УДК 004.056.53
DOI: 10.31673/2409-7292.2024.020003

**Гайдур Г. І., Гахов С. О.,
Марченко В. В., Гайдур К. В.**

КОНЦЕПТУАЛЬНА МОДЕЛЬ ВИЯВЛЕННЯ ФІШИНГОВИХ АТАК НА ОСНОВІ ВИКОРИСТАННЯ МЕТОДІВ ОПОРНИХ ВЕКТОРІВ

У статті досліджено проблему виявлення кіберзагроз в інформаційній системі організацій, на прикладі фішингових атак. Саме фішинг є початковим вектором проведення атак щодо досягнення мети зловмисника та дозволить отримати інформацію щодо облікових записів користувача, даних мережі або дані адміністратора. Для вирішення проблеми виявлення кібератак на інформаційну систему організації, в роботі запропоновано концепцію виявлення фішингових атак. Основна ідея концепції полягає у застосуванні методів машинного навчання, які дозволяють проводити аналіз великих обсягів даних. Фішингові атаки можуть виявлятися саме через аналіз великих обсягів даних. Однією перевагою методів машинного навчання є те, що такі методи будуть виявляти неправомірні тенденції, відомі тактики та практики проведення атак такого типу. В роботі показано особливості визначення ознак щодо вхідних даних запропонованої моделі. Саме визначення ознак відомими техніками проведення фішингових атак, дозволили отримати набір даних, в якості вхідних ознак. Вхідні дані було покладено в основу методу опорних векторів, який класифікує отримані дані на фішингові та правомірні. В результаті дослідження отримано якісні характеристики моделі виявлення фішингових атак. Обчислені точність та чутливість моделі застосовано для SVM з лінійним ядром та радіально-базисною функцією (RBF). Перевірка адекватності та точності обраних моделей показано на прикладі ROC кривої, яка показує прогнози щодо визначення фішингу. Отже, запропонована концептуальна модель дозволяє розширити напрями досліджень, щодо виявлення фішингових атак в інформаційній системі організації шляхом додаткових методів обробки вхідних даних та удосконалення методів машинного навчання.

Ключові слова: інформаційна система, модель, кібербезпека, атака, фішинг, аномалія, ознаки виявлення вторгнень, інформаційна безпека, машинне навчання.

Вступ та постановка проблеми

Робота в IT-секторі відбувається в інформаційному просторі, яке складається з мереж, пристроїв, інформаційних систем та інформаційних технологій. Все це тісно переплітається між собою, що в свою чергу викликає актуальну проблему, яка виникає у зв'язку з кіберзагрозами, забезпечення кібербезпеки інформаційних систем організацій. Зростання рівня розвитку інформаційних технологій пропорційне зростанню різних типів атак на IT-сектор та IT-середовище.

Для розуміння поняття кіберзагрози в [1] надається визначення поняття кіберзагрози. Кіберзагрози - наявні та потенційно можливі явища і чинники, що створюють небезпеку життєво важливим національним інтересам України у кіберпросторі, справляють негативний вплив на стан кібербезпеки держави, кібербезпеку та кіберзахист її об'єктів. Дане визначення можливо інтерпретувати, з точки зору функціонування інформаційної системи будь-якої організації, як будь-які явища, які можуть порушити виконання організацією своїх бізнес-процесів та вплинути на їх стан кібербезпеки і, як наслідок, отримати фінансові збитки або втрату репутації.

Отже, в даному досліджені розглянуто, на прикладі фішингових атак, виявлення кіберзагроз в інформаційній системі організації, які впливають на стан кібербезпеки організацій. В роботі досліджено шляхи щодо виявлення фішингових атак в IT- середовищі з застосуванням методів машинного навчання на основі запропонованої концептуальної моделі.

