Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя
(повне найменування вищого навчального закладу)
Факультет комп'ютерно-інформаційних систем і програмної інженерії
(назва факультету)
Кафедра кібербезпеки
(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр
(освітній рівень)

на тему: *"Розробка інформаційної системи для виявлення шахрайства з кредитними картками"*

Виконав: студент (ка)

Спеціальності:

125 «Кібербезпека»
(шифр і назва напряму підготовки, спеціальності)

Безруков О.О.
підпис (прізвище та ініціали)

Керівник *Максимчук О.О.*
Підпис (прізвище та ініціали)

Нормоконтроль *Лобур Т.Б.*
підпис (прізвище та ініціали)

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

Рецензент
підпис (прізвище та ініціали)

м. Тернопіль – 2022

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)
«__» _____ 2022 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня Бакалавр
(назва освітнього ступеня)

за спеціальністю 125 Кібербезпека
(шифр і назва спеціальності)

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

1. Тема роботи Розробка інформаційної системи для виявлення шахрайства з кредитними картками

Керівник роботи Максимчук О.О., асистент каф. КБ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «23» 03 2022 року № 4/7-178

2. Термін подання студентом завершеної роботи 17.06.2022

3. Вихідні дані до роботи

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

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

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Безпека життєдіяльності, основи хорони праці	Пулька Ч.В., проф. кафедри МТ		

7. Дата видачі завдання 16.02.2022 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи	16.02 – 19.02	*Виконано*
2.	Аналіз, формалізація завдання моніторингу шахрайських банківських транзакцій	20.02 – 27.02	*Виконано*
3.	Опрацювання джерел в галузі дослідження	28.02 – 16.03	*Виконано*
4.	Розроблення програмного коду	17.03 – 20.03	*Виконано*
5.	Проведення експериментальних досліджень	20.03-05.04	*Виконано*
6.	Оформлення розділу «Теоретична частина»	06.03 – 17.04	*Виконано*
7.	Оформлення розділу «Теоретичні дослідження та визначення специфікацій програмного забезпечення»	18.04 – 29.04	*Виконано*
8.	Оформлення розділу «Практична частина. Проєктування та розробка програмних компонентів для системи виявлення шахрайства»	30.04 – 13.05	*Виконано*
9.	Виконання завдання до підрозділу «Безпека життєдіяльності, основи хорони праці»	14.05 – 21.05	*Виконано*
10.	Оформлення кваліфікаційної роботи	22.05 – 05.06	*Виконано*
11.	Нормоконтроль	06.06 – 12.06	*Виконано*
12.	Перевірка на плагіат	10.06 – 15.06	*Виконано*
13.	Попередній захист кваліфікаційної роботи	16.06 – 19.06	*Виконано*
14.	Захист кваліфікаційної роботи	23.06.2022	

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

Керівник роботи _____ *Максимчук О.О.*
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Розробка інформаційної системи для виявлення шахрайства з кредитними карткам // Кваліфікаційна робота ОР «Бакалавр» // Безруков Олександр Олександрович // Тернопільський національний технічний університет імені Івана Пулюя, факультет комп'ютерно-інформаційних систем і програмної інженерії, кафедра кібербезпеки, група СБс-42 // Тернопіль, 2022 // С. , рис. – , табл. – , кресл. – , додат. – .

Ключові слова: ЕЛЕКТРОНИЙ БАНКІНГ, ФРАД-МОНІТОРІНГ, ШАХРАЙСТВО, ТРАНЗАКЦІЇ, RANDOM FOREST

Актуальність роботи визначається необхідністю впровадження систему фрод-моніторингу для мінімізації збитків від шахрайства у банківських транзакціях шляхом проектування та розробки системи моніторингу шахрайських банківських транзакцій. Розроблені заходи дозволять знизити ризики фінансових втрат клієнтами банку від шахрайства з кредитними картами при проведенні платежів.

Метою дипломного проекту є мінімізації збитків від шахрайства у банківських транзакціях шляхом проектування та розробки системи моніторингу шахрайських банківських транзакцій.

Для досягнення поставленої мети потрібно вирішити такі завдання:

– виконати аналіз, формалізацію завдання моніторингу шахрайських банківських транзакцій;

– виконати аналіз існуючих методів вирішення поставленого завдання;

– виконати теоретичні дослідження та визначення специфікацій програмного забезпечення;

– виконати проектування програмного забезпечення.;

– провести експериментальні дослідження розробленого програмного забезпечення.

ANNOTATION

Development of an Information System to Identify any Credit Cards Frauds // Qualification thesis of educational level "Bachelor" // Bezrukov Oleksandr Oleksandrovych // Ternopil Ivan Puluj National Technical University, Faculty of Computer Information Systems and Software Engineering, Department of Cybersecurity, СБс-42 group // Ternopil, 2022 // P. , fig. - , table. - , chair. - , added. - .

Keywords: ELECTRONIC BANKING, FRAUD MONITORING, FRAUD, TRANSACTIONS, RANDOM FOREST

The relevance of the work is determined by the need to implement a fraud monitoring system to minimize losses from fraud in banking transactions by designing and developing a monitoring system for fraudulent banking transactions. The developed measures will reduce the risks of financial losses by the bank's customers from credit card fraud when making payments.

The aim of the diploma project is to minimize losses from fraud in banking transactions by designing and developing a system for monitoring fraudulent banking transactions.

To achieve this goal, you need to solve the following tasks:

- perform analysis and formalization of the task of monitoring fraudulent bank transactions;
- perform an analysis of existing methods for solving the problem;
- perform theoretical research and determine software specifications;
- perform software design.;
- conduct experimental studies of the developed software

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ІС	Інформаційна система
VPN	Віртуальна приватна мережа
IPsec	Internet Protocol Security
CAP	Central Access Policy
SSL / TLS	Transport Layer Security

ВСТУП

Вітчизняні банківські установи, як і інші суб'єкти економіки, які здійснюють свою діяльність в умовах непередбачуваності, невизначеності, загроз та небезпек. У той самий час роль банків постійно посилюється. Банківська система є важливим елементом економіки та надає значний, різнобічний вплив на всі аспекти життєдіяльності суспільства. Банки виконують широкий спектр операцій, таких як: акумулювання коштів та заощаджень, здійснення кредитно-розрахункових та інших операцій. Крім того, майже всі фінансові потоки фізичних та юридичних осіб проходять через банки.

Сучасні банки функціонують за умов дестабілізуючого впливу як зовнішніх, і внутрішніх чинників. Наразі гостро постає питання забезпечення економічної безпеки банківської системи держави як основи фінансової системи.

Наявність вищезгаданих проблем зумовлюється низкою факторів, серед яких, перш за все, слід зазначити низький рівень економічної безпеки банків, зумовлений недоліками функціонування діючих систем управління безпекою банківського бізнесу, забезпечувати реалізацію основних інтересів, пріоритетних цілей банків, захист від впливу дії негативних факторів.

Однією з основних складових системи економічної безпеки є моніторинг банківських операцій як форма протидії шахрайству з різних функціональних сфер діяльності банку.

Постійний моніторинг шахрайських дій як один із обов'язкових елементів системи управління економічною безпекою банку є, на наш погляд, важливим та актуальним питанням у контексті забезпечення сталого та ефективного функціонування вітчизняної банківської системи.

Функціонування банківських установ на сучасному етапі розвитку економіки спонукає їх до використання інноваційних технологій банківського обслуговування, які дають можливість зайняти провідні конкурентні позиції на ринку. Впровадження різних форм електронного банкінгу підвищило ефективність обслуговування клієнтів та зробило його економічно вигідним, проте такі види діяльності банків стали особливо вразливими до шахрайських

дій злочинців. В електронному банківському обслуговуванні клієнтів все частіше застосовуються такі види злочинів, як шахрайство з платіжними картками, розповсюдження комп'ютерних вірусів, незаконне зняття коштів із банківських рахунків, викрадення конфіденційної інформації тощо. За таких умов дослідження видів основних шахрайств електронного банкінгу та методів мінімізації ризиків,

Теоретичні основи сутності та форм електронного банкінгу та видів шахрайств, а також методи управління ризиками електронної форми обслуговування клієнтів банків досліджували такі вчені, як О. Волчанка, Б. Кінг, М. Зубко, І. Домінова, Г. Шпаргало, Л. Капінус, І. Пасічник, Г. Карчева, В. Бутузов та інші.

Актуальність роботи визначається необхідністю впровадження систему фрод-моніторингу для мінімізації збитків від шахрайства у банківських транзакціях шляхом проектування та розробки системи моніторингу шахрайських банківських транзакцій. Розроблені заходи дозволять знизити ризики фінансових втрат клієнтами банку від шахрайства з кредитними картами при проведенні платежів.

Об'єктом дослідження є процеси фрод-моніторингу шахрайства у банківських транзакціях.

Предметом дослідження є методи та інструменти проектування та розробки системи моніторингу шахрайських банківських транзакцій.

Метою дипломного проекту є мінімізації збитків від шахрайства у банківських транзакціях шляхом проектування та розробки системи моніторингу шахрайських банківських транзакцій.

Для досягнення поставленої мети потрібно вирішити такі завдання:

– виконати аналіз, формалізацію завдання моніторингу шахрайських банківських транзакцій;

– виконати аналіз існуючих методів вирішення поставленого завдання;

– виконати огляд аналогічних рішень;

– виконати теоретичні дослідження та визначення специфікацій програмного забезпечення;

– виконати проектування програмного забезпечення.;

– виконати розробку макетів програмних компонентів;

– провести експериментальні дослідження розробленого програмного забезпечення.

1 ТЕОРЕТИЧНА ЧАСТИНА

1.1 Аналіз, формалізація завдання моніторингу шахрайських банківських транзакцій

Електронний банкінг набуває неймовірної популярності, тому дослідження специфіки цієї форми банківського обслуговування клієнтів має особливе значення. При електронній формі обслуговування клієнтів небезпекою номер один є шахрайства та кіберзлочини, а тому для банківських установ, які використовують інформаційні технології у процесі банківського обслуговування клієнтів, одним із пріоритетних завдань є ідентифікація та оцінка загроз шахрайству електронного банкінгу.

Дослідження визначень сутності електронного банкінгу у науковій літературі та за змістом не збігаються. Так роботі [1] електронний банкінг розглядається як діяльність банку з надання комплексу послуг клієнтам за допомогою комп'ютерних технологій.

На думку авторів, під електронними банківськими послугами слід розуміти дії банку, створені задля вдосконалення і реалізацію звичних банківських операцій шляхом використання інформаційних систем [1].

Автори роботи [2] ототожнюють електронний банкінг та електронну банківську діяльність як синонімічні поняття та визначають його як «процес надання послуг у банку та здійснення банківських операцій з використанням автоматизованих систем, у тому числі електронних каналів зв'язку».

Поняття електронного банкінгу досліджує Механчук Н.М. інтерпретуючи його як послугу банку, що передбачає дистанційне управління рухом фінансових коштів на картковому рахунку за допомогою електронних мереж та обладнання [3]. Тищенко О.І. наводить подібне визначення електронного банкінгу, проте зазначає, що електронне банківське обслуговування є спеціальним інструментом, щоб надати послуги, використовуючи новітні інформаційні та телекомунікаційні технології [4].

Часто у науковій економічній літературі використовується таке визначення електронного банкінгу, як «забезпечення можливостей для клієнтів банків отримувати віддалений доступ до своїх банківських рахунків через інформаційно-телекомунікаційні системи та, як мінімум, здійснювати переказ фінансових коштів між ними» [5].

Чіткого визначення економічного змісту електронного банкінгу також не має і у іноземних учених. Більшість зарубіжних учених описують електронний банкінг як електронний зв'язок (через комп'ютер або мобільний телефон) між банком та клієнтом для підготовки, управління та контролю за проведенням фінансових операцій та послуг [6].

Нерідко поняття електронного банкінгу використовують як узагальнене визначення на дослідження особливостей дистанційного електронного обслуговування клієнтів банку.

Визначення електронного банкінгу наводить Базельський комітет із банківського нагляду «електротронний банкінг – це надання роздрібних та незначних за обсягом банківських продуктів та послуг через електронні банківські канали, а також значних за обсягом електронних платежів та інших оптових банківських послуг електронним шляхом» [7].

Наведені визначення дозволяють зробити висновок, що немає єдиного підходу до розуміння електронного банкінгу серед науковців та практиків банківської справи. Під сутнісним змістом більшість учених вважають, що під електронним банкінгом слід розуміти вид дистанційного банківського обслуговування, через який здійснюється процес надання банківських послуг лише за допомогою використання Інтернету та мобільного зв'язку. Проте вважаємо, що «електронний банкінг» та «дистанційне банківське обслуговування» не є синонімічними поняттями.

Так, електронне банківське обслуговування клієнтів не завжди здійснюється дистанційно, адже обслуговування через термінали самообслуговування та банкомати зазвичай відбувається у відділеннях банківських установ, а дистанційне банківське обслуговування відбувається лише тоді, коли клієнт не обслуговується у відділенні. Отже, поняття

«електронний банкінг» ширше за поняття «дистанційне банківське обслуговування». Спільним між електронним банкінгом та дистанційним обслуговуванням є те, що клієнт сам виступає операціоністом при використанні цих способів банківського обслуговування.

Враховуючи наведені вище визначення електронного банкінгу та специфічні характерні ознаки цього способу обслуговування клієнтів вважаємо, що під поняттям «електронний банкінг» слід розуміти спосіб банківського обслуговування, за допомогою якого надаються традиційні та інформаційні послуги банківського обслуговування через різні автоматизовані форми інформаційних технологій.

В умовах функціонування та розвитку форм електронного банкінгу збільшується загроза кіберзлочинності та шахрайства. Багато вітчизняних та зарубіжних учених приділяють увагу дослідженню питань безпеки банківської діяльності та протидії шахрайству електронного банкінгу. Так Малишевська О.О. доводить, що безпека банківської діяльності - це стан стійкої життєдіяльності, при якому забезпечується реалізація мети банку та основних його інтересів, захист від внутрішніх та зовнішніх факторів дестабілізації незалежно від умов функціонування. Вважається, що суттєвою загрозою як зовнішнього, так і внутрішнього походження безпеки банку є ризик шахрайства, наголошується, що предметом шахрайських посягань насамперед є гроші (75%), товарно-матеріальні цінності (20%), а 5% шахрайств посідає викрадення інтелектуальної розробки банку [8, 9]. Домінова І.В. зазначає, що злочином при використанні платіжних карток вважається:

- грошові кошти, які власники платіжних карток зберігають на своїх рахунках;
- інформація, завдяки якій можна здійснювати перекази коштів;
- різні послуги та майнове забезпечення підприємств, які здатні здійснювати карткові розрахунки.

На інформаційній безпеці в умовах електронного банкінгу наполягає і закордонний дослідник Бретт Кінг, зазначаючи, що шахрайство з персональними даними клієнтів є однією з головних проблем обслуговування клієнтів через

електронні канали [11]. Базельський комітет з банківського нагляду також наголошує на необхідності ідентифікації шахрайства електронного банкінгу та мінімізації ризику цієї форми банківського обслуговування клієнтів. Так, у Базелі III зазначено, що ризик шахрайства є частиною операційного ризику електронного банкінгу, та зазначено, що внутрішнє та зовнішнє шахрайство у банку описується як події, пов'язані з операційним ризиком банку [12]. Тобто в цьому випадку шахрайство є не окремим підвидом ризику, а загрозою – потенційними чи реальними діями певних суб'єктів, здатних завдати конкретному банку матеріальної чи моральної шкоди [8]. Особливістю загроз є те, що вони є конкретними та призводять до фінансових втрат банків та їхніх клієнтів. Узагальнюючи сказане, можна визначити такі основні об'єкти шахрайства електронного банкінгу, як:

1) конфіденційна інформація про клієнтів банку (пін-коди, CVV-код)

2) логіни та паролі доступу до форм електронного банкінгу (Інтернет банкінг та мобільний банкінг);

3) фінансові ресурси банку та клієнтів банку, доступ до яких можливий за умови викрадення вищевказаних об'єктів шахрайства [11].

Отже, можна констатувати, що недостатньо уваги приділяється процесу ідентифікації та оцінювання ризику шахрайства електронного банкінгу, що ускладнює подальший процес досліджень.

Ідентифікація шахрайства електронного банкінгу пов'язана з виявленням джерел, що призводять до появи ризику, властивого цьому виду банківської діяльності. Наразі найбільшими ризиками шахрайства електронного банкінгу є:

1) шахрайства з платіжними картками;

2) шахрайства через мобільний телефон та Інтернет.

З метою ідентифікації джерел прояву ризиків шахрайства, властивих кожній із наведених форм функціонування електронного банкінгу, проаналізуємо їх види. Так, шахрайства з платіжними картками, з мобільним телефоном та Інтернетом переважно здійснюються за допомогою таких методів, як соціальна інженерія (до яких належить фішинг, вишинг), фармінг, трешинг, основною метою є отримання конфіденційної інформації, зазначеної на

платіжній картці. Шахрайства з платіжними картками є найпопулярнішим видом шахрайства в Україні. Розглянемо докладніше кожен із видів шахрайства, що дозволить надалі вчасно їх ідентифікувати та розробити методи запобігання цим ризикам. Соціальна інженерія - це метод управління діями та поведінкою людини з метою отримання від неї певної інформації або здійснення певних дій. Так, Домінова І.В. наголошує на необхідності якісного підбору персоналу, щоб уникнути кадрового ризику, одним із проявів якого є шахрайство або розкриття конфіденційної інформації працівником банку [13]. Найбільш популярними видами шахрайства з платіжними картками з використанням методів соціальної інженерії є фішинг та вишинг, які спрямовані на неуважних чи довірливих користувачів банківськими платіжними картками. одним із проявів якого є шахрайство чи розкриття конфіденційної інформації працівником банку [13]. Найбільш популярними видами шахрайства з платіжними картками з використанням методів соціальної інженерії є фішинг та вишинг, які спрямовані на неуважних чи довірливих користувачів банківськими платіжними картками. одним із проявів якого є шахрайство чи розкриття конфіденційної інформації працівником банку [13]. Найбільш популярними видами шахрайства з платіжними картками з використанням методів соціальної інженерії є фішинг та вишинг, які спрямовані на неуважних чи довірливих користувачів банківськими платіжними картками.

