «усувати», «закривати». Ця обставина і визначає терор як особливу форму політичного насильства, що характеризується жорстокістю, цілеспрямованістю й уявленою ефективністю. Ці особливості визначили широке використання терору упродовж людської історії як засобу політичної боротьби в інтересах держави, організацій чи окремих угруповань. Безпосередньо сам факт привселюдної страти кримінальних чи політичних злодіїв, чи процес «аутодафе» в період середньовікової інквізиції, є класичною формою терору в інтересах держави чи католицької церкви.

Правовою основою боротьби з міжнародним тероризмом є «Декларація про заходи для ліквідації міжнародного тероризму», що затверджена на 49-й сесії Генеральної асамблеї ООН (резолюція 49/60 від 9 грудня 1994 р.)

Цей документ встановлює принципи відносин світової спільноти і програму заходів з метою ліквідації такого огидного суспільного явища, як

міжнародний тероризм, а також встановлює подальше співробітництво між державами для невідкладної ліквідації будь-яких форм і проявів терористичної діяльності. Характерним для розвитку світової спільноти є те, що наявність лідера (провідної країни чи провідної сили) народжує відповідну реакцію – формування нижчого за рангом (рівнем) іншого лідера (іншої країни чи іншої провідної сили). Має місце формування біполярності, виникають реалії антагонізму на різних рівнях світового суспільства На другому етапі формування ці політичні, злочинні та інші сили (групи) шукають собі відповідні «ніші» існування; економічну, політичну, наукову та інші види підтримок; формують свої озброєні сили, відповідні професійні кадри, джерела озброєння, територію знаходження тощо. При цьому використовуються всі «блага» цивілізації особистого розвитку і поширення впливу на світову спільноту.

Міжнародний тероризм, створюючи свій плацдарм, може викликати кризи (системні) в світовій, моральній, політичній, економічній системі відносин і зруйнувати та усунути всі передумови розвитку світової спільноти.

В Україні, за даними служби безпеки, за останні два роки скоєно понад 560 злочинів терористичного характеру, внаслідок цього 90 осіб (із них 15 представників владних структур) загинуло. В Україні зростає активність міжнародних терористичних організацій, насамперед із країн Ближнього Сходу («Хезбола», «Абу Ніджаль», «Хамас», «Брати мусульмани»), які прагнуть використати територію України для транзиту своїх бойовиків до країн західної Європи, підготовки терористичних акцій.

Головними принципами попередження та боротьби з міжнародним тероризмом має стати постійне удосконалення відповідної законодавчої бази, співробітництво з правоохоронними організаціями, консолідація з іншими країнами й організація напрямів запобігань поширенню будь-яких терористичних організацій і угруповань.

Терористичний акт не має безпосередніх можливостей досягнення оголошеної кінцевої мети і звичайно складається з таких елементів: насильницька дія у різноманітних її формах, політичний мотив в основі здійснення самого терористичного акту; сам акт спрямовано проти осіб, організацій, націй, національностей і меншин, державних інститутів чи їх представників з метою їх залякування чи виконання окремих вимог. Терор щодо націй, етнічної, расової чи релігійної групи, що здійснюється для її повного чи часткового усунення, розглядається світовою спільнотою вже як акт геноциду.

Варіанти комбінацій за спрямованістю суб'єкт—об'єкт здійснення терористичного акту багатоспрямовані, тому важко дати універсальне визначення «терору». Проте деякі критерії певної класифікації можна встановити:

- індивідуальний, організований терор і терор як політика держави;
- терор як метод внутрішньополітичної боротьби і терористичні акти міжнародного характеру.

4.2.2 Структура системи БЖД

Поняття «життєдіяльність» стосується тільки людини. Людина живе і працює в безпосередньому зв’язку з навколишнім середовищем.

Життєдіяльність (ЖД) – це складна фізіологічна система, яка має назву «система ЖД».

Системою називають сукупність взаємозв’язаних елементів, функціонування яких спрямоване на досягнення певної загальної мети.