Аналіз останніх досліджень та публікацій

Аналіз щорічного звіту IBM [2], показав організаціям інформацію про фінансові наслідки кіберзагроз. Виток конфіденційної інформації, внаслідок кіберзагроз, за останні роки з зросла на 15%. Все це вказує на те, що виток інформації і вплив її на стан кібербезпеки може відбуватися в будь-якій галузі: охорони здоров'я, енергетичній, фінансовій, промисловій. Розуміння такого впливу дозволить, організаціям приймати рішення, щодо побудови своєї системи кібербезпеки, спираючись на конкретні дані щодо кіберзагроз. Використання п'яти функцій NIST Cybersecurity Framework, як моделі кібербезпеки, є

хорошим планом побудови системи кібербезпеки: ідентифікація, захист, виявлення, реагування та відновлення. Така модель працюватиме незалежно від того, якого розміру буде організація [3].

Спираючись на дані звіту IBM, основою кіберзагрози є кібератака, яка спрямована на досягнення відповідної мети [1, 2]. Початковий вектор атаки показує, яким найпоширенішим способом можна отримати початкові дані для досягнення мети кібератаки. Згідно [1] атака типу фішинг займає перше місце і за витратами організацій складає 4,76 млн. доларів.

Фішингова атака – це спроба зловмисників обманом змусити вас поділитися інформацією або виконати дії, які надають їм доступ до облікових записів, комп'ютера чи навіть до мережі [3].

Аналіз останніх наукових публікацій показав, якщо розглядати фішингову атаку згідно NIST Cybersecurity Framework [3], велика увага приділена саме виявленню (detect) фішингової атаки.

В [4] автори сфокусували своє дослідження на виявленні фішингових атак з точку зору різних видів його прояву. В результаті аналізу отримали таксономію щодо традиційних та не традиційних методів виявлення фішингових атак. Але це не вирішує проблему їх виявлення в сучасних інформаційних системах. В [5] автори чітко визначили основну проблему щодо виявлення фішингових атак, як один із різновидів методів соціальної інженерії. В роботі зроблено аналіз щодо підходів виявлення фішингу: технічний та людський. В роботі застосовано ентропійний підхід для виявлення атаки та ознаки за якими це відбувається. Серед запропонованих методів виявлення автори пропонують застосувати моделі машинного навчання. Але в запропонованому алгоритмі недостатньо чітко сформульовані ознаки і їх кількість для подальшої оцінки моделі.

В [6] автори запропонували досліджувати виявлення фішингу саме методами штучного інтелекту, з використанням методів класифікації URL адрес. Але в роботі не прослідковується чітка концепція, за якою потрібно проводити дослідження. Аналіз [7] показав що методи штучного інтелекту дозволяють за рахунок зменшення розмірності функцій виявляти процеси притаманні фішинговим атакам.

Отже, можна зробити висновок, про необхідність створення моделі, яка дозволить за рахунок вхідних даних (ознак) покращити точність моделі щодо виявлення фішингових атак.

Мета статті полягає у побудові концептуальної моделі виявлення фішингових атак. Дана модель враховує вхідні параметри або функції, які притаманні і можуть бути ознаками фішингових атаках. Дана модель може бути застосована для виявлення кіберзагроз, а саме фішингових атак на інформаційну систему організації.

Виклад основного матеріалу

Концептуальна модель виявлення фішингових атак

В основі побудови концептуальної моделі виявлення фішингових атак, закладено основні типи поширених фішингових атак, з якими частіше за все можуть стикатися організації [8], які показано в таблиці 1.

Розуміння природи і фізичного змісту ознак фішингових атак дозволить на першому кроці отримати вхідні дані моделі $X(x_1, x_2, x_3 \dots x_n)$. Другим кроком буде визначення методики обробки вхідних даних. Третій крок полягає в застосуванні алгоритмів машинного навчання, які базуються на різних математичних концепціях та методах обробки даних. Наприклад, використання алгоритмів класифікації та регресії базуються на регресійних або класифікаційних моделях. Четвертий крок передбачає отримання результату моделі виявлення фішингових атак за обраними ознаками цільової функції Y (0, 1). Останній крок повинен включати в себе перевірку адекватності побудованої моделі виявлення фішингових атак в інформаційній системі організації (рис 1).