Фішинг (від fishing – риболовля) – спосіб отримання персональної інформації (номерів карт, паролів, банківських рахунків) шляхом розсилки електронних листів від імені банку або відомих компаній, що містять посилання на підроблені сайти, емітують роботу справжніх. Надалі зазначена інформація використовується для ініціювання з-за кордону неналежних грошових транзакцій поштових відправлень [14]. Техніка фішингу була описана ще 1987 року, проте термін phishing почали використовувати у 1996 року у США у зв'язку з активним розвитком цього виду шахрайства.

Наступним видом шахрайства вішинг (від англ. Voice – «голос» – є телефонне шахрайство, пов'язане з виманюванням конфіденційної інформації про реквізити банківських карток та їх паролів та відміною до переказу коштів

на карту злодіїв (шахрай дзвонить потенційній жертві шахрайства та надається працівником правоохоронних органів). служби безпеки банку (у 94% випадків) та змушує власника платіжної картки назвати її реквізити) [12, 13, 14].

Наступний різновид шахрайства з платіжними картками – це фармінг. До цього виду шахрайства є те, що фармінг-технології дозволяють змінити IP-адресу сайту, і при вході на web-сторінку легітимної організації проводиться перенаправлення на підроблену, створену для збору конфіденційної інформації [13, 14]. Слід зазначити, що під впливом психологічних методів чи неуважності людина стає жертвою шахраїв, якщо йдеться про фішинг і вишинг, тоді як під час фармінгу шахрай не контактує з жертвою, тому захистом від фармінгу може бути встановлення на персональний комп'ютер ліцензійного антивірусу.

Наступна підгрупа шахрайств електронного банкінгу – це шахрайства з мобільним телефоном та Інтернетом. До цієї групи віднесено всі види шахрайства, для реалізації яких використовується мобільний телефон та мережа Інтернет. Так, одним із видів шахрайства за допомогою телефону та Інтернету є змішинг, коли шахраї надсилають жертві SMS-повідомлення для переходу на фішинговий сайт або для відправки у відповідь на SMS-повідомлення реквізитів платіжної картки [15].

Отже, за походженням шахрайські операції у банківській сфері можуть бути зовнішніми (здійснюються клієнтами банку або третіми особами) та внутрішніми (здійснюються персоналом банку).

У роботі розглядаються шахрайські операції транзакційного типу, об'єктом яких є банківські платіжні картки.

При банківському шахрайстві (як, втім, будь-якому іншому) слід відзначити найбільш характерну особливість - гнучкість. Справді, з появою нових банківських технологій виникають і нові шахрайські технології, спрямовані на використання всіх їхніх можливих прогалин. Таким чином, шахрайство у сфері банківських транзакцій можна охарактеризувати як навмисний акт бездіяльності або вчинення будь-якої дії при здійсненні банківської операції, що призводить до неправомірної вигоди для будь-якої особи за рахунок одночасної шкоди для іншої особи або банку.

Як відомо, банки вдаються до комплексних, системних заходів запобігання шахрайським транзакціям. Так, у більшості банків використовується скорингова модель визначення ймовірності шахрайства [16]. Наприклад, при онлайн-оплаті будь-якою карткою банку необхідно пройти двофакторну ідентифікацію (перший ступінь – реквізити картки, другий ступінь – підтвердження оплати за допомогою СМС-коду). У цілому нині, типова система для онлайн-банкінгу скорингова модель використовує визначення нетипових операцій із допомогою патернів. В результаті визначено підозрілі операції, які зазнають більш детального контролю та можуть бути автоматично скасовані.

Тривожним є те, що попри всі зусилля протидії з боку правоохоронних органів (насамперед кіберполіції) тенденція до зростання шахрайства з використанням соціальних технологій є явною. Посилення протидії даному виду шахрайства можливе лише за умови широкої інформаційної кампанії, спрямованої на максимальне висвітлення можливих шахрайських технологій, зростання фінансової та технічної грамотності користувачів і рівня їх поінформованості про нові види шахрайських схем. Гнучкості та адаптованість зловмисників необхідно протиставити оперативне викриття шахрайських схем та доведення інформації до користувачів платіжних карток.

Про можливості протидії шахрайським операціям, слід зазначити, що існує два напрямки: технічний та соціальний, залежно від сфери застосування шахрайських технологій. Технічний шлях запобігання банківському шахрайству активно вдосконалюється, постійно розвиваються технічні інструменти безпеки банківських транзакцій, розкриваються нові шахрайські схеми. Цей напрямок «приречений» на постійну еволюцію. Також дуже важливо звернути увагу на сегмент мобільних та безконтактних платежів.

Можливі заходи щодо запобігання та припинення шахрайства:

1. Удосконалювати законодавчу базу у частині мобільного зв'язку та електронного грошового обігу, у тому числі посилити відповідальність за злочини у сфері високих технологій.

2. Сформувати єдині правила для всіх операторів мобільного зв'язку із встановленням відповідальності за бездіяльність при шахрайстві з використанням обладнання або програмного забезпечення оператора.

3. Забезпечити ефективний державний нагляд за належним проведенням ідентифікації клієнтів з метою повного дотримання законодавства про протидію легалізації (відмиванню) доходів, одержаних злочинним шляхом, та фінансування тероризму.

4. Посилити роботу з формування відповідальної економічної поведінки та підвищення фінансової грамотності населення, особливо щодо безпечного використання електронних засобів платежів.

Доцільним є створення єдиної інформаційної системи як єдиного сервісу отримання інформації від операторів зв'язку. Можливості єдиної інформаційної системи:

– Додаткова верифікація користувачів;

– Актуалізація абонентських баз колекторських служб;

– підвищення рівня захищеності громадян – абонентів рухомого радіотелефонного зв'язку;

– підвищення рівня безпеки переказів коштів із використанням мобільного телефону;

– належне виконання кредитними та іншими фінансовими організаціями нормативних вимог.

Дані переваги дозволяють звузити сферу передумов банківського шахрайства, які у сфері дистанційних платежів можуть бути такими:

– широке поширення комунікаційних пристроїв серед населення, не підготовленого до протидії шахрайству;

– використання технічних засобів для здійснення платіжних операцій в автоматичному режимі без особистої присутності власника коштів для здійснення платежу або передачі грошей іншій особі;

– рух грошових коштів на основі єдиних принципів та правил комунікації та платежів;

– недостатність заходів протидії шахраям.

Найскладнішою є проблема захисту безконтактних, у тому числі «мобільних» платежів, оскільки в даному випадку основною загрозою є робота шкідливого ПЗ, яке здійснює втручання у платіжну систему без відома клієнта та обминаючи ідентифікацію при здійсненні транзакції [13, 14].

Тому вже сьогодні й потрібне створення відповідної нормативної бази для закріплення зон відповідальності мобільних операторів у сфері технічного забезпечення проведення фінансових транзакцій. Створивши адекватну систему запобігання шахрайству, заблокувавши можливості розсилки фішингових повідомлень SMS та можливості несанкціонованої заміни SIM карт, посиливши контроль за реалізацією контрактів мобільних операторів, можна суттєво ускладнити для злочинців процес використання викрадених коштів, здійснити оперативне блокування та повернення викрадених сум.

Отже, найбільш ефективним заходом протидії шахрайським операціям з боку клієнтів банків є максимальна поінформованість користувачів про необхідні дії у разі шахрайських атак (Наприклад, психологічного тиску при використанні технологій соціальної інженерії). Також користувачі платіжних карток повинні бути повністю поінформовані про технічні заходи щодо запобігання втраті коштів.

Превентивні заходи протидії випадків банківського шахрайства мають бути обов'язковими, а також постійно оновлюватися і доповнюватися відповідно до розвитку технологій.

1.2 Аналіз існуючих методів вирішення поставленого завдання.

Дослідженню питань виявлення шахрайств у банках, які здійснюються персоналом банку, присвячено багато робіт вітчизняних та зарубіжних учених. У роботі [17] наведено класифікацію способів шахрайства в банківській сфері залежно від наміру, типових слідів, етапу та способу внесення свідомо неправдивої інформації в документи, виду банківської операції, характеру здійснення банківської операції, наявності певних домовленостей, дій з

підготовки та приховування шахрайства. У роботі [18] розглянуті шахрайства, які здійснюються персоналом банку, у контексті спотворення фінансової звітності банку. Останнє може бути результатом фальсифікацій з первинними документами чи незаконного присвоєння активів (грошових надходжень, викрадення фізичних активів чи інтелектуальної власності, використання майна у власних цілях тощо).

У роботі [19] наведено характеристику шахрайських дій залежно від виду банківської операції. У роботі [20] розглянуто технологію Data Mining (асоціативний аналіз), яка може бути використана для виявлення шахрайств у банках, що здійснюються персоналом банку. У роботі [21] наведено приклади використання штучних нейронних мереж виявлення шахрайств персоналу з урахуванням публічної фінансової звітності.

Незважаючи на наявність значної кількості наукових публікацій з досліджуваної проблематики, нині відсутня систематизація методів економіко-математичного моделювання для виявлення шахрайств у банках, які здійснюють персонал банку.

Згідно з визначенням Базеля II, шахрайство є частиною операційного ризику банку та класифікується як внутрішнє та зовнішнє. Внутрішнє шахрайство – це ризик несподіваних фінансових, матеріальних та репутаційних втрат банку внаслідок шахрайських дій його персоналу. Найбільш поширеними шахрайствами у банках є відмивання «брудних» грошей, шахрайства з кредитами та незаконне привласнення активів [22]. Першою причиною шахрайства є фінансові труднощі шахрая. Другий – існування можливості для здійснення шахрайства. Третьою – впевненість шахрая у існуванні вагомих причин скоєння їм шахрайських дій.

Відомо, що шахрайство з кредитними картками, включаючи незаконне використання кредитної картки або її інформації без відома власника. У роботі [24] зазначено, що сьогодні виявлення таких шахрайств широко застосовуються: логістична регресія, яка здатна вирішувати категоріальні класифікаційні завдання; метод опорних векторів (SVM, Support Vector Machine), який здатний обробляти незбалансовані дані та складні зв'язки між змінними; зручні у

використанні дерева рішень; випадковий ліс (Random forest) самоорганізуються карти Кохонена (SOM, Self-Organizing Map); нечітка логіка, що підвищує ефективність управлінських рішень.

На думку авторів роботи [25], за наявності невизначеностей найкращі результати дає застосування нечітких методів. Однак слід враховувати, що основним недоліком останніх є їхня надто висока точність, тому з метою її підвищення краще використовувати гібридні нейро-нечіткі системи (ANFIS, Adaptive Neuro-Fuzzy Inference System). Незважаючи на цілком пристойні результати, які дає метод опорних векторів, він є чутливим до збільшення кількості даних і не може підтримувати великі набори даних [24].

Для виявлення спотворень фінансової звітності у банківській сфері широко застосовуються: нейронні мережі, які здатні впоратися із завданнями без алгоритмічного вирішення; мережі Байєса, що використовуються для виявлення аномалій; генетичні алгоритми, що використовуються для бінарної класифікації; текст майнінг (text mining), який використовується для кластеризації та виявлення аномалій. У роботі [24] наголошується також, що сучасною тенденцією виявлення шахрайства є використання гібридних методів, які використовують сильні сторони різних методів.

Виявлення фінансового шахрайства включає моніторинг поведінки власників карткових рахунків з метою виявлення їхньої небажаної поведінки.

У роботі [26] при цьому використовується генетичний алгоритм, у якому замість максимізації кількості правильно класифікованих транзакцій визначається цільова функція зі змінними, що показують втрати від хибної класифікації. Таким чином, правильна класифікація одних транзакцій важливіша за інші. На першому етапі запропонованого в [26] алгоритму вводяться вихідні дані - транзакції власника карткового рахунку, кожна з яких має набір стандартизованих атрибутів, що описують поведінку власника карткового рахунку.

Вихідні дані включають, наприклад, такі змінні:

– кількість разів використання картки;

– місцезнаходження картки у момент її використання;

– баланс, доступний на картковому рахунку;

– середньодобова сума грошей, яка знімалася власником карткового рахунку тощо.

На другому етапі в результаті роботи генетичного алгоритму розраховуються критичні значення вищезгаданих змінних. Далі ці критичні значення використовуються разом із технологіями Data Mining.

У роботі [27] для моніторингу поведінки власників карткових рахунків використовується прихована марківська модель (HMM, Hidden Markov Model), яка спочатку вчиться нормальних дій власника картки, а потім використовується для виявлення шахрайської поведінки. У роботі [28] моніторингу поведінки власників карткових рахунків використовується теорія нечіткої логіки.

З того часу, як багато банківських і платіжних операцій перейшли в область інформатизації, шахрайство в цій сфері активно розвивається. Найбільш відомі атаки на банківські системи за останні кілька років були виконані злочинними угрупованнями Cobalt, Carbanak, Lazarus та Lurk.

Зловмисники здійснюють атаки на системи міжбанківських переказів, картковий процесинг, управління банкоматами, інтернет-банкінг та платіжні шлюзи. За даними звіту Positive Technologies, зловмисники використовують простий сценарій для здійснення атаки, що складається з 5 послідовних етапів:

– Попередня розвідка та підготовчі роботи.

– Проникнення у внутрішню мережу.

– Закріплення у внутрішній мережі та розвиток атаки.

– Компрометація банківських систем та розкрадання коштів.

– Приховування слідів.

Ці етапи актуальні при фішинг, зараженні комп'ютера або смартфона жертви відомим раніше шкідливим, проведенні атак типу man-in-the-middle, використанні кейлогерів і навіть вразливостей нульового дня

Системи протидії шахрайству можуть мати у своєму арсеналі такі технології та можливості:

1. Текстова аналітика, яка виконується за допомогою технологій пошуку, категоризації контенту та вилучення сутностей.

2. Розрахунок статистичних параметрів, який використовується для виявлення відхилень, які могли б вказати на шахрайство.

3. Мережна аналітика, що використовується для ідентифікації з'єднань, виявлення закономірностей. Gap-тестування має на увазі виявлення будь-яких відсутніх елементів у послідовних даних там, де їх не повинно бути.

4. Підтвердження дати входу використовується для оцінки невідповідного або підозрілого часу для розміщення або введення інформації.

5. Контрольоване машинне навчання, яке виготовляють основі історичних даних, що дозволяє виявляти певні шаблони.

6. Навчання без вчителя, що має на увазі аналіз та оцінку даних, які не містять відомостей про виявлене шахрайство. Використовується виявлення нових аномалій.

Функція у всіх антифрод-систем єдина — виявляти та запобігати шахрайству. Однак вони можуть по-різному вирішувати це завдання і порівнювати антифрод-системи без проведення додаткової класифікації є неправильним рішенням. Так, наприклад, є так звані core-системи - потужні аналітичні платформи, що дозволяють реалізовувати логіку в окремих сегментах (ДБО або процесинг банківських карток), також існують спеціалізовані системи, що контролюють параметри пристроїв та ризики на їхньому боці. І в той же час розробляються окремі системи, ув'язнені під розпізнавання фото, відео, мови. Багато хто з систем не конкурує, а, навпаки, доповнює функції один одного.

Компанія Featurespace заснована у 2008 році (м. Кембридж, Великобританія). Компанія була створена професором з Кембриджського університету з метою розробки механізму адаптивної поведінкової аналітики, яка дозволяє здійснювати захист від шахрайства на основі виявлення аномалій. Система ARIC White Label (рис.1.1) від компанії Featurespace належить до класу загально аналітичних платформ. Система використовує технології машинного навчання для забезпечення захисту від шахрайських транзакцій за різними типами платежів (карти, електронні гаманці та ін.) як реального часу. В ARIC

White Label створюються моделі нормальної поведінки клієнтів, відхилення у яких надалі реєструє система. Для різних клієнтів можуть створюватися різні правила аналізу,

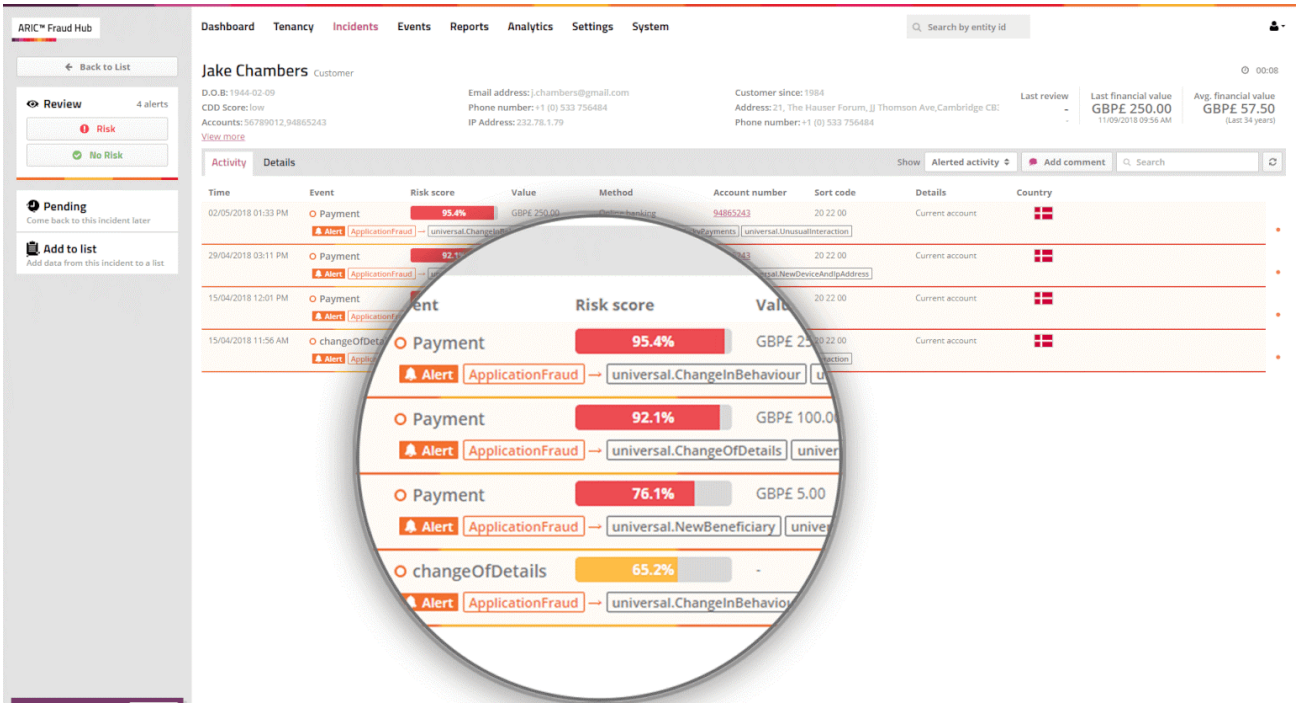

Рисунок 1.1 – Інтерфейс системи ARIC White Label

Особливості ARIC White Label:

– Індивідуальне налаштування інтерфейсу системи та надання доступу до системи.