Система ЖД складається із взаємопов’язаних елементів: життя, діяльності людини, навколишнього середовища, – і має підтримувати комфортне та безпечне існування людини, забезпечити сталий розвиток людства.

Розглянемо характеристики елементів системи ЖД.

Життя – це форма існування матерії, яка характеризується обміном речовин, здатністю до розмноження і розвитку, вмінням пристосовуватись до навколишнього середовища.

Людина – вища форма розвитку живої матерії, і її існування – дуже складний процес, що не тільки підтримує її фізіологічний стан, але й задовольняє духовні потреби. Крім того, на життя людини суттєво впливають умови проживання та праці, медичний догляд і багато інших факторів, що виникають завдяки діяльності самих людей.

Діяльність – це специфічна форма ставлення людей до навколишнього середовища та одне до одного, яка має задовольняти потреби та інтереси людини. Це соціальна категорія, нерозривно зв'язана із суспільством. Тільки завдяки діяльності людини створено всі блага, які має людство.

Основні види діяльності такі:

- виробнича;
- наукова;
- мистецька;
- освітня.

Однією із специфічних форм діяльності людини є праця – перша й основна умова існування людини (людства).

Праця – цілеспрямована діяльність людини, у процесі якої вона впливає на природу і використовує її з метою виробництва матеріальних та інших благ, необхідних для задоволення своїх потреб.

Потреби – це необхідність для людини того, що забезпечує її існування і самозабезпечення (фізіологічне, матеріальне, соціальне, духовне та ін.).

Навколишнє середовище (довкілля) або середовище існування – це все, що оточує людину впродовж її життя. Навколишнє середовище, у свою чергу, поділяють на такі види:

- природне середовище;
- штучне середовище.

Природне середовище (біосфера) – це частина Землі і простору навколо неї, де зосереджено все живе. Біосфера включає:

- атмосферу (газоподібна частина);
- гідросферу (рідка водна частина);
- літосферу (тверда частина).

На ЖД людей найбільше впливає частина біосфери від поверхні Землі вглиб на 15–20 км і до висоти 20–22 км, де починається озоновий шар. Природне середовище є джерелом природних ресурсів для існування людини: повітря, води, деревини, корисних копалин, ґрунту та ін.

Штучне середовище – це складова довкілля, створена людством за тривалий час його існування. Штучне середовище умовно можна поділити на два види:

- виробниче середовище;
- побутове середовище.

Виробничим називають середовище, в якому людина реалізує свою трудову діяльність (підприємства, установи, навчальні заклади тощо).

Побутовим є середовище, де люди мешкають або проводять вільний час. Воно охоплює сукупність житлових будинків, комунально-побутових об’єктів, місця відпочинку та ін.

Організм людини може нормально функціонувати тільки тоді, коли умови (параметри) зовнішнього середовища відповідають оптимальним. Якщо умови середовища змінюються, стають несприятливими, то на протидію їм організм людини включає спеціальні механізми, які зберігають постійність параметрів внутрішнього середовища (всередині організму) чи змінюють їх у межах допустимого.

Можливість функціонування організму в середовищі, параметри якого постійно змінюються, забезпечується завдяки механізму, який називають адаптацією.

Адаптація (лат. adapto – пристосування) – динамічний процес пристосування організму до мінливих умов зовнішнього середовища, який спостерігається в будь-якому виді діяльності щоразу, коли виникають значні зміни в системі «людина – середовище». Адаптація може бути фізіологічною, психологічною, соціальною.

Отже, для функціонування системи ЖД середовище має обов'язково відповідати природним параметрам. Відхилення можливі в межах допустимого, коли організм людини здатний адаптуватися, захистити себе, підтримувати існування. Усе, що існує за цими межами, становить загрозу життю, тому виникає потреба захисту ЖД людей. Отже, безпека – важлива складова системи ЖД.

Розглядаючи систему ЖД як взаємодію людей з навколишнім середовищем, слід зауважити, що вона завжди підпорядкована певним принципам, правилам, умовам життя, природним умовам, традиціям тощо.