Таблиця 1

Типи фішингових атак

Тип фішингової атаки	Пояснення
Фішинг електронної пошти /Email Phishing	Найпоширеніша форма фішингу. Зловмисник надсилає оманливий електронний лист, який, здається, надійшов із законного джерела. Електронні листи часто вимагають конфіденційну інформацію, таку як облікові дані для входу, номери соціального страхування або фінансові дані.
Фішинг/ Spear Phishing	Більш цілеспрямована форма атаки. Зловмисник попередньо досліджує особу, щоб створити персоналізовані повідомлення. Це може підвищити ймовірність успіху, оскільки відправник виглядає більш надійним і поінформованим.
Китобійний промисел/Whaling	Націлено на високопоставлених осіб, старших менеджерів або керівників. Зловмисник адаптує листування до людей, які працюють в цій же сфері і часто заохочують суб'єкта переказати кошти або відмовитися від іншої важливої інформації. Це дозволяє зловмиснику отримати подальший доступ до системи.
Фармінг/Pharming	Включає перенаправлення користувачів на шахрайські веб-сайти, які відображають справжній веб-сайт. Зловмисник прагне змусити користувача ввести особисту інформацію на веб-сайті-дзеркалі для отримання подальшого доступу.

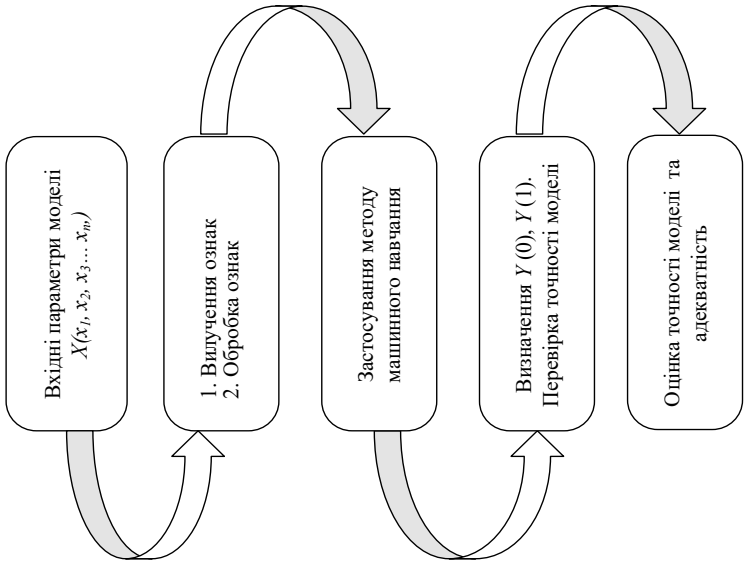


Рис. 1. Концептуальна модель виявлення фішингових атак

Призначення даної концепції полягає у виявлені з високою точністю та низьким рівнем помилкових спрацьовувань, щоб захистити користувачів інформаційної системи організації від кіберзагроз, які несуть в собі фішингові атаки.

Ознаки фішингових атак

Виявлення ознак фішингових атак, яке передбачає дослідження на базі машинного навчання та передбачає визначення вхідних даних (ознак) та аналіз різних ознак вебзастосунків та їх вмісту для виявлення шаблонів, типових для спроб проведення фішингових атак. На рис. 2 показано основні частини URL-адрес, які будуть використовуватись як вхідні дані для вилучення ознак. Опишемо основні ознаки, які моделі

машинного навчання можуть використовувати для виявлення фішингу з використанням аналізу URL-адрес:

невідповідність URL-адрес: вимагають перевірку, розбіжностей між відображеною URL-адресою та фактичною цільовою URL-адресою;

URL Shorteners: виявлення використання служб скорочення URL, які можуть приховати справжнє призначення;

маніпуляції субдоменами: пошук зміни в субдоменах, щоб імітувати законні веб-сайти;

перевірка IP-адреси: порівняння домену URL-адреси з її IP-адресою, щоб виявити підозрілі відхилення.