– Запобігання не лише типовим шахрайським діям, а й заходам з відмивання грошей (AML).

– Використання технологій машинного навчання та поведінкового аналізу, що дозволяють забезпечити захист від шкідливих програм, атак ботів та шахрайства зі зворотними платежами.

– Можливість використання ARIC White Label як хмарного сервісу.

Компанія FICO заснована у 1956 році (м. Сан-Хосе, штат Каліфорнія, США). Компанія спеціалізується на розробці програмного забезпечення для передиктивної аналітики та прийняття рішень, у тому числі рішень з оцінки кредитних ризиків, а також зменшення збитків від шахрайських дій. Система

FICO Application Fraud Manager (рис.1.2) від компанії FICO відноситься до загальноаналітичних платформ та здійснює ідентифікацію спроб шахрайства в режимі реального часу за рахунок аналітичної системи, яка використовує технології машинного навчання та адаптивного аналізу. Рішення може бути встановлене як локально, так і використовуватися за технологією SaaS. Система дозволяє запобігати спробам шахрайства з боку третіх осіб, а також спробам навмисного зловживання привілеями облікових записів,

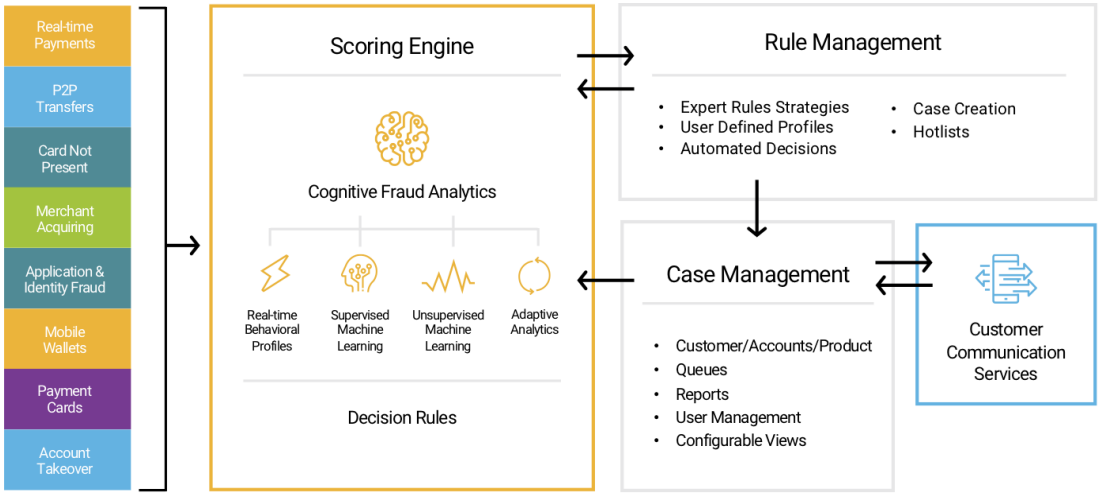

Рисунок 1.2 – Схема роботи системи FICO Application Fraud Manager

Особливості FICO Application Fraud Manager:

– Використання технологій машинного навчання.

– Можливість використання платформи як хмарного сервісу.

– Розширений аналіз посилань та соціальних мереж для моніторингу клієнта.

– Розслідування інцидентів з можливістю розподілу ролей та наданням звітності.

Компанія «Фродекс» заснована у 2011 році. Компанія спеціалізується на послугах із забезпечення інформаційної безпеки, розробки та впровадження інтелектуальних систем виявлення шахрайських платежів, систем обробки даних, проведення розслідувань у сфері інформаційної безпеки. Флагманським

рішенням «Фродекс» є система виявлення шахрайських платежів FraudWall, якій надано клас інформаційних систем для вирішення специфічних галузевих завдань. Систему FraudWall (рис.1.3) від компанії Фродекс можна віднести до класу загальноаналітичних платформ. Вона призначена для запобігання крадіжці коштів клієнта в системах дистанційного банківського обслуговування (ДБО), боротьби з внутрішнім шахрайством (наприклад, несанкціоновані платежі в АБС), запобігання крадіжці коштів банку через АРМ КБР. Коли система виявила підозрілий платіж, вона здійснює дзвінок клієнту і веде із нею живе спілкування, розпізнаючи відповіді клієнта. Після завершення дзвінка FraudWall приймає рішення про виконання платежу або зупинення операції.

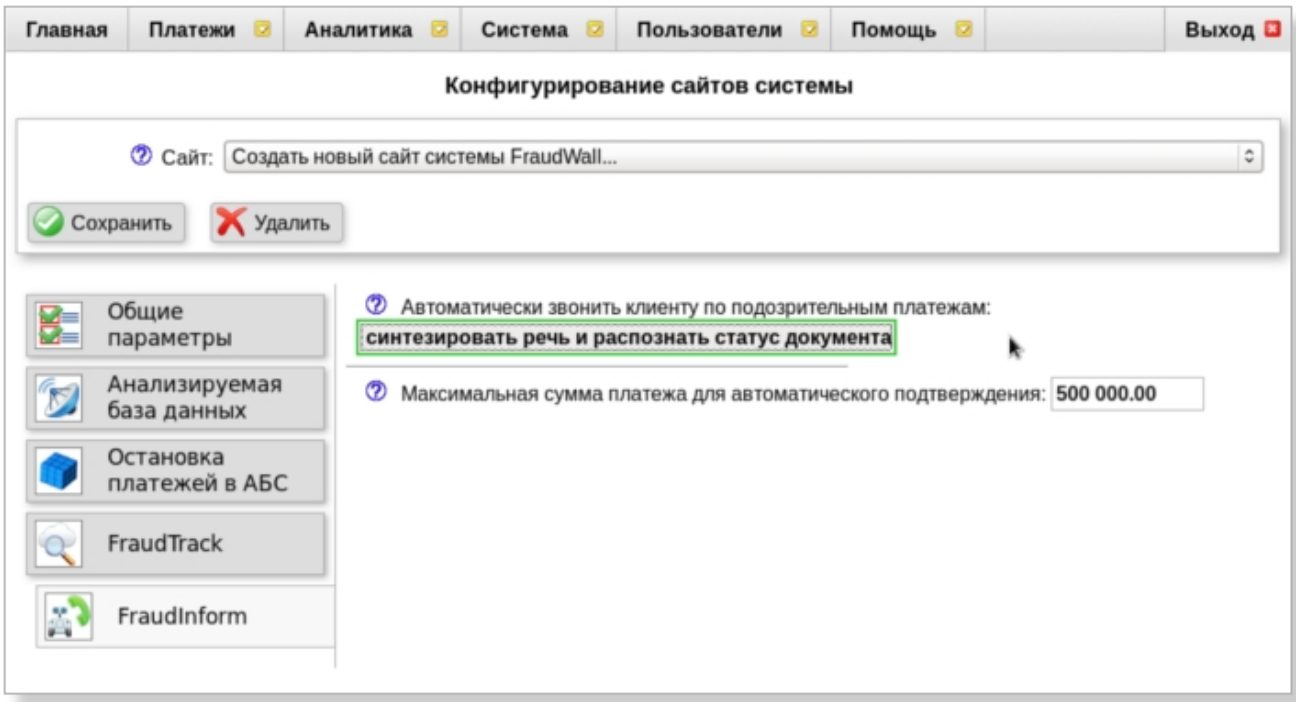

Рисунок 1.3 – Інтерфейс системи FraudWall

Особливості FraudWall:

– Масштабованість рішення.

– Інтеграція зі спеціалізованою системою розпізнавання мови VoiceNavigator від Центру мовних технологій.

– Інтеграція з міжбанківською системою FraudMonitor та чорними списками.

– Декілька варіантів впровадження в мережу банку (безпосереднє підключення до бази даних ДБО або АБС, підключення в ролі веб-сервера системи ДБО, підключення в ролі проміжного проксі-сервера між клієнтом банку та веб-сервером системи ДБО).

Шахрайство у банківській сфері продовжує прогресувати з кожним роком. А тому зростає ринок систем протидії банківському шахрайству. Лідерами у цій сфері є США. Проте забезпечення безпеки від фроду є актуальним і для російських фінансових організацій. При проектуванні та розробці системи моніторингу шахрайства у банківських транзакціях необхідно насамперед визначитися з тим, які завдання їй слід виконувати.

У таблиці 1.1 наведено порівняльний аналіз існуючих систем виявлення шахрайства.

Таблиця 1.1 – Порівняльні параметри програм виявлення шахрайства.

Параметр	ARIC White Label	FICO Application Fraud Manager	FraudWall
Можливість масштабування	Так	Так	Так
Типи шахрайства, що виявляється	Типові шахрайські дій, та заходи щодо відмивання грошей (AML).	Типові шахрайські дій	Типові шахрайські дій
Використовувані технології	Технології машинного навчання та поведінкового аналізу	Технології машинного навчання та аналіз соціальних мереж	Технології машинного навчання
Хмарні сервіс	так	так	так
Аналітична звітність	Ні	Так	ні
Аналіз соціальних мереж	Ні	Так	Ні
Інтеграція з іншими системами	Ні	Ні	Так

1.3 Висновки до розділу 1

У розділі проаналізовано сучасні економіко-математичні методи виявлення шахрайств у банках. Якісні методи враховують невизначеність за допомогою суб'єктивних експертних оцінок. Кількісні методи базуються на традиційному математичному апаратові, а методи машинного навчання – на технологіях штучного інтелекту. Вони враховують невизначеність за допомогою засобів статистики та теорії ймовірностей. Оптимальними для врахування невизначеності та виявлення шахрайства у банках є гібридні методи, що використовують сильні сторони різних підходів.

Підсумовуючи викладене, слід зазначити, що обслуговування банками клієнтів з використанням різних форм електронного банкінгу пов'язане з різними видами шахрайства, тому з боку банківських установ вимагає управління ризиками, які притаманні такому виду банківського бізнесу.

З метою запобігання ризикам шахрайства, що виникають під час електронного обслуговування клієнтів, банківські установи повинні будувати чітку систему їхньої ідентифікації та мінімізації. До основних методів мінімізації ризиків шахрайства електронного банкінгу можна віднести:

1) підвищення фінансової грамотності клієнтів за їх поінформованість про види шахрайств електронного банкінгу з метою скорочення фінансових втрат від шахрайства. цього можна досягти такими способами, як:

– надання банківським працівником у вигляді флаєра або брошури повної інформації щодо видів шахрайств з платіжними картками при отриманні клієнтом банківської картки;

– надсилання на електронну пошту клієнтів листів про нові форми шахрайства електронного банкінгу;

– поширення банками інформації про види шахрайства електронного банкінгу через засоби масової інформації та соціальні мережі;

– інформування банками своїх клієнтів щодо порядку дій у разі втрати чи крадіжки картки з метою своєчасного блокування їх рахунків.

2) модернізація та захист власних програмних комплексів банків відповідно до сучасних вимог безпеки, а також проведення постійного моніторингу ринку електронного банківського обслуговування з метою виявлення нових видів шахрайства.

У зв'язку з цим актуальною є тема цього дослідження, пов'язана з проектуванням та розробкою системи моніторингу шахрайських банківських транзакцій.

Для досягнення мети роботи необхідно вирішити такі завдання:

– Виконати аналіз проблеми шахрайства в банківських транзакціях;

– виконати проектування та розроблення вимог до системи моніторингу шахрайських банківських транзакцій;

– Виконати проектування системи моніторингу шахрайських банківських транзакцій;

– розробити макети системи моніторингу шахрайських банківських транзакцій;

2 ТЕОРЕТИЧНІ ДОСЛІДЖЕННЯ ТА ВИЗНАЧЕННЯ СПЕЦИФІКАЦІЙ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

2.1 Розробка архітектури програмного забезпечення

У роботі пропонується використання клієнт-серверної архітектури, що дозволяє клієнтським компонентам взаємодіяти з ресурсами даних.

Загалом дворівнева архітектура - це моделі програмування, які дозволяють розподіляти функціональність додатків за двома незалежними системами, як правило:

– Клієнтські компоненти, що працюють на клієнтських комп'ютерах (рівень перший);

– Процеси, запущені на віддалених серверах (другий рівень);

– Дискретна колекція баз даних, менеджерів ресурсів та додатків мейнфреймів (другий рівень)

На рисунку 2.1 представлена структурна схема пакета, з розподілом за двома рівнями. Рівні логічні. Вони можуть працювати або не працювати на одному фізичному сервері.

Перший рівень. Відповідає за уявлення та взаємодію з користувачем. Ці клієнтські компоненти дозволяють користувачеві взаємодіяти з процесами другого рівня безпечним та інтуїтивно зрозумілим способом. Клієнти не мають доступу до послуг рівня бази даних та бізнес-логіки безпосередньо. Клієнтський компонент надає діалогову форму, в якій адміністратор вводить дані про клієнтів та їх заявки. Клієнтський компонент відправляє це замовлення процесам другого рівня, які перевіряють бази даних послуг та виконують завдання, необхідні для замовлення послуги.

Рисунок 2.1 – Структурна схема пакета

Другий рівень. Процеси другого рівня зазвичай називають рівнем бізнес-логіки програми. Ці процеси управляють бізнес-логікою програми, і їм дозволено доступом до сервісів бази даних. На рівні логіки програми відбувається більша частина обробки. Декілька клієнтських компонентів можуть одночасно звертатися до процесів другого рівня, тому цей рівень логіки програми повинен керувати власними транзакціями.

Без рівня логіки програми клієнтські компоненти отримують прямий доступ до бази даних. База даних потрібна для управління власними з'єднаннями, зазвичай блокуючи доступ до запису. Поділ рівня бізнес-логіки та рівня даних знижує навантаження на служби бази даних, підтримує більш

ефективне управління з'єднаннями та може підвищити загальну продуктивність мережі.

Служби бази даних захищені від прямого доступу до клієнтських компонентів, що знаходяться в захищеній мережі. Взаємодія має відбуватися через процеси другого рівня.

Спілкування між рівнем. Усі рівні повинні спілкуватися один з одним. Відкриті стандартні протоколи та відкриті API полегшують цей зв'язок. У роботі створюється клієнтська програма з використанням мови програмування C#. Ці клієнти працюють у будь-якій операційній системі, спілкуючись із шаром логіки програми. Бази даних можуть бути будь-якого дизайну, якщо прикладний рівень може вимагати їх і маніпулювати ними. Ключем до цієї архітектури є рівень логіки програми.

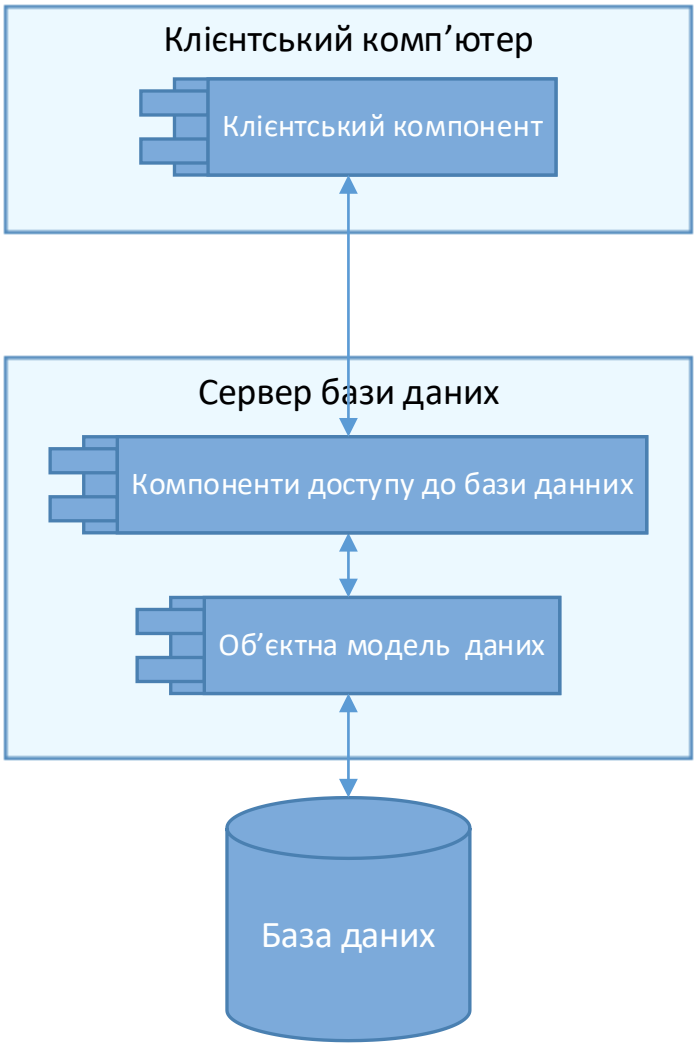

Рисунок 2.2 – Схема пакетів для модуля менеджера

Для модуля Менеджера пропонується використання клієнт-серверної моделі системи баз даних (рис. 2.2).