Система ЖД має такі характерні ознаки:

- її функціонування підпорядковане об'єктивним законам природи;
- це динамічна система, яка розвивається, удосконалюється, пристосовується до змін умов існування;
- тяжіє до сталого розвитку, вживаючи заходів захисту від упливу негативних факторів.

Основні принципи забезпечення ЖД такі:

- своєчасність, достатність, якість забезпечення людей необхідними для життя засобами високої якості і заходами в потрібний час у належній кількості;
- безпека ЖД (захист ЖД від упливу негативних факторів, що виникають унаслідок як природних явищ, так і діяльності людей).

Рівень реалізації цих принципів значною мірою залежить від способів забезпечення ЖД. Виходячи із сказаного, можна визначити такі головні способи забезпечення ЖД:

1. Організація ефективної трудової діяльності людей в суспільстві з максимальним залученням усіх ресурсів (створення робочих місць, упровадження високопродуктивного виробництва і технологій, нормування праці тощо).

2. Організація та удосконалення освіти і підготовка кадрів, розвиток науки відповідно до вимог часу.

3. Розвиток сфери послуг (комунальних, транспортних, торговельних, побутових і т. ін.).

4. Розширення мережі культурних, спортивних, розважальних установ.

5. Проведення заходів щодо збереження здоров'я людей (диспансеризація, оздоровлення, кваліфіковане медичне обслуговування і лікування, санітарно-епідеміологічний стан).

6. Розроблення законодавчих і нормативно-правових актів із забезпечення прав, свобод і захисту людей і суспільства в цілому.

Залежно від того, якою мірою реалізуються принципи та способи забезпечення ЖД, визначається рівень життя людей окремих країн і загальний розвиток людства.

4.3 Висновки до 4 розділу

Таким чином, у результаті аналізу вимог щодо охорони праці користувачів комп'ютерів, визначено особливості організації робочих місць, вимог з електробезпеки, природного та штучного освітлення для ефективної і безпечної роботи.

Також розглянуто питання міжнародного тероризму, структури системи БЖД.

ВИСНОВКИ

Під час виконання кваліфікаційної роботи освітнього рівня «Магістр» було спроєктовано програмне забезпечення для виявлення аномалій в базі даних транзакцій. Для дослідження використано відкритий набір даних з перетвореними ознаками методом головних компонент з метою приховування конфіденційної та чутливої інформації про клієнтів

Було досягнуто виконання наступних завдань:

- Здійснено огляд останніх статистичних даних по шахрайству, який засвідчив важливість даного дослідження, особливо для України.
- Проведено аналіз основних загроз для фінансових установ.
- Проаналізовано види шахрайств в залежності від організації.
- Проведено огляд основних підходів до створення систем виявлення шахрайських транзакцій
- Розглянуто існуючі технічні рішення для виявлення шахрайських транзакці;
- Запропоновано моделі керованого машинного навчання та проведено їх тестування на наборі даних, який є у відкритому доступі;
- Запропоновано метрики і проведено оцінку точності запропонованих моделей.
- Розглянуто окремі питання з охорони праці і безпеки в надзвичайних ситуаціях.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. PwC’s Global Economic Crime and Fraud Survey [Електронний ресурс]. URL: https://www.pwc.com/gx/en/forensics/gecsm-2022/pdf/PwC%E2%80%99s-Global-Economic-Crime-and-Fraud-Survey-2022.pdf

2. Total value of losses due to card fraud worldwide - split between the United States and rest of the world - from 2014 to 2021 Survey [Електронний ресурс]. URL: https://www.statista.com/statistics/1264329/value-fraudulent-card-transactions-worldwide/

3. Пандемія шахрайства: в Україні зафіксовано рекордна активність шахраїв [Електронний ресурс]. URL: https://opendatabot.ua/analytics/fraud-pandemic