Сервер у моделі клієнт/сервер – це просто СУБД, а клієнт – це додаток бази даних, що обслуговується СУБД. У роботі ролі СУБД виступає MS SQL Server.

Базова модель клієнт/сервер нічого не говорить про розташування різних компонентів. Проте, оскільки компоненти є різними, їх можна знайти на різних комп'ютерах. Модель розподіленої клієнт-серверної системи, в якій клієнт знаходиться на одному комп'ютері, а сервер та база даних - на іншому, настільки популярна, що її зазвичай називають моделлю клієнт-сервер (рис. 2.3). Модель віддаленої бази даних стосується випадку, коли клієнт і сервер знаходяться на одному комп'ютері, але база даних знаходиться на віддаленому комп'ютері (рис. 2.4).

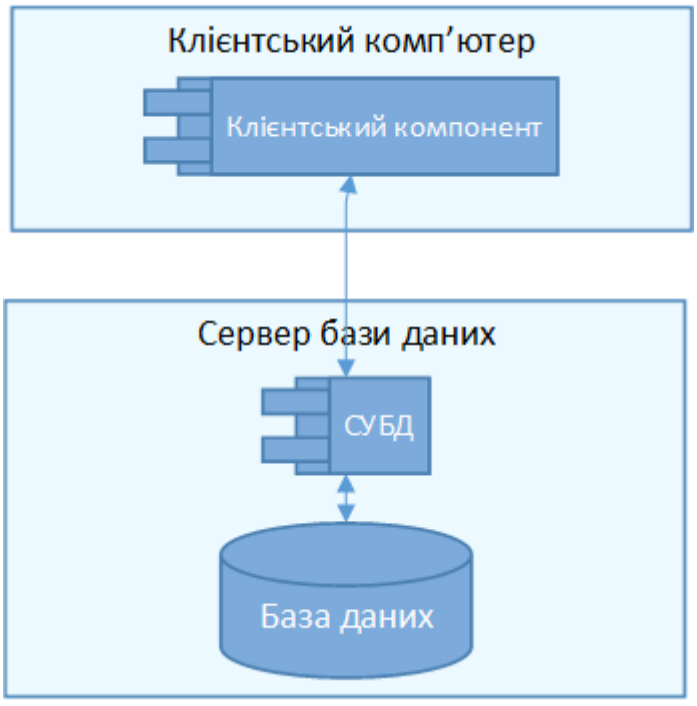

Рисунок 2.3 – Розподілена клієнт-серверна система

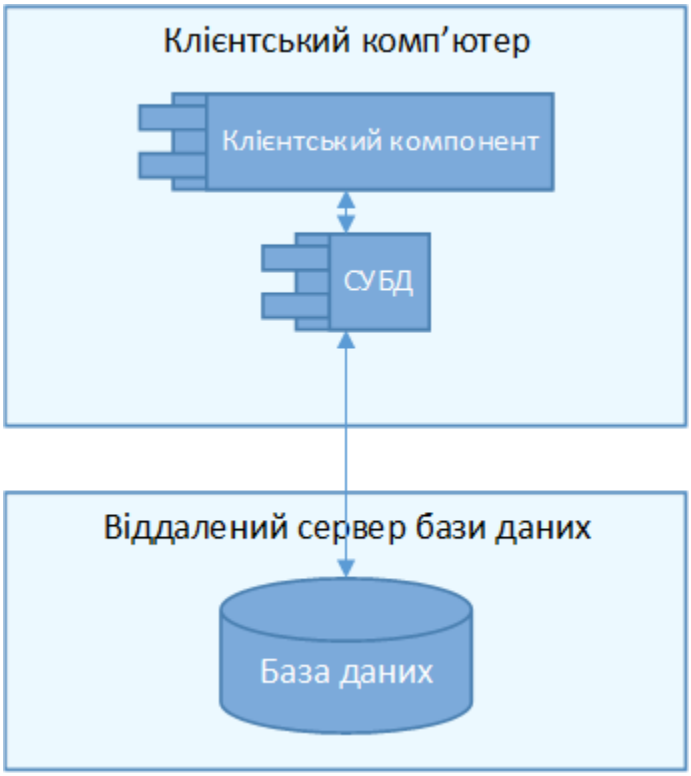

Рисунок 2.4 – Модель віддаленої бази даних

2.2 Розробка інформаційних специфікацій

Дослідження предметної області дозволило виявити такі об'єкти:

– Клієнт;

– Менеджер;

– Послуга;

– Договір;

– Реєстр транзакцій.

З виявлених об'єктів предметної області можна сформувати інформаційну модель (рис.2.5).

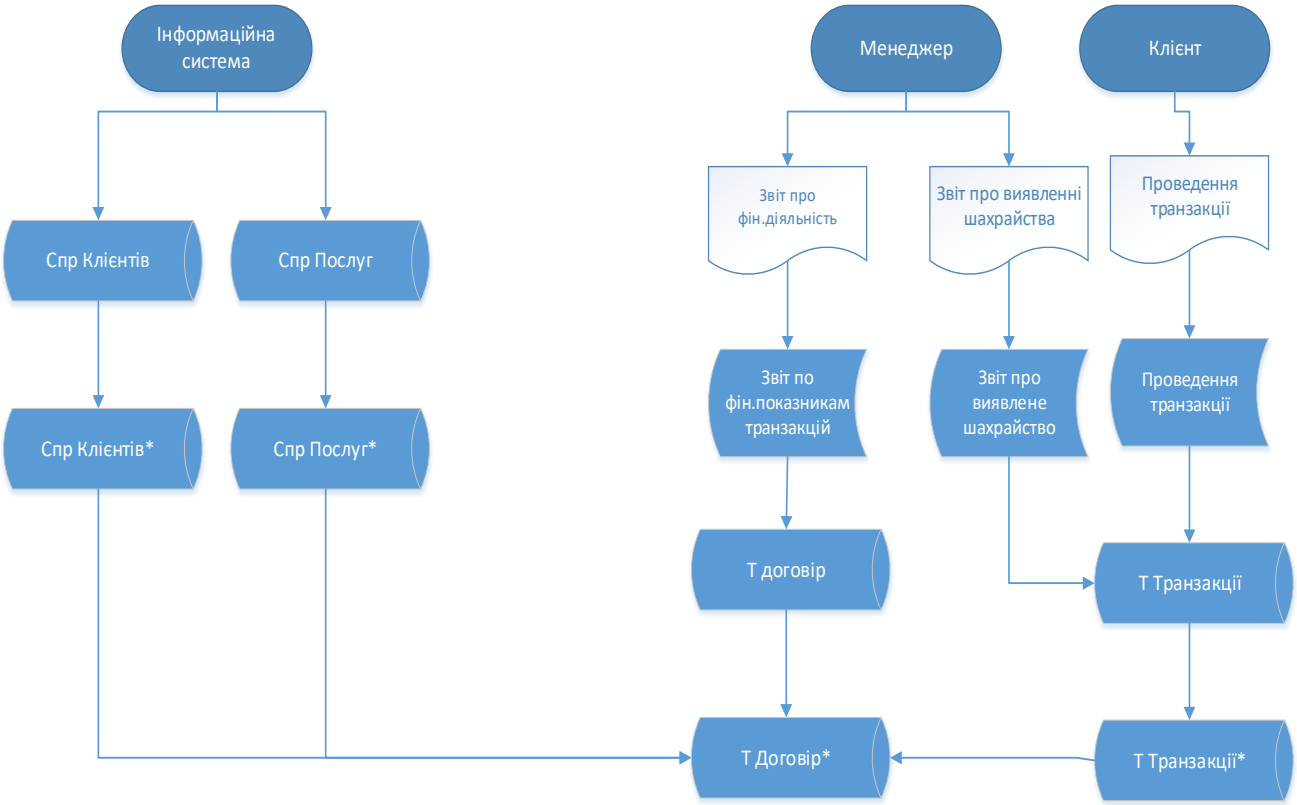

Рисунок 2.5 – Інформаційна модель

Для формування глобальної логічної моделі даних необхідно кожному виділеного об'єкта логічної моделі даних визначити набір атрибутів із зазначенням джерела даних їх заповнення. Опис сутностей та його атрибутів представлено у таблиці 2.1.

Таблиця 2.1 - Характеристика вхідних даних

Дані	Реквізити	Джерело
Відомості про клієнтів	ПІБ дата народження Серія та номер паспорта дата видачі паспорта Ким виданий паспорт Адреса	Паспорт клієнта
	Телефон Електронна пошта	За словами клієнта

Продовження таблиці 2.1

Відомості про договори	Номер договору Дата ув'язнення Клієнт Адміністратор Послуга	Договір
Відомості про відвідування занять клієнтом	Вид послуги дата відвідування Час відвідування Інструктор Коментарі	Реєстр відвідувань
Відомості про транзакції	Клієнт Дата транзакції Опис транзакції Сума	Транзакція клієнта
Відомості про послуги	Послуга Опис послуги Вартість	Транзакція клієнта
Відомості про адміністратора	ПІБ Телефон Електронна пошта	Дані компанії

Проведемо опис глобальної логічної моделі даних. Опис сутностей глобальної логічної моделі даних та атрибутів сутностей представлено у таблиці 2.2.

Таблиця 2.2 - Опис глобальної логічної моделі даних

Сутність	Атрибут	Опис
1	2	3
Клієнт	кодКлієнта	Унікальний ідентифікатор
	ПІБКлієнта	Прізвище, ім'я та по батькові клієнта
	Дата народження	Дата народження клієнта
	Паспорт	Серія та номер паспорта клієнта
	Дата видачі	Дата видачі паспорта клієнта
	Ким виданий	Ким виданий паспорт
	Адреса	Адреса проживання клієнта
	Телефон	Контактний телефон клієнта
	Email	Електронна адреса клієнта
Послуга	кодПослуги	Унікальний ідентифікатор послуги
	Назва Послуги	Назва послуги
	ОписПослуги	Опис послуги
	Вартість послуги	Вартість послуги
Договір	Номер договору	Унікальний ідентифікатор договору
	кодЗаявки	Код заявки
	кодАдміністратора	Код Адміністратора
	ДатаДоговору	Дата оформлення договору
Адміністратор	кодАдміністратора	Унікальний ідентифікатор
	ПІБАдміністратора	ПІБ адміністратора
	ТелАдмін	Телефон адміністратора
	emailАдмін	Електронна адреса адміністратора
Реєстр транзакцій	Код	Унікальний ідентифікатор
	кодКлієнта	Код клієнта
	кодІнструктора	Код Інструктора
	ДатаТранзакції	Дата проведення транзакції клієнтом
	Сума транзакції	Сума транзакції
	Код отримувача	Код отримувача транзакції

Для встановлення зв'язків між сутностями логічної моделі даних необхідно визначити первинні та зовнішні ключі в батьківських та дочірніх сутностях, а також визначити тип та потужності зв'язку між сутностями. Опис ключів у сутності бази даних, а також опис зв'язків між сутностями представлено у таблицях 2.3-2.4.

Таблиця 2.3 - Опис ключів сутностей глобальної логічної моделі даних

Сутність	Ключове поле	Тип ключа
Клієнт	кодКлієнта	Первинний
Послуга	кодПослуги	Первинний
Договір	Номер договору	Первинний
	кодЗаявки	Зовнішній
	кодАдміністратора	Зовнішній
Адміністратор	кодАдміністратора	Первинний
Реєстр транзакцій	Код	Первинний
	кодКлієнта	Зовнішній
	кодОдержувача	Зовнішній

Таблиця 2.4 - Опис зв'язків між сутностями глобальної логічної моделі даних

Батьківська сутність	Дочірня сутність	Тип зв'язку	Потужність зв'язку
Клієнт	Договір	НЕІД	1:М
Послуга	Договір	НЕІД	1:М
Клієнт	Реєстр транзакцій	НЕІД	1:М
Адміністратор	Договір	НЕІД	1:М

2.3 Розробка поведінкових специфікацій

При роботі із системою передбачається три користувача: Адміністратор, Клієнт. Дерево функцій для модуля Адміністратор представлено рисунку 2.6.

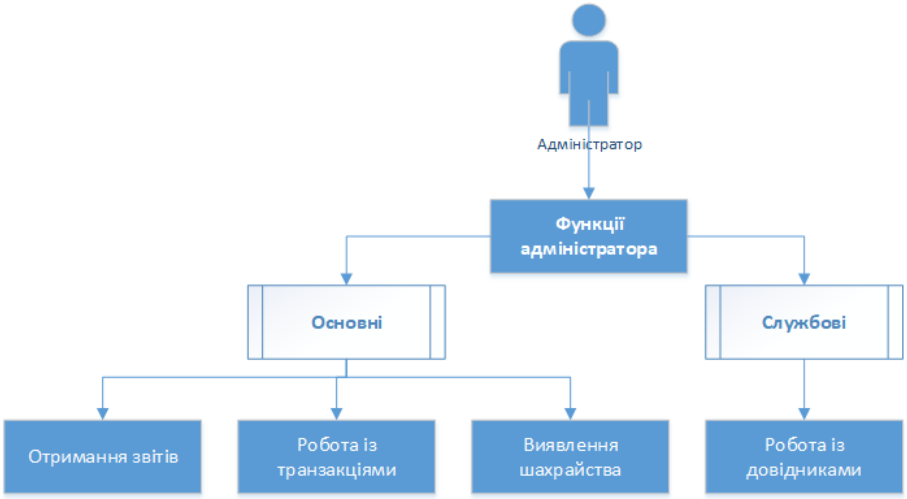

Рисунок 2.6 – Дерево функцій Адміністратора

Дерево функцій для модуля Клієнта представлене рисунку 2.7.

Рисунок 2.7 – Дерево функцій для модуля Інструктора

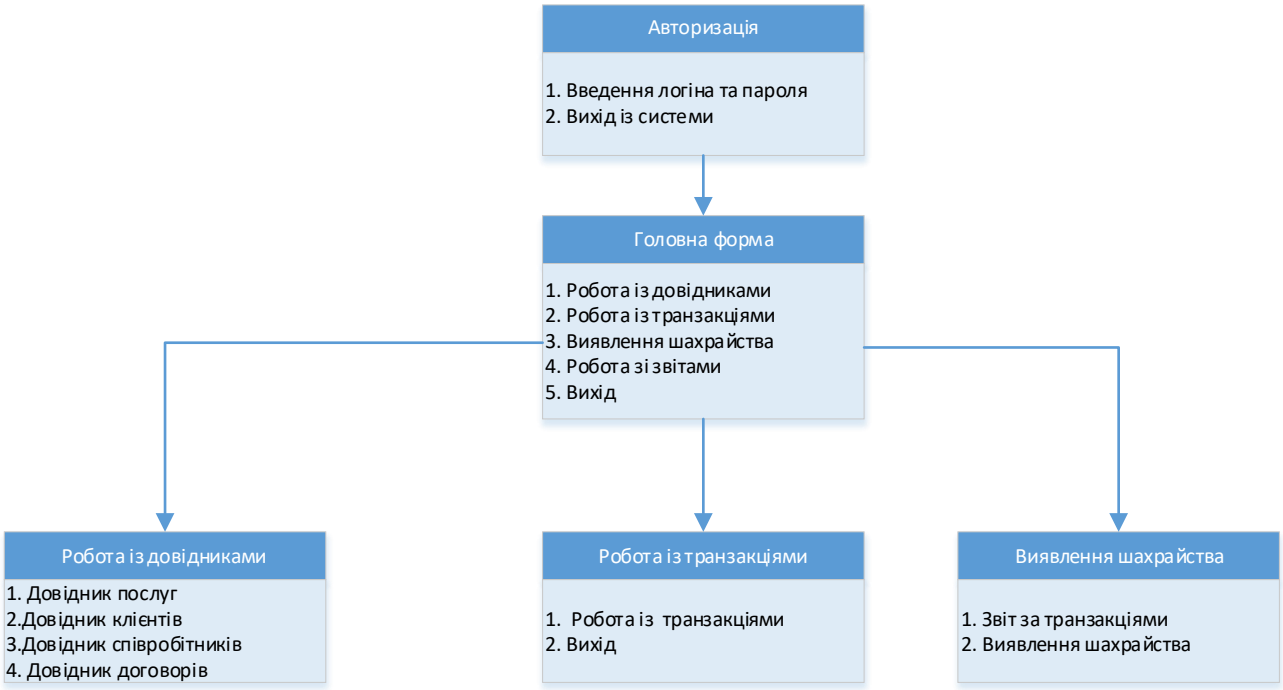

Рисунок 2.8 – Сценарій діалогу Адміністратор

На підставі розроблених дерев функцій для модулів Адміністратора та Клієнта визначимо сценарії діалогу. На рисунку 2.8 представлений сценарій діалогу для Адміністратора. На рисунку 2.9 представлено сценарій діалогу для Клієнта.

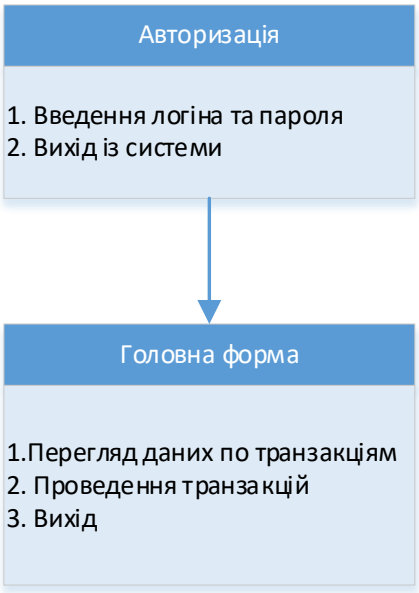

Рисунок 2.9 – Сценарій діалогу Клієнта

Рисунок 2.10 – Схема роботи системи виявлення шахрайських транзакцій

Робочий процес оцінки шахрайських транзакцій містить такі кроки:

1 Клієнтська програма надсилає інформацію програмі з визначення шахрайських транзакцій.

2 Модель виявлення шахрайських транзакцій визначає оцінку ризику (в діапазоні 0-100) для вхідних даних за допомогою моделі машинного навчання, що навчається з використанням історичних даних. Оцінка 0 вказує на те, що прогноз вважається таким, що має найменший можливий ризик, а оцінка 100 вказує на те, що прогноз вважається таким, що має найбільший можливий ризик.

3 Якщо оцінка ризику для конкретного прогнозу опускається нижче заданого граничного значення, подальші дії не здійснюються.

4 Якщо оцінка ризику перевищує заданий поріг (наприклад, 90), цикл запускається автоматично та надсилає прогнози для перевірки персоналом банку. Персонал банку розглядають транзакцію та виносять рішення (схвалюють, відхиляють або відправляють для подальшої перевірки).

5 Результат схвалення або відхилення зберігається у БД. Дані БД можуть використовуватися для перенавчання моделі виявлення шахрайських транзакцій.

2.4 Висновки по розділу 2

У другому розділі була проведена розробка програмного забезпечення для виявлення шахрайства із фінансовими транзакціями по карточним рахункам. Розглянуто архітектурну модель застосування для виявлення шахрайства. У якості архітектури обрано клієнт-серверну модель. Розроблено інформаційну модель для бази даних застосування, визначені основні сутності їх атрибути та описано зв’язки між сутностями. Наведено сценарії діалогів для Адміністратора та Клієнта застосування. Розглянуто схему виявлення транзакцій за допомогою застосування.

3 ПРАКТИЧНА ЧАСТИНА. ПРОЄКТУВАННЯ ТА РОЗРОБКА ПРОГРАМНИХ КОМПОНЕНТІВ ДЛЯ СИСТЕМИ ВИЯВЛЕННЯ ШАХРАЙСТВА

3.1. Проєктування структури програмного забезпечення.

Діаграми розгортання використовуються для візуалізації топології фізичних компонентів системи, де розгортаються програмні компоненти.

Діаграми розгортання застосовуються для опису статичного уявлення розгортання системи. Діаграми розгортання складаються з вузлів та їх взаємозв'язків.

Сам термін "Розгортання" визначає призначення діаграми. Діаграми розгортання використовуються для опису апаратних компонентів, де розгортаються програмні компоненти. Діаграми компонентів та діаграми розгортання тісно пов'язані.

Діаграми компонентів використовуються для опису компонентів, а діаграми розгортання показують як вони розгортаються в апаратному забезпеченні.

UML здебільшого призначений для фокусування на програмних артефактах системи. Однак ці дві діаграми є спеціальними діаграмами, які використовуються для фокусування на програмних та апаратних компонентах.

Більшість діаграм UML використовуються для обробки логічних компонентів, але діаграми розгортання призначені для того, щоб зосередитись на апаратній топології системи. Схеми розгортання застосовуються системними інженерами.

Мета діаграм розгортання може бути описана таким чином:

– Візуалізується апаратною топологією системи.

– Описує апаратні компоненти для розгортання програмних компонентів.

– Описує вузли обробки часу виконання.

Діаграма розгортання представляє уявлення про розгортання системи. Він пов'язані з діаграмою компонентів, оскільки компоненти розгортаються з допомогою діаграм розгортання. Схема розгортання складається із вузлів. Вузли це не що інше, як фізичне обладнання, що використовується для розгортання програми.

Діаграми розгортання корисні системним інженерам. Ефективна схема розгортання дуже важлива, оскільки вона контролює такі параметри:

– Подання.

– Масштабованість.

– Ремонтопридатність.

– Портативність.

Перед побудовою схеми розгортання слід визначити такі артефакти:

– Вузли

– Відносини між вузлами

Нижче наведено приклад діаграми розгортання інформаційної системи виявлення шахрайства у банківських транзакціях (рис.3.1).

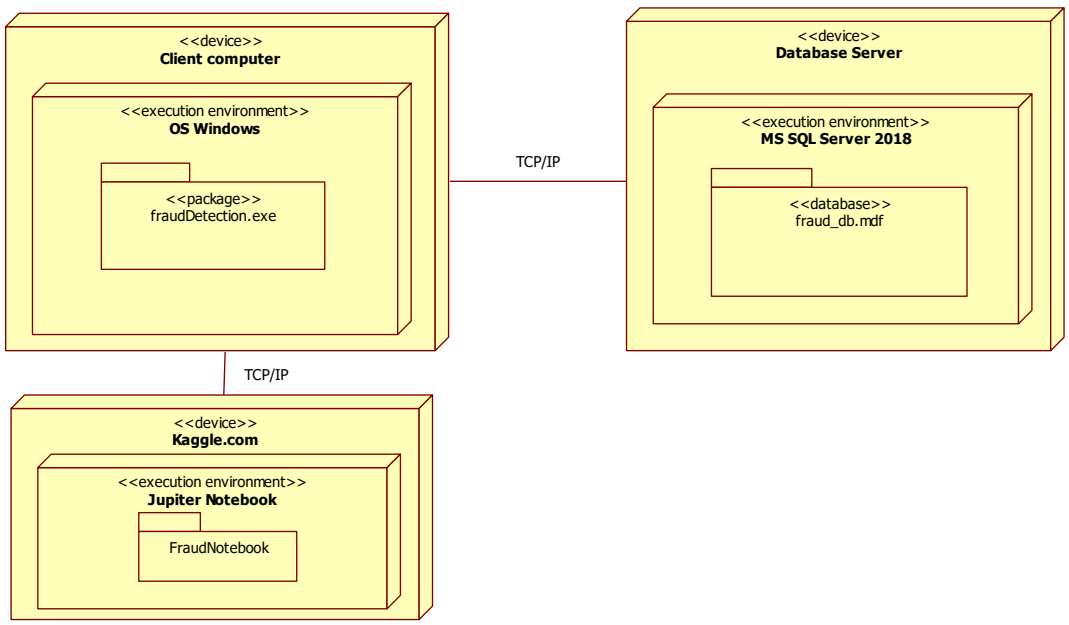

Рисунок 3.1 – Діаграма розгортання програмної системи

Програмна система поєднуватиме клієнт-серверну програму з хмарним сервісом на основі аналітичної платформи Kaggle. Ця програма розгорнута в мережі підприємства з використанням сервера баз даних під управлінням СУБД MS SQL Server 2018. Модуль аналізу та виявлення шахрайства у фінансових транзакціях представлений у вигляді Jupiter Notebook, розгорнутої на хмарній аналітичній платформі Kaggle.

Діаграми розгортання переважно використовуються системними інженерами. Ці діаграми використовуються для опису фізичних компонентів (апаратних засобів), їх розподілу та об'єднання.

Діаграми розгортання можна візуалізувати як апаратні компоненти/вузли, де знаходяться програмні компоненти.

Програмні компоненти розробляються для моделювання складних бізнес-процесів. Ефективних програмних програм недостатньо для задоволення бізнес-вимог. Бізнес-вимоги можна описати як необхідність підтримки кількості користувачів, швидкий час відгуку і т.д.

3.2. Проєктування компонентів програмного забезпечення.

Діаграма класів - це статична діаграма, яку використовують для опису, документування та візуалізації різноманітних фаз системи, також для побудови виконуваного коду програмної програми. Її можна визначити за допомогою обмежень, які накладаються на систему, а також завдяки атрибутів і операцій класів. Такі діаграми показують набори інтерфейсів, класів, взаємодій та обмежень. Вона також відома як структурна схема.

Статистична діаграма дуже часто використовується, щоб змоделювати об’єктивно-орієнтовані системи. Вони є єдиними UML-діаграмами, які можуть бути безпосередньо зіставлені з мовами об’єктивно-орієнтованими системами. Метою діаграми класів є моделювання статичного уявлення програми. Діаграми класів - це єдині діаграми, які можна безпосередньо зіставлені з об’єктивно-орієнованими мовами і тому широко використовуються під час побудови.

UML-діаграми, такі як діаграма активності, діаграма послідовностей, можуть дати лише потік послідовностей програми, проте діаграма класів трохи відрізняється. Це найпопулярніша UML-діаграма у спільноті кодерів.

Мета діаграми класів може бути узагальнена таким чином:

– Аналіз та проектування статичного уявлення програми.

– Опис обов'язків системи.

– База для діаграм компонентів та розгортання.

– Пряма та зворотна інженерія.

Діаграми класів - це найбільш популярні UML-діаграми, що використовуються для створення програмних додатків. Дуже важливо вивчити процедуру малювання діаграми класів.

Діаграми класів мають багато властивостей, які необхідно враховувати при малюванні, але тут діаграма розглядатиметься з погляду верхнього рівня.

Діаграма класів в основному є графічним уявленням статичного уявлення системи і представляє різні аспекти додатку. Набір діаграм класів представляє всю систему.

При побудові діаграми класів слід пам'ятати наступні моменти:

– Ім'я діаграми класів має сенс для опису аспекту системи.

– Кожен елемент та його взаємозв'язку мають бути визначені заздалегідь.

– Відповідальність (атрибути та методи) кожного класу має бути чітко визначена

– Для кожного класу має бути зазначено мінімальну кількість властивостей, оскільки непотрібні властивості ускладнять схему.

На рисунку 3.2 представлена діаграма класів для системи обліку транзакцій та виявлення шахрайства у фінансових транзакціях.

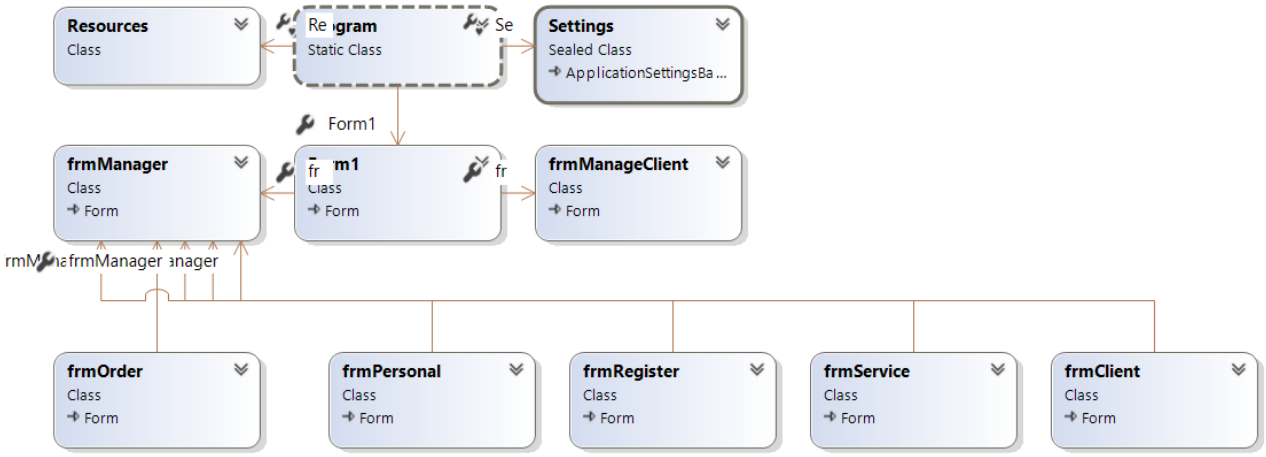

Рисунок 3.2 – Діаграма класів програмного забезпечення

Опис елементів діаграми класів представлено таблиці 3.1.

Таблиця 3.1 - Опис елементів діаграми класів

Елемент	Опис
Program	Точка входу в систему
Resources	Ресурси програми
Settings	Установки програми
Form1	Головна форма авторизації та реєстрації користувачів
frmManager	Головна форма для роботи менеджера
frmManageClient	Форма для роботи клієнта із системою
frmOrder	Форма для роботи з таблицею Договору
frmPersonal	Форма для роботи з таблицею Адміністратор
frmRegister	Форма для роботи з таблицею Реєстр транзакцій
frmService	Форма для роботи з таблицею
frmClient	Форма для роботи з таблицею Клієнт

3.4. Проєктування алгоритмів

Виявлення фінансових махінацій – складне завдання. Існує велика кількість робіт із створення моделі розпізнавання фінансових махінацій [29-31]. У роботі Дж. Хансена [32] було запропоновано узагальнену якісну модель реагування (EGB2) для виявлення управлінського шахрайства. У роботі

використовувалися дані міжнародної бухгалтерської фірми, щоб оцінити модель, і вона показала відмінну здатність до прогнозування. Автори виявили, що модель може вирішити асиметричні помилки вартості від помилок типу II та I, вона може ефективно запобігти втраті від судового позову від скоєння помилки типу II. Є. Кіркос та інші автори [33] використовували метод інтелектуального аналізу даних для створення трьох шахрайських ідентифікаційних моделей на основі Дерев рішень (DT), Нейронних мереж (NN) та Байєсівської мережі (BBN), а також шляхом десятикратної перехресної перевірки. Результати показують, що байєсовська мережа (BBN) має найвищу точність. Ravisankar та інші в роботі [34] використовували багатошарові нейронні мережі зворотного зв'язку, машини опорних векторів, імовірнісні нейронні мережі та інші методи побудови моделей. Результати показали, що продуктивність ймовірнісної нейронної мережі (PNN) була найвищою. Фаннінг і Коггер [35] побудували модель на основі нейронної мережі з незалежними змінними, що включали фінансові коефіцієнти та якісні змінні. Вони також використовували деякі інші традиційні статистичні методи для побудови моделей, щоб порівняти з нейронні мережі. Отже вияснилось що найбільшу високу точність мають нейронні мережі, ніж традиційні статистичні методи. Результати показують, що байєсовська мережа (BBN) має найвищу точність. Ravisankar та інші в роботі [34] використовували багатошарові нейронні мережі зворотного зв'язку, машини опорних векторів, імовірнісні нейронні мережі та інші методи побудови моделей.

Саммерс та Суїні [36] створили логістичну модель для перевірки взаємозв'язку між внутрішніми транзакціями та шахрайством. Вони виявили, що менеджери зменшуватимуть запаси акцій компанії шляхом частих торгів, коли відбувається шахрайство.

Еббот і Паркер [37] досліджували ефективність присутності незалежного комітету з аудиту, який знижує ймовірність корпоративного шахрайства шляхом регресійного аналізу. Результати аналізу показують, коли незалежний аудит комітету та корпоративні річні збори, що проводяться не менше двох разів, повідомляли компанії про зниження помилок у фінансовій звітності. Чен та інші [38] вивчали взаємозв'язок між корпоративним управлінням та структурою

власності корпоративного фінансового шахрайства на вибірці китайських лістингових компаній. Результати показують, що структура корпоративної власності та характеристики ради директорів є важливими показниками інтерпретації шахрайства. Вони також встановили, що співвідношення кількості незалежних директорів,

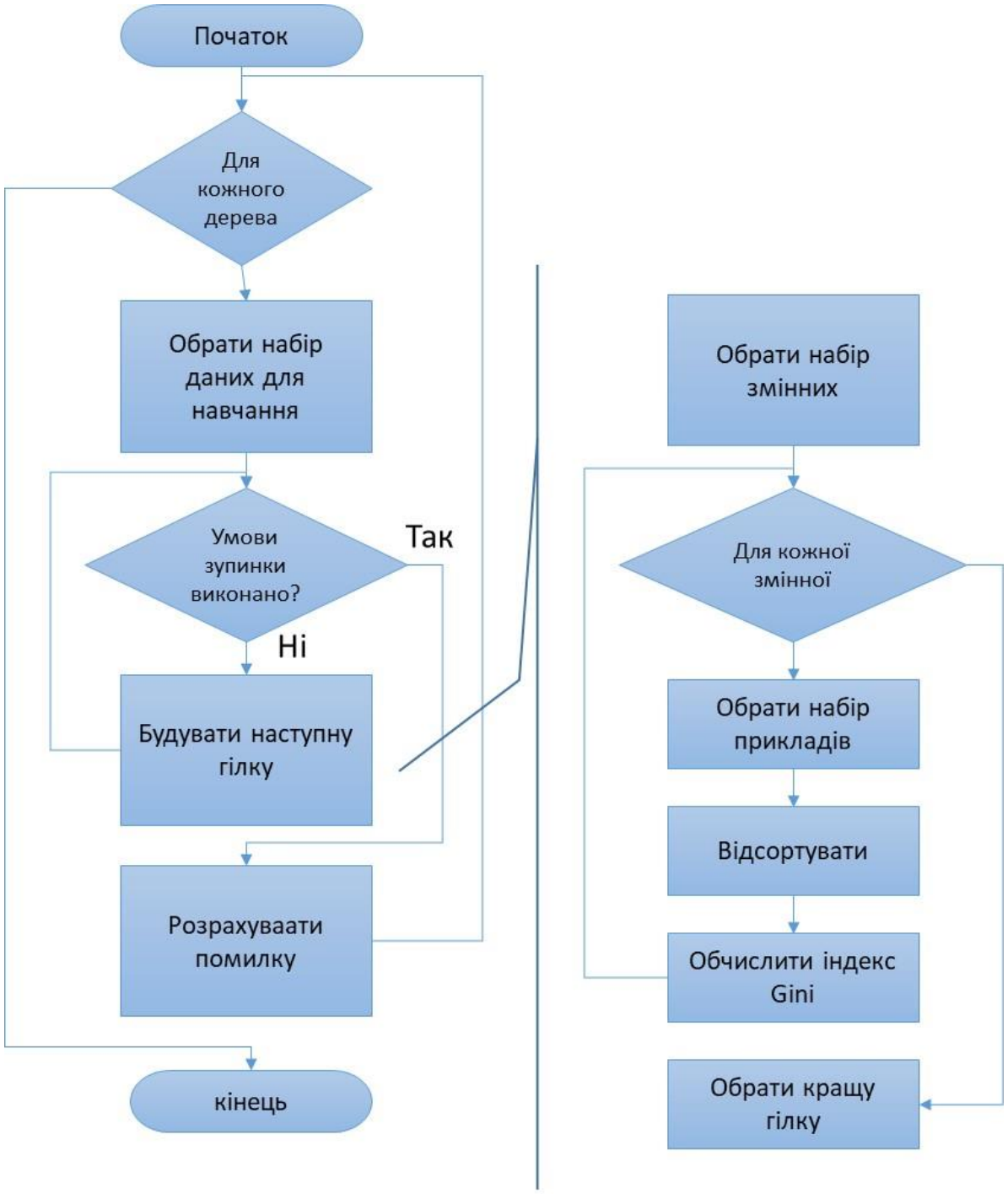

Рисунок 3.3 – Схема алгоритму Random Forest

У цій роботі пропонується використання алгоритму Random Forrest (рис.3.3) виявлення фінансового шахрайства в транзакціях клієнтів банку. Як вихідні дані будуть використані дані про фінансові транзакції клієнтів банку.