4. N. Zagorodna, I. Kramar. Economics, Business and Security: Review of Relations. Business Risk in Changing Dynamics of Global Village BRCDGV-2020: Monograph / Edited by Pradeep Kumar, Mahammad Sharif. India, Patna: Novelty & Co., Ashok Rajpath,. 446 p., pp.25-39. Отримано з: http://elartu.tntu.edu.ua/handle/lib/33034

5. Alice Hutchings. Computer security threats faced by small businesses in Australia, Trends and issues in crime and criminal justice, No 433, 2012, Australian Institute of Criminology ISSN: 1836-2206

https://www.aic.gov.au/publications/tandi/tandi433

6. Dacosta, Italo & Chakradeo, Saurabh & Ahamad, Mustaque & Traynor, Patrick. (2012). One-Time Cookies: Preventing Session Hijacking Attacks with Stateless Authentication Tokens. ACM Transactions on Internet Technology - TOIT. 12. 10.1145/2220352.2220353.

7. Sheng X.S. (2009). A policy analysis of phishing countermeasures. Pittsburgh: Carnegie MellonUniversity. http://gradworks.umi.com/3383412.pdf

8. Symantec (2019). Cloud Security Threat Report (CSTR). Adapting to the New Reality of Evolving Cloud Threats

9. Amit Gupta, M. C. Lohani & Mahesh Manchanda (2021) Financial fraud detection using naive bayes algorithm in highly imbalance data set, Journal of Discrete Mathematical Sciences and Cryptography, 24:5, 1559-1572, DOI: 10.1080/09720529.2021.1969733
https://www.tandfonline.com/doi/pdf/10.1080/09720529.2021.1969733

10. Zhu, Xiaoqian & Ao, Xiang & Qin, Zidi & Chang, Yanpeng & Liu, Yang & He, Qing & Li, Jianping. (2021). Intelligent Financial Fraud Detection Practices in Post-Pandemic Era: A Survey. The Innovation. 2. 100176. 10.1016/j.xinn.2021.100176.
https://www.researchgate.net/publication/355514839_Intelligent_Financial_Fraud_Detection_Practices_in_Post-Pandemic_Era_A_Survey

11. Top 10 Credit Card Fraud Detection Solutions in 2023 [Електронний ресурс]. URL: https://cybeready.com/top-10-credit-card-fraud-detection-solutions-in-2023

12. S. B. Kotsiantis Supervised Machine Learning: A Review of Classification Techniques. Informatica 31 (2007) 249-268 https://datajobs.com/data-science-repo/Supervised-Learning-[SB-Kotsiantis].pdf

13. R.P.Shanthi Rani, Allimalli Durgabhavani, R.Reeja Igneshia Malar, Amit Kumar Singh SMOTE: Credit Card Fraud Detection Using Supervised Machine Learning Methods // Journal of Engineering Sciences ICETT- Vol 14 Issue 05(S),2023 https://jespublication.com/specialissue/2023-V14I5017.pdf

14. Sridhar Ramaswamy, Rajeev Rastogi, and Kyuseok Shim. “Efficient algorithms for mining outliers from large data sets”. In: Proceedings of the 2000 ACM SIGMOD international conference on Management of data. 2000, pp. 427–438

15. Zengyou He, Xiaofei Xu, and Shengchun Deng. “Discovering cluster-based local outliers”. In: Pattern recognition letters 24.9-10 (2003), pp. 1641–1650

16. T. Zoppi, A. Ceccarelli and A. Bondavalli, "Into the Unknown: Unsupervised Machine Learning Algorithms for Anomaly-Based Intrusion Detection," 2020 50th Annual IEEE-IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S), Valencia, Spain, 2020, pp. 81-81, doi: 10.1109/DSN-S50200.2020.00044.

17. ГОСТ 12.1.005-88. ССБТ. Загальні санітарно-гігієнічні вимоги до повітря робочої зони. [Електронний ресурс]. URL: http://online.budstandart.com/ua/catalog/doc-page.html?id_doc=6264.

18. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної та заочної (дистанційної) форм навчання «БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ» / В.С. Стручок –Тернопіль: ФОП Паляниця В.А., – 156 с. Отримано з https://elartu.tntu.edu.ua/handle/lib/39196.

ДОДАТКИ

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ПУЛЮЯ

М А Т Е Р І А Л И

XI НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

«ІНФОРМАЦІЙНІ МОДЕЛІ, СИСТЕМИ ТА ТЕХНОЛОГІЇ»

13-14 грудня 2023 року

ТЕРНОПІЛЬ
2023

УДК 004.056
О.Безруков, Стадник Марія, к.т.н., доц.
(Тернопільський національний технічний університет імені Івана Пулюя)

ВИЯВЛЕННЯ ШАХРАЙСЬКИХ ТРАНЗАКЦІЙ З ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ

O. Bezrukov, Stadnyk Mariia, Ph.D., Assoc. Prof.

DETECTION OF FRAUD TRANSACTIONS USING MACHINE LEARNING METHODS

Виявлення шахрайства – це набір процесів і методів, які дозволяють підприємствам виявляти та запобігати несанкціонованій фінансовій діяльності, до якої, зокрема, належать шахрайські операції з кредитними картками, крадіжки, кіберзломи, шахрайство зі страхуванням тощо. Виявлення шахрайських транзакцій є однією з центральних питань безпеки таких фінансових установ, як банки, та стосується процесу моніторингу транзакцій і поведінки клієнтів, щоб точно визначити шахрайську діяльність і боротися з нею.

Шахрайство з транзакцією — це купівля товарів і послуг за допомогою викрадених платіжних даних. Кіберзлочинці або викрадають платіжну інформацію зі слабко захищених корпоративних баз даних, або ж отримують потрібні дані методами соціальної інженерії, або купують її на темних форумах. Жертви часто не знають про це, поки не перевірять свої виписки та не помітять одну або кілька несанкціонованих транзакцій, лише після того можуть подати запит до банку. Ознаками таких транзакцій часто є нетипові для людини замовлення, поспіх з доставкою або ж інша адреса доставки.

Часто виявлення шахрайських транзакцій належить до обов'язків дата аналітиків, основними завданнями яких є аналіз історичних даних та створення моделей та методів, які дозволять швидко визначати підозрілі транзакції та блокувати їх. Машинне навчання є одним з ключових методів, що дозволяє виявляти злочинні перекази. Виявлення шахрайських транзакцій можна трактувати як задачу класифікації, коли фінансова установа має історичні промічтковані дані, на основі яких проводить аналіз нових даних. Або ж як задачу кластеризації, коли серед існуючого набору транзакцій необхідно виділити підозрілі та ті, які не викликають підозр.

Основними проблемами, які виникають при аналізі є те, що процес виявлення підозрілих транзакцій можна віднести до задачі виявлення аномалій (Anomaly detection). Класичні методи машинного навчання чудово працюють зі збалансованими наборами даних, коли кількість екземплярів кожного класу орієнтовно однакова. Натомість виявлення аномалій, за звичай, передбачає, що нестандартних випадків є значно менше, ніж стандартних, а отже вимагає особливих методів для роботи з такими датасетами. Ще однією специфікою таких задач є необхідність виявлення підозрілих фінансових операцій майже в реальному часі, адже сучасна екосистема транзакцій створена для швидкості та зручності. Навіть такий відносно складний процес, як заявка на кредит, можна здійснити за допомогою смартфона, тоді як більш рутинні покупки завершуються кількома натисканнями клавіш. Це високошвидкісне середовище може полегшити шахраям завершити свої злочини та зникнути до того, як їх вдасться виявити.

Ще однією проблемою може стати надмірно старанна система виявлення шахрайства, адже це може призвести до більшої кількості помилкових спрацьовувань. Це незручно для клієнтів, які в результаті можуть стати менш лояльними, і дорого для компаній, які повинні витрачати час і ресурси на обробку попередження. В роботі буде продовжено дослідження виявлення аномальних для клієнтів транзакцій з метою усунення вище наведених проблем.