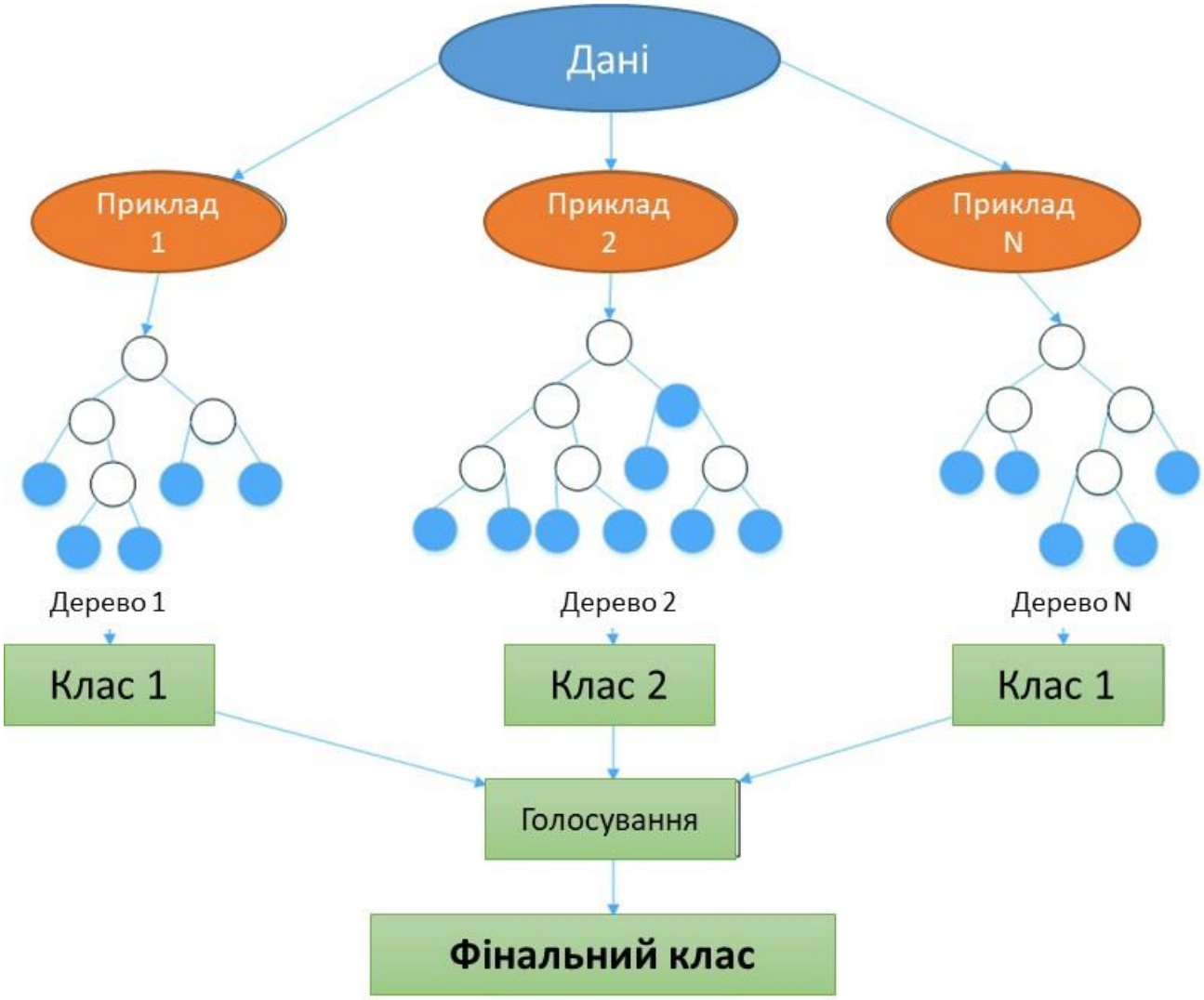

Рисунок 3.5 – Схема роботи алгоритму Random Forrest

Random Forrest - це комбінований метод класифікації, запропонований відомими вченими [39] та включає метод дерева рішень (DT) [40] є основним класифікатором та встановлює велику кількість дерев. Конкретний процес алгоритму полягає в наступному:

Крок 1: Вибірка “m” (кількість обраних вибірок) (m < M, де M-число для всієї вибірки) вибірки випадково з усіх вибірок за допомогою методу bootstrap (рис.3.5).

Крок 2: Формування дерева рішень із вилученим зразком, який не обрізається.

Крок 3: Повтор кроків 1, 2 та побудова великої кількості дерев рішень та розробка класифікаційної послідовності дерева рішень {h1(X), h2(X), ..., hntree(X)}.

Крок 4: Остаточна класифікація визначається за кожним рекордним голосуванням за результатами класифікації дерева рішень.

Його можна виразити так: hi - це єдине дерево рішень модельних дерев, Y- вихідна змінна (або цільова змінна), а I () - індикаторна функція.

3.5 Розробка макетів програмних компонентів

Програмний компонент виявлення шахрайства у банківських транзакціях розроблено на платформі Kaggle, яка надає можливості розробки програмних компонентів для аналізу даних та використання алгоритмів машинного навчання. Як ядро розробки використовується Jupyter Notebook.

Jupyter Notebook – це дуже потужний інструмент для представлення розробки та представлення проектів у галузі наук даних.

Проект Jupyter є наступником більш раннього проекту IPython Notebook, який вперше був опублікований як прототип у 2010 році. Хоча в Jupyter Notebooks можна використовувати з багатьма різними мовами програмування, у цій роботі використовується Python, оскільки він є найпоширенішим варіантом використання.

Для початку створення програмного компонента для виявлення шахрайства у фінансових транзакціях необхідно встановити бібліотеки для дослідницького аналізу та машинного навчання (рис.3.6).

[1]:
```
import pandas as pd
import numpy as np
```

[2]:
```
import warnings
warnings.filterwarnings("ignore", category=DeprecationWarning)
```

[3]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

Рисунок 3.6 – Встановлення необхідних бібліотек

Для початку роботи необхідно завантажити дані для аналізу та побудови моделі машинного навчання. Програмний код для завантаження даних та виведення статистичної інформації про дані представлено на рисунку 3.7. Результати виконання програмного коду представлені рисунках 3.8 – 3.9.

[4]:
```
data = pd.read_csv('../input/PS_20174392719_1491204439457_log.csv')
```

[5]:
```
print('Data does not have any NULL value.')
data.isnull().any()
```

Рисунок 3.7 – Програмний код для завантаження вихідних даних

```
Data does not have any NULL value.
```

Out[5]:
```
step              False
type              False
amount            False
nameOrig          False
oldbalanceOrg     False
newbalanceOrig    False
nameDest          False
oldbalanceDest    False
newbalanceDest    False
isFraud           False
isFlaggedFraud    False
dtype: bool
```

Рисунок 3.8 – Результат виконання програмного коду

Out[6]:

	step	type	amount	nameOrig	oldbalanceOrg	newbalanceOrig	nameDest	oldbalanceDest	newbalanceDest	isFraud	isFlaggedFraud
0	1	PAYMENT	9839.64	C1231006815	170136.0	160296.36	M1979787155	0.0	0.0	0	0
1	1	PAYMENT	1864.28	C1666544295	21249.0	19384.72	M2044282225	0.0	0.0	0	0
2	1	TRANSFER	181.00	C1305486145	181.0	0.00	C553264065	0.0	0.0	1	0
3	1	CASH_OUT	181.00	C840083671	181.0	0.00	C38997010	21182.0	0.0	1	0
4	1	PAYMENT	11668.14	C2048537720	41554.0	29885.86	M1230701703	0.0	0.0	0	0

Рисунок 3.9 – Фрагмент файлу з вихідними даними

Надані дані містять дані про фінансові транзакції, а також цільову змінну isFraud, яка є фактичним статусом шахрайства транзакції, а isFlaggedFraud це індикатор, який використовується для позначення транзакції з використанням деякого порогового значення.

Одним із основних етапів при підготовці моделі машинного навчання є очищення даних. Для очищення даних слід виконати наступний програмний код (рис.3.10).

```
import numpy as np
import pickle
```

[27]:
```
import warnings
warnings.filterwarnings("ignore", category=DeprecationWarning)
```

[28]:
```
data_fraud = pd.read_csv('../input/PS_20174392719_1491204439457_log.csv')
```

+ Code + Markdown

[29]:
```
data_fraud = data_fraud.replace(to_replace={'PAYMENT':1,'TRANSFER':2,'CASH_OUT':3,
                                            'CASH_IN':4,'DEBIT':5,'No':0,'Yes':1})
```

[30]:
```
data_fraud.drop(['nameOrig','nameDest','isFlaggedFraud'],axis=1,inplace=True)
```

[31]:
```
data_fraud.head()
```

Рисунок 3.10 – Програмний код для очищення та підготовки вихідних даних

У процесі підготовки даних відбувається заміна текстових значень поля тип транзакції на числові значення. Також видаляються поля, які не є

інформативними та необхідними для побудови моделі машинного навчання: nameOrig, nameDest, isFlaggedFraud. Результати очищення даних представлені рисунку 3.11.

	step	type	amount	oldbalanceOrg	newbalanceOrig	oldbalanceDest	newbalanceDest	isFraud
0	1	1	9839.64	170136.0	160296.36	0.0	0.0	0
1	1	1	1864.28	21249.0	19384.72	0.0	0.0	0
2	1	2	181.00	181.0	0.00	0.0	0.0	1
3	1	3	181.00	181.0	0.00	21182.0	0.0	1
4	1	1	11668.14	41554.0	29885.86	0.0	0.0	0

Рисунок 3.11 – Результати очищення та підготовки даних

Для створення моделі машинного навчання пропонується використання бібліотеки Scikit-Learn (рис.3.12).

```
[33]: from sklearn.model_selection import train_test_split
      train_X, test_X, train_y, test_y = train_test_split(X, y, test_size = 0.2, random_state = 121)
```

Рисунок 3.12 – Використання бібліотеки Scikit-Learn

Бібліотека Scikit-Learn розроблена на базі SciPy, яка спочатку установлюється, а потім починається робота.

Як алгоритм машинного навчання для класифікації шахрайських транзакцій пропонується використання Random Forrest.

Random Forrest - це тип контрольованого алгоритму машинного навчання, що базується на ансамблевому навчанні. Ансамблеве навчання - це тип навчання, при якому поєднуються різні типи алгоритмів або один і той же алгоритм запускається кілька разів, щоб сформувати більш потужну модель прогнозування. Алгоритм Random Forrest поєднує кілька алгоритмів одного і того ж типу, тобто кілька дерев рішень, у результаті чого утворюється ліс дерев, звідси і назва "Випадковий ліс". Алгоритм випадкового лісу можна використовувати як регресійних, так класифікаційних завдань.

Нижче наведено основні кроки, пов'язані з виконанням алгоритму Random Forrest:

1. Виберіть N випадкових записів із набору даних.

2. Побудуйте дерево рішень на основі цих записів N.
3. Виберіть потрібну кількість дерев у вашому алгоритмі та повторіть кроки 1 та 2.

Перевагами використання випадкового лісу, є:

- Стабільність. Як і в будь-якому алгоритмі, у його використанні є свої недоліки та переваги. Коли в набір даних вводяться нові точки даних, то це не призводить до сильної зміни загального алгоритму. Внесені зміни можуть змінити тільки одне дерево, а не усі.
- Числові та категоріальні ознаки.
- Маштабовані або пропущенні значення в даних (хоча ми виконали масштабування об'єктів у цій статті тільки для демонстрації).

Недоліки використання випадкового лісу:

- Складність – це є основним недоліком випадкових лісів. Велика кількість дерев рішень і їх з’єднання вимагають багато обчислювальних ресурсів.

Через свою складність вони вимагають набагато більше часу для навчання, ніж інші порівнянні алгоритми.

Далі розглянемо як бібліотека Scikit-Learn Python може бути використана для реалізації алгоритму випадкового лісу для вирішення задач класифікації шахрайських фінансових транзакцій.

3.6 Проведення експериментальних досліджень

Оскільки немає «відсутніх» і «сміттєвих значень», немає необхідності додаткової очистці даних, але все одно потрібно виконати аналіз даних, оскільки дані містять величезні варіації значення у різних стовпцях. Нормалізація підвищить загальну точність моделі машинного навчання.

Проведемо аналіз сум переказів за типами транзакцій. Результат аналізу представимо у вигляді стовпчикової діаграми (рис.3.13).

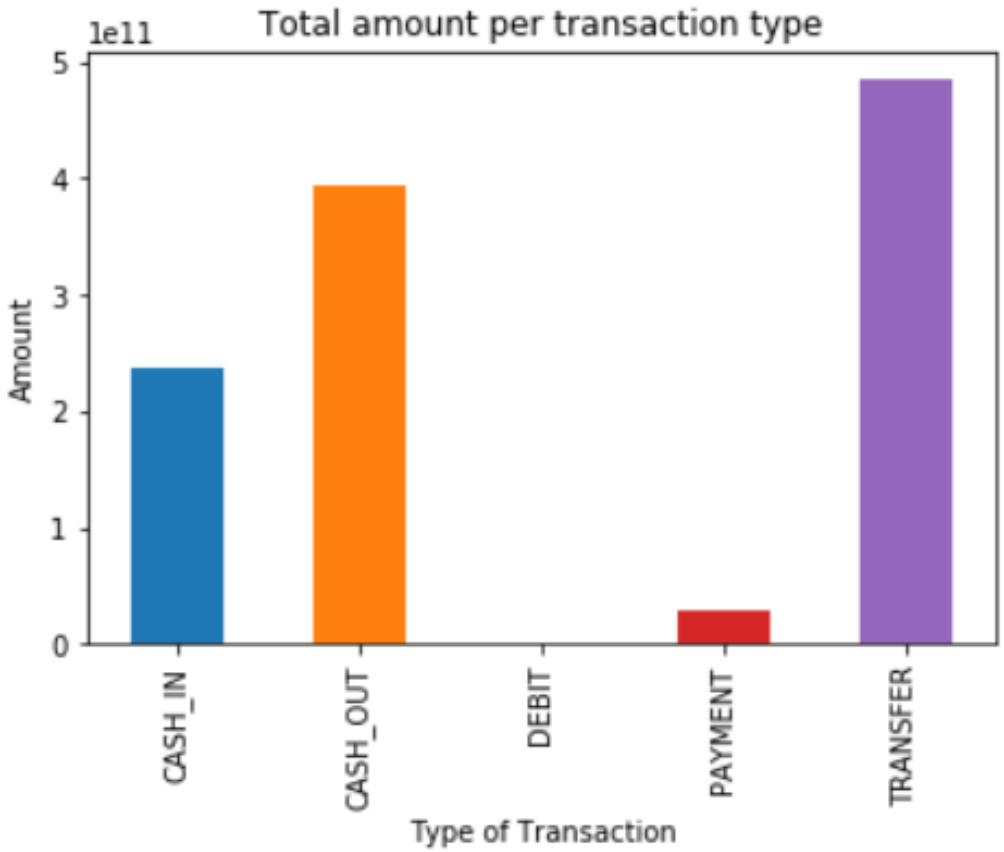

Рисунок 3.13 – Аналіз транзакцій за типом

Як видно з графіка, найбільший обсяг транзакцій виконується у вигляді переказів (TRANSFER), на другому місці транзакції, пов'язані зі зняттям готівки з рахунку (CASH_OUT) та на третьому місці внесення грошей на рахунок (CASH_IN).

Графік вище показує, що TRANSFER та CASH_OUT також є єдиним способом шахрайства. Таким чином, ми зосередимося на цьому виді угод.

Побудуємо теплову картку (heatmap) (рис. 3.14). Теплова карта — це матричне представлення даних, у якому відображається кожне значення за допомогою певного кольору. Кожній величині відповідає свій колір, а матриця індексів зіставляє 2 елементи або їх характеристики. Теплові карти показують зв'язки кількох змінних між собою, відображаючи їх величини у вигляді певних кольорів. Також, подивившись інші точки на теплової карті, можна побачити, як у наборі даних один зв'язок порівнюється з усіма іншими. Саме кольори

дозволяють легко та швидко проаналізувати дані, оскільки таке позначення є інтуїтивно-зрозумілим.

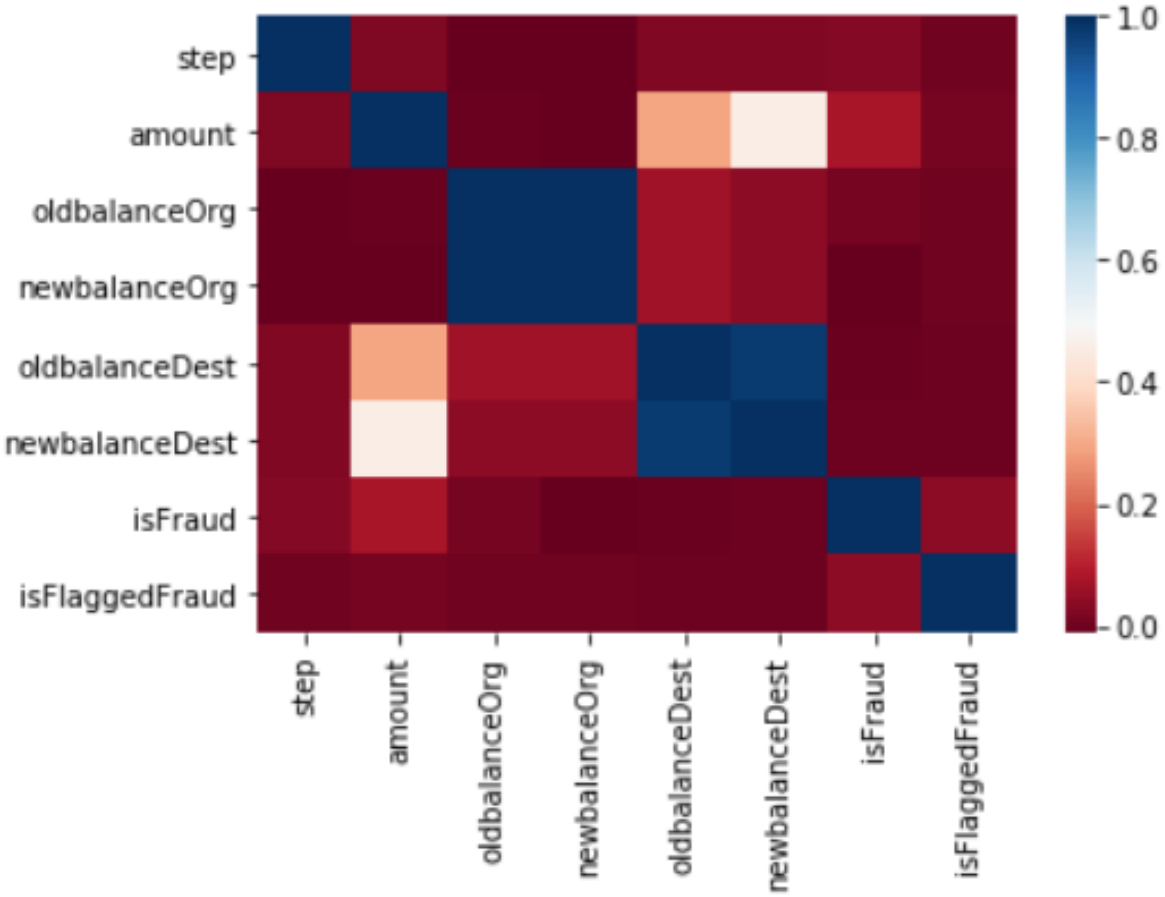

Рисунок 3.14 – Теплова картка

Аналіз теплової карти, представленої на рисунку 3.14, показав, що:

- значення полів OldbalanceOrg та NewbalanceOrg сильно корелюють.
- Значення полів OldbalanceDest та NewbalanceDest сильно корелюють.
- значення поля Amouny корелює з isFraud (цільовою змінною).

Між цими ознаками немає великого зв'язку, тому треба зрозуміти, де зв'язок між ними залежить від типу транзакції та суми. Для цього потрібно по-різному представити теплову карту шахрайських та не шахрайських транзакцій (рис.3.15).

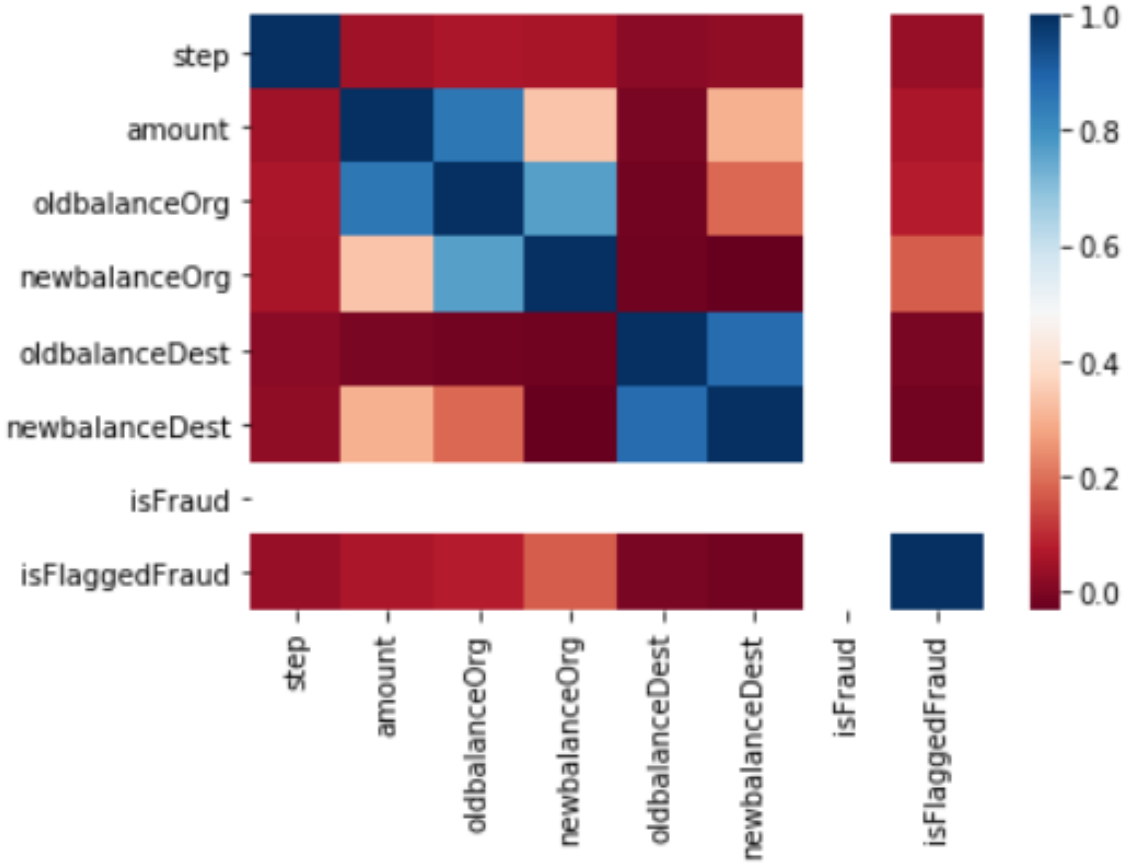

Рисунок 3.15 – Теплова картка

Як видно з рисунку 3.15, є 2 поля, які несуть корисну інформацію для моделі виявлення шахрайства: isFraud та isFlaggedFraud column. Виходячи з гіпотези, isFraud - це індикатор, який вказує на "фактичні шахрайські транзакції", тоді як isFlaggedFraud - це те, що система запобігає транзакції через спрацьовування "деяких порогових значень". З наведеної вище теплової карти видно, що існує певний зв'язок між іншими стовпцями і isFlaggedFraud, отже, має бути зв'язок між isFraud.

Побудувавши діаграму виявлених шахрайських (рис.3.16) транзакцій, можна побачити, що виявляється системою пороговим значенням лише 0,2% таких транзакцій.

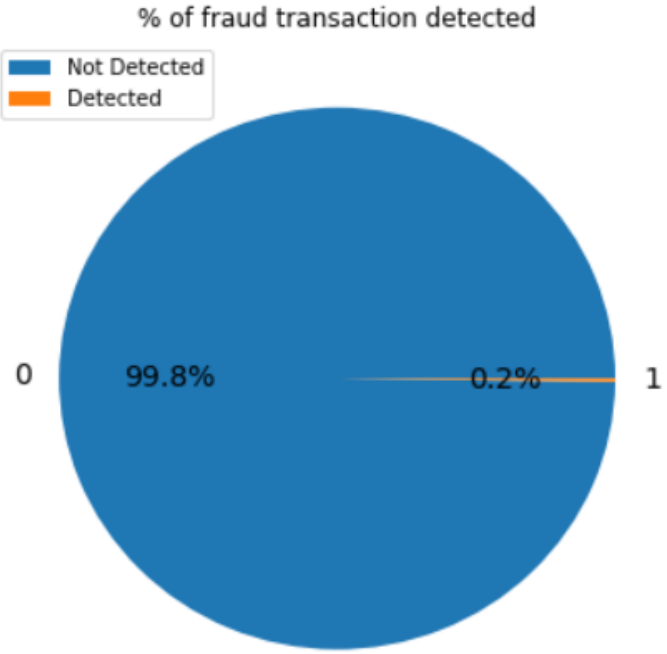

Рисунок 3.16 – Відсоток виявлених шахрайських транзакцій

Також побудуємо графік виявлених шахрайських транзакцій, що були відзначені як шахрайські (рис.3.17).

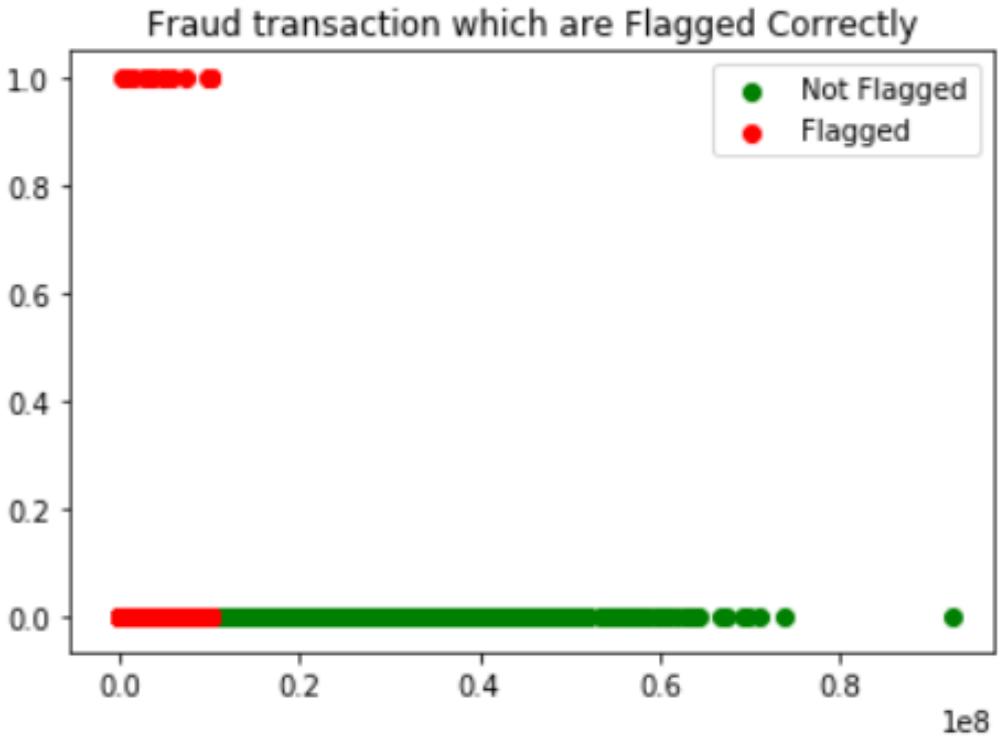

Рисунок 3.17 – Графік зазначених шахрайських транзакцій

Наведені вище графіки ясно показують необхідність програмного модуля, який швидко і надійно виявляв транзакції, які є шахрайством.

У запропонованому, у роботі, програмному модулі пропонується використання алгоритму Random Forrest для класифікації набору даних з фінансових транзакцій.

Random Forrest – це алгоритм класифікації та регресії. Загалом це набір класифікаторів дерева рішень. Випадковий ліс має перевагу перед деревом рішень, оскільки він виправляє звичку надмірно пристосовуватися до їхнього тренувального набору.

Підмножина навчального набору вибирається випадковим чином для навчання кожного окремого дерева, а потім будується дерево рішень, кожен вузол потім розбивається на об'єкт, вибраний з випадкового підмножини повного набору об'єктів. Навіть для великих наборів даних з багатьма функціями та екземпляри даних навчаються надзвичайно швидко у випадковому лісі і тому, що кожне дерево навчається незалежно від інших. Було виявлено, що алгоритм випадкового лісу забезпечує хорошу оцінку помилки узагальнення та стійкий до надмірного підганяння.

Особливості запропонованого програмного модуля:

– Випадковий ліс ранжує важливість змінних у регресійному чи класифікаційному завданні природним чином, що може бути зроблено Випадковим лісом.
– Поле "amount" – це сума транзакції. Ознака "isFraud" є цільовим класом для бінарної класифікації, і він набуває значення 1 для позитивного випадку (шахрайство) та 0 для негативного випадку (не шахрайство).

Random Forrest - це тип контрольованого алгоритму машинного навчання, що базується на ансамблевому навчанні. Ансамблеве навчання - це тип навчання, при якому поєднуються різні типи алгоритмів або один і той же алгоритм кілька разів, щоб сформувати більш потужну модель прогнозування. Алгоритм випадкового лісу поєднує кілька алгоритмів одного й того ж типу, тобто кілька дерев рішень, у результаті чого виходить ліс дерев, звідси і назва

"Випадковий ліс". Алгоритм випадкового лісу можна використовувати як регресійних, так класифікаційних завдань.

Нижче наведено основні кроки, пов'язані з виконанням алгоритму випадкового лісу

1. Виберіть N випадкових записів із набору даних.

2. Побудуйте дерево рішень на основі цих записів.

3. Виберіть потрібну кількість дерев в алгоритмі та повторіть кроки 1 та 2.

4. Для завдання класифікації кожне дерево в лісі передбачає категорію, до якої належить новий запис. Зрештою, новий рекорд присвоюється тій категорії, яка отримує більшість голосів.

Так як перші три етапи вже виконані, перейдемо до етапу побудови моделі машинного навчання.

Існує один заключний етап підготовки даних: поділ даних на навчальні та тестові набори. Під час навчання ми дозволяємо моделі "бачити" відповідь. Ми очікуємо, що існує певний зв'язок між усіма ознаками та цільовим значенням, і завдання моделі-вивчити цей зв'язок під час навчання. Потім, коли приходить час оцінити модель, ми просимо її зробити прогнози на тестовому наборі, де вона має доступ лише до функцій (а не відповідей)! Оскільки у нас є фактичні відповіді для набору тестів, ми можемо порівняти ці передбачення з справжнім значенням, щоб судити, наскільки точна модель. Як правило, під час навчання моделі ми випадково розбиваємо дані на навчальні та тестові набори, щоб отримати представлення всіх точок даних.

Наступний код розбиває набори даних на тестові та навчальні.

```
from sklearn.model_selection import train_test_split
train_X, test_X, train_y, test_y = train_test_split(X, y, test_size = 0.2, random_state = 121)
```

Рисунок 3.18 – Програмний код формування навчальної та тестової вибірки

Далі скориставшись бібліотекою Scikit-learn та алгоритмом випадкового лісу зі skicit-learn, створюємо екземпляр моделі та навчаємо модель на навчальних даних (рисунок 3.19).

```
from sklearn.ensemble import RandomForestClassifier
clf = RandomForestClassifier(n_estimators=15)
```

Рисунок 3.19 – Програмний код для навчання моделі

Після навчання моделі, можна дізнатися відносини між функціями та цілями. Наступний крок оцінити якість моделі. Для цього необхідно виконати визначення шахрайських транзакцій за тестовими характеристиками (моделі ніколи не дозволяється бачити тестові відповіді). Потім необхідно порівняти отримані моделлю відповіді відомими відповідями. (Рисунок 3.20).

```
if True:
    probabilities = clf.fit(train_X, train_y.values.ravel()).predict(test_X)
```

+ Code + Markdown

```
from sklearn.metrics import average_precision_score
if True:
    print(average_precision_score(test_y,probabilities))
```

Рисунок 3.20 – Оцінка точності класифікації шахрайських транзакцій

Проведена оцінка показала, що запропонована у роботі модель вміє визначати шахрайські транзакції з точністю 77% (рис.3.21).

```
from sklearn.metrics import average_precision_score
if True:
    print(average_precision_score(test_y,probabilities))
```

```
0.7718845638217104
```

Average precision score is 0.7687.

Рисунок 3.21 – Оцінка точності виявлення шахрайських транзакцій

3.4 Висновки до розділу 3

Таким чином можна зробити висновок, що алгоритм випадкового лісу буде працювати краще з великою кількістю навчальних даних, але швидкість під час тестування та застосування страждатиме. Також допомогло б застосування більшої кількості методів попередньої обробки.

4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ХОРОНИ ПРАЦІ

4.1 Соціальне значення охорони праці

Соціальне значення охорони праці полягає в сприянні зростанню ефективності суспільного виробництва шляхом безперервного вдосконалення і поліпшення умов праці, підвищення його безпеки, зниження виробничого травматизму і захворюваності.

У зв'язку з цим соціальне значення охорони праці проявляється, перш за все, у впливі на зміну наступних трьох основних показників, що характеризують рівень розвитку суспільного виробництва.

1. Зростання продуктивності праці в результаті збільшення фонду робочого часу за рахунок:

– скорочення внутрішньо змінних простоїв шляхом попередження передчасного стомлення, а також зниження числа або ліквідації мікротравм, обумовлених несприятливими умовами праці. Попередження передчасного втоми з допомогою раціоналізації умов праці, введення оптимальних режимів праці та відпочинку та інших заходів на харчових підприємствах сприяє збільшенню ефективного використання робочого часу. Цей же результат дає ліквідація мікротравм, так як кожна з них супроводжується втратою до 2-х годин робочого часу;

– скорочення цілоденних втрат робочого часу в результаті зниження рівня або ліквідації тимчасової непрацездатності через виробничого травматизму, професійної і загальної захворюваності. Цей показник має важливе значення для харчових виробництв, на яких кожна травма в даний час супроводжується втратою працездатності в середньому більш ніж на 26 днів.

2. Збереження трудових ресурсів і підвищення професійної активності працюючих за рахунок:

– поліпшення стану здоров'я працюючих і збільшення середньої тривалості їх життя шляхом поліпшення умов праці, що також супроводжується збільшенням виробничого стажу працюють при їх високої трудової активності;

– підвищення професійного рівня внаслідок зростання кваліфікації і майстерності у зв'язку зі збільшенням виробничого стажу;

– можливості використання залишкової трудової активності, великого практичного досвіду та професійних знань пенсіонерів по старості та інвалідів на доступних для них роботах і забезпеченні відповідних їх фізичним можливостям умов праці.

3. Збільшення сукупного національного продукту за рахунок поліпшення зазначених вище показників і складових їх компонентів.

4.2 Ергономічні проблеми безпеки життєдіяльності

Під ергономікою розуміють галузь знань, яка комплексно вивчає трудову діяльність людини у системах ЛМС з метою забезпечення її ефективності, безпеки та комфорту.

Мета ергономіки — підвищення ефективності системи ЛМС, забезпечення безпеки праці.

Завдання ергономіки:

– розробка основ проектування діяльності людини-оператора з врахуванням специфіки експлуатації технічних систем та факторів навколишнього середовища;

– вивчення закономірностей взаємодії людини з технічними системами та навколишнім середовищем;

– формування принципів побудови системи Л МС та алгоритмів дії у них людини-оператора;

– розробка перспективних форм праці людини і пов'язаних з нею технічних систем, факторів навколишнього середовища;

– розробка методів дослідження, проектування та експлуатації системи Л МС, які забезпечують безпеку людини, ефективність праці.

Предметом ергономіки є трудова діяльність людини у процесі взаємодії з технічними системами та в умовах особливого впливу на неї факторів навколишнього природного середовища.

У системі ЛМС завжди є 3 елементи: предмет праці, засоби праці та суб'єкт праці. Найменшою цільною одиницею, де наявні вказані елементи, є місце праці.

Місце праці — це зона, де є необхідні технічні засоби, де відбувається трудова діяльність людини. Місце праці обладнане засобами відображення інформації, органами керування та допоміжним обладнанням.

Організацією місця праці називається проведення системи заходів шодо його обладнання засобами та предметами праці і їх розташуванням у визначеному порядку з метою досягнення:

– оптимізації умов трудової діяльності;

– безпеки праці;

– максимальної ефективності;

– комфортності роботи людини.

До робочого місця ставляться такі вимоги:

– достатній робочий простір, який дає змогу працюючій людині здійснювати необхідні рухи та переміщення;

– достатні фізичні, зорові та слухові зв'язки між людиною та обладнанням, а також між людьми під час виконання спільного трудового завдання;

– необхідний рівень освітлення;

– наявність необхідних засобів захисту;

– оптимальне розташування робочих місць, а також безпечні та достатні проходи для працюючих людей.

Органи керування повинні забезпечити перехід дій від людини до машини. Вони мають бути надійними у роботі та зручними в користуванні, не допускати аварій, травм при перевантаженнях та помилкових діях людини. При організації робочого місця враховують основні антропометричні дані людини.

Найважливішою характеристикою робо чого місця є зона досягнення моторного поля.

Моторне поле — це простір робочого місця, в якому розміщені органи керування та інші технічні засоби, в якому людина здійснює рухові дії для виконання робочого завдання.

Розрізняють зони легкого та оптимального досягнення.

Легке досягнення — при русі руку плечовому суглобі з опорою Оптимальне досягнення — рух у ліктьових суглобах з опорою.

При організації місця праці потрібно враховувати:

– ступінь рухливості оператора (сидячи, стоячи);

– конфігурацію і спосіб розміщення каналів індикаторів та органів керування;

– потребу в огляді робочого простору;

– необхідність використання робочої поверхні для писання та інших робіт, розміщення телефонів, розташування інструкцій тощо.

Велике значення має правильний вибір робочого сидіння Конструкція робочого сидіння повинна забезпечити підтримку основної робочої пози, не утруднювати робочих рухів, зміну положення, забезпечити умови для відпочинку.

Продуктивність праці, працездатність людини в багатьох випадках визначаються правильним встановленням режиму праці та відпочинку, що означає зміну періодів праці та відпочинку протягом доби, тижня та довшого терміну.

Реалізація основних ергономічних вимог до режимів праці та відпочинку дає змогу забезпечити необхідний рівень працездатності, зменшити втому, зберегти здоров'я людей.

Для операторів, які працюють з екранами дисплеїв та інших індикаторів, можуть бути рекомендовані такі режими праці та відпочинку.

Тривалість безперервної праці не повинна перевищувати 4—6 год. В іншому випадку працездатність через втому зору раптово знижується. Під час праці, яка не допускає відхилення уваги, її тривалість слід скорочувати. Наприклад, оператор, який стежить за екраном індикатора, найуважніше і найточніше працює протягом

перших 30 хв чергування. За цей час він допускає мінімальну кількість помилок (пропусків та хибних тривог). Надалі, внаслідок втоми зорового аналізатора, кількість помилок зростає май же в два рази та залишається незмінною до кінця другої години. Тому для підтримки високої ефективності праці може бути рекомендований 30-хвилинний період чергування з наступною 30-хвилинною перервою.

Для обслуговуючого персоналу, при роботі якого допускаються нерегламентовані перерви і не потрібне постійне перебування на місці праці, тривалість безперервної праці може перевищувати 6 год.

Тривалість відпочинку повинна бути у 2 рази (а при інтенсивному навантаженні — у 3 рази) більшою, ніж тривалість безперервної роботи.

Максимальний інтервал між періодами праці не повинен перевищувати 48 год, тому що більша тривалість відпочинку призводить до значного збільшення часу спрацьованості (у 4—10 разів).

Організація відпочинку має дві мети:

– зняти втому, яка виникла внаслідок попередньої праці;

– забезпечити швидке включення у роботу відпочиваючої зміни (збереження трудової готовності).

При організації праці протягом тижня, місяця потрібно враховувати ту обставину, що з часом організм людини пристосовується до нічної праці і часто злам складеного стереотипу негативно впливає на його працездатність. Разом з тим тривала праця в нічну зміну порушує соціальні та інші зв'язки, що викликає негативну психологічну реакцію. Тому доцільніше чергувати роботу у денну та нічну зміни.

ВИСНОВКИ

У міру того, як інтернет-банкінг стає все більш поширеним, шахрайство стає все більш важливою проблемою. Сучасні методи боротьби з шахрайством включають чорні списки, правила ручної обробки та перевірку окремих транзакцій експертами-людьми. Ці методи досягли хороших результатів, але значна кількість шахрайських транзакцій все ще відбувається, і процес виявлення шахрайства є дорогим через його велику залежність від експертів.

Мета роботи – підвищити якість виявлення шахрайства у фінансових транзакціях та зменшити необхідність дорогого людського аналізу окремих транзакцій. У роботі пропонується проектування та розробка інформаційної системи обліку транзакцій та програмного модуля для виявлення шахрайських транзакцій на основі класифікатора Random Forests.

Для досягнення мети в даній роботі були сформовані та вирішені наступні завдання:

- виконано аналіз проблеми шахрайства у банківських транзакціях;
- виконано проектування та розроблення вимог до системи моніторингу шахрайських банківських транзакцій;
- виконано проектування системи моніторингу шахрайських банківських транзакцій;
- розроблено макети системи моніторингу шахрайських банківських транзакцій.

СПИСОК ЛІТЕРАТУРНИХ ДЖЕРЕЛ

1. Петрик Ю. М. ЕЛЕКТРОННИЙ БАНКІНГ В СУЧАСНИХ УМОВАХ //Збірник матеріалів Звітної студентської наукової конференції за результатами науково-дослідної роботи у 2019 р./за заг. ред. ЯС Янишина, ГВ Марків, РІ Содоми–м. Кам'янка-Бузька, 2020.-464 с. – 2020. – С. 83.
2. Дубина М. В., Шеремет О. М. Розвиток e-banking: світовий та вітчизняний досвід //Проблеми і перспективи економіки та управління. – 2019. – №. 2 (18). – С. 154-162.
3. Механчук Н. М. ОСОБЛИВОСТІ ЕЛЕКТРОННОГО БАНКІВСЬКОГО ОБСЛУГОВУВАННЯ //Сучасне українське студентство: проблеми та ціннісні орієнтації. – С. 245.
4. Тищенко О. І. Огляд сучасних тенденцій на ринку онлайн-банкінгу в Україні //Економіка і суспільство. – 2017. – №. 13. – С. 1237-1243.
5. Домінова І. В. Ризики електронного банкінгу та їх класифікація //Облік і фінанси. – 2016. – №. 3. – С. 69-76.
6. Chaimaa B., Najib E., Rachid H. E-banking Overview: Concepts, Challenges and Solutions //Wireless personal communications. – 2021. – Т. 117. – №. 2. – С. 1059-1078.
7. Risk Management for Electronic Banking and Electronic Money Activities. Basel : Basel Committee on Banking Supervision, Березень 1998
8. Малишевська О. О. ЕЛЕКТРОННИЙ БАНКІНГ: СУТНІСТЬ ТА РИЗИКИ //СЕКЦІЯ 1. ПУБЛІЧНІ ФІНАНСИ: СТАН ТА ПЕРСПЕКТИВИ РЕФОРМУВАННЯ. – 2018. – С. 217.
9. Литвиненко Д. Е. Управління ризиками електронного банкінгу : дис. – Сумський державний університет, 2021.
10. Домінова І. В., Доминова И. В. Шляхи мінімізації ризику шахрайства електронного банкінгу. – 2019.
11. Кінг Б. Банк 4.0 Нова фінансова реальність. - Litres, 2018.

12. Basel III : Global regulatory framework для більше résilient banks and banking systems. Basel : Basel Committee on Banking Supervision, June 2011.
13. Домінова І. В. Ризик шахрайства в умовах функціонування електронного банкінгу //Бізнес-навігатор. – 2017. – №. 4-2. – С. 92-98.
14. Мілентьєва А. М., Ісмайлов К. Ю. ФІШИНГ ЯК СУЧАСНА ПРОБЛЕМА СУСПІЛЬСТВА //Всі матеріали надані в авторській редакції та виражають персональну позицію учасника конференції. – С. 12.
15. Коляда О. Р., Раковська А. А. РОЗВИТОК ДІДЖИТАЛ ТЕХНОЛОГІЙ ТА РИЗИК ШАХРАЙСТВА В ІНТЕРНЕТ-БАНКІНГАХ КЛІЄНТІВ //ББК 65.262. 101_21я431. – С. 307.
16. Коваленко І. О. Обставини, що підлягають встановленню під час розслідування шахрайства у сфері використання банківських електронних платежів //Прикарпатський юридичний вісник. – 2021. – №. 1. – С. 98-101.
17. Коновалова І. О., Пивоваров В. В. ШАХРАЙСТВО В УМОВАХ ДІДЖИТАЛІЗАЦІЇ СУСПІЛЬСТВА //д-р юрид. наук, проф. АП Гетьман; д-р юрид. наук, проф. БМ Головкін; к. ю. н., доц. ОВ Таволжанський; к. ю. н. КС Остапко. – С. 168.
18. Курман О. В. и др. ФІНАНСОВЕ ШАХРАЙСТВО У ПРОСТОРАХ ІНТЕРНЕТ //Вісник студентського наукового товариства «ВАТРА» Вінницького торговельно-економічного інституту КНТЕУ. Вінниця: Редакційно-видавничий. – С. 217.
19. John SN, Okokpujie Kennedy O., Anele C., Olajide F., Chinyere Grace Kennedy (2016). Realtime fraud detection in the banking sector using data mining techniques/algorithm. 2016 International conference on computational science and computational intelligence (CSCI). P. 1186-1191. DOI: 10.1109/CSCI.2016.223
20. Kurt Fanning, Kenneth O. Cogger (1998) Neural network detection of management fraud using published fnancial data. Intelligent systems in accounting, fnance and management. Volume 7, issue 1. P. 21-41
21. Zuraidah Mohd Sanusi, Mohd Nor Firdaus Rameli, Yusarina Mat Isa (2015) Fraud schemes in banking institutions: prevention measures to avoid severe fnancial loss. Procedia economics and fnance. № 28. P. 107-113

22. Мальцева С. М. та ін. До ПИТАННЯ ПРО МЕТОДИ ЗАХИСТУ ВІД ФІНАНСОВОГО КИБЕРМОШЕНСТВА //Азимут наукових досліджень: економіка та управління. - 2020. - Т. 9. - №. 2. - С. 332-324.
23. Удалова З. У., Столбовой У. З. Огляд зарубіжного досвіду використання методу регресійного аналізу виявлення шахрайства з показниками фінансової звітності //Облік і статистика. - 2019. - №. 4. - С. 39-48.
24. Jarrod West, Maumita Bhattacharya, Rafgul Islam (2014) Intelligent fnancial fraud detection practices: an investigation. Процедури міжнародної конференції на охорони здоров'я та захисту в комунікаційних мережах. Volume 153. P. 186-203. DOI: 10.1007/978-3-319-23802-9_16
25. Левашов М. В., Овчинніков П. В. Ефективність класифікаторів для виявлення фрода у фінансових транзакціях // Питання кібербезпеки. - 2019. - №. 5 (33).
26. Rinky D. Patel, Dheeraj Kumar Singh (2013) Credit card fraud detection & prevention of fraud using genetic algorithm. International journal of soft computing and engineering (IJSCE). Volume 2, issue 6. P. 292-294
27. MohdAvesh Zubair Khan, JabirDaud Pathan, Ali Haider Ekbal Ahmed (2014) Credit card fraud detection system using hidden Markov model and k-clustering. International journal of advanced research in computer and communication engineering. Volume 3, issue 2. P. 5458-5461
28. Balamurugan M., Mathiazhagan P. (2015) Credit card transaction fraud detection system using fuzzy logic and k-means algorithm. International Journal of Innovative Research in Technology. Volume 2, issue 3. P. 171-176
29. Perols J. Financial statement fraud detection: Analysis of statistical and machine learning algoritms //Auditing: A Journal of Practice & Theory. - 2011. - Т. 30. - №. 2. - С. 19-50.
30. Awoyemi JO, Adetunmbi AO, Oluwadare SA Кредитна картка fraud detection using machine learning techniques: A comparative analysis //2017 International Conference on Computing Networking and Informatics (ICCNI). - IEEE, 2017. - С. 1-9.

31. Varmedja D. та ін. Credit card fraud detection-machine learning methods //2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH). - IEEE, 2019. - С. 1-5.
32. Yee OS, Sagadevan S., Malim NHAH Кредитна картка fraud detection використовуючи машину освітлення як data mining technic // Journal of Telecommunication, Electronic and Computer Engineering (JTEC). - 2018. - Т. 10. - №. 1-4. - С. 23-27.
33. Abbasi A. та ін. Metafraud: meta-learning framework for detecting financial fraud //Mis Quarterly. - 2012. - С. 1293-1327.
34. Kirkos E., Spathis C., Manolopoulos Y. Data mining techniques for detection of fraudulent financial statements //Expert systems with applications. - 2007. - Т. 32. - №. 4. - С. 995-1003.
35. Ravisankar P. та ін. Detection of financial statement fraud and feature selection using data mining techniques //Decision support systems. - 2011. - Т. 50. - №. 2. - С. 491-500.
36. Fanning KM, Cogger KO Neural network detection of management fraud using published financial data //Intelligent Systems in Accounting, Finance & Management. - 1998. - Т. 7. - №. 1. - С. 21-41.
37. Summers SL, Sweeney JT Fraudulently misstated financial statements and insider trading: An empirical analysis //Accounting Review. - 1998. - С. 131-146.
38. Abbott LJ, Park Y., Parker S. Ефекти з audit committee activity and independence on corporate fraud //Managerial Finance. - 2000.
39. Chen G. та ін. Ownership structure, corporate governance, і fraud: Evidence from China //Journal of Corporate Finance. - 2006. - Т. 12. - №. 3. - С. 424-448.
40. Maxwell AE, Warner TA, Fang F. Implementation of machine-learning classification in remote sensing: An applied review //International Journal of Remote Sensing. - 2018. - Т. 39. - №. 9. - С. 2784-2817.
41. Strobl C., Malley J., Tutz G. Введення в recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests //Psychological methods. - 2009. - Т. 14. - №. 4. - С. 323.