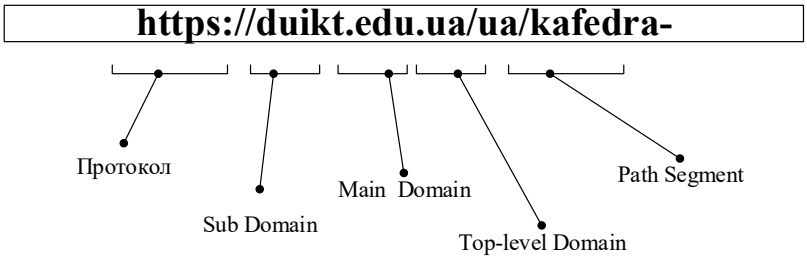


Рис.2. Складові URL-адреси

Отже, прийняття рішення про фішингову атаку для відповідних ознак можна визначити наступним чином:

1) *Скорочення послуги* (x_1)

Сервіси скорочення URL-адрес скорочують довгу URL адресу до значно коротшої довжини і перенаправляють на початкове посилання цільового веб-сайту.

Наприклад: URL-адресу "https://www.techopedia.com/phishing-statistics" скорочено до "surl.li/xxx/". Скорочене посилання surl.li або bit.ly перенаправляє на оригінальну URL-адресу https://www.techopedia.com/phishing-statistics". Ці скорочені посилання законно використовуються для аналітики. Навіть якщо кожна скорочена URL-адреса не є фішинговою, більшість з них потенційно можуть бути фішинговими [6].

$$f(URL) = \begin{cases} Phishing, \text{якщо } URL \text{ короткий} \\ Legitimate, \quad \text{інакше} \end{cases} \tag{1}$$

Зловмисники використовують ці сервіси для маскування несправжніх URL-адрес.

2) *Наявність IP-адреси* (x_2)

Використання IP-адреси замість URL-домену є тривожною ознакою зловмисних намірів. Приклад: "http://125.98.3.123/fake.html"

$$f(domain) = \begin{cases} Phishing, \text{ якщо доменна частина має } IP \\ Legitimate, \text{ Інакше} \end{cases} \tag{2}$$

3) *PageRank* (x_3)

PageRank (PR) - це імовірнісний алгоритм, який використовує Google у своїй пошуковій системі для оцінки якості веб-сайтів і ранжування веб-сторінок відповідно до результатів пошуку. PageRank працює, перевіряючи кількість і характер асоціацій зі сторінкою, щоб приблизно оцінити, наскільки важливим є сайт.

У контексті алгоритму PageRank виявлення фішингу включає ідентифікацію шаблонів або характеристик на веб-сторінках, які вказують на те, що вони ймовірно шахрайські або зловмисні. Основні ознаки, за якими можна виявити фішинг в рамках алгоритму PageRank:

Шаблони зв'язків. Неприродні посилання – це фішингові веб-сайти можуть мати неприродні шаблони посилань, наприклад раптовий наплив посилань із низькоякісних або підозрілих веб-сайтів. Ферми посилань включають в себе сторінки, які беруть участь у фермах посилань або інших схемах маніпулювання посиланнями, імовірно, мають нижчу якість і можуть вказувати на спроби фішингу.

Якість вмісту. Низькоякісний вміст полягає у тому що фішингові сторінки часто містять неякісний вміст, зокрема орфографічні помилки, граматичні помилки або безглуздий текст. Дублювання вмісту зосереджено на тому, що сторінки з дубльованим або плагіатним вмістом можуть свідчити про фішинг або спам.

Характеристики URL-адреси. Оманливі URL-адреси - це фішингові веб-сайти, які часто використовують оманливі URL-адреси, які імітують законні сайти. Виявлення URL-адрес із неправильно написаними доменними іменами чи додатковими субдоменами може допомогти виявити спроби фішингу.

Переспрямування URL-адреси. Фішингові сторінки можуть використовувати методи переспрямування URL-адрес, щоб приховати справжню URL-адресу призначення, спрямовуючи користувачів на шкідливі сайти.

Репутація сторінки. Фішингові веб-сайти часто використовують нещодавно зареєстровані домени, щоб уникнути виявлення. Виявлення нещодавно зареєстрованих доменів може допомогти визначити можливі спроби фішингу.

Репутація домену. Сторінки, розміщені в доменах з історією зловмисної діяльності або низькою репутацією, швидше за все, є фішинговими сайтами.

Структура сторінки та функціональність. Фішингові сторінки часто містять фальшиві форми входу або запитують конфіденційну інформацію від користувачів. Виявлення сторінок із підозрілими формами чи полями введення може допомогти виявити спроби фішингу.

Отже, основною гіпотезою для визначення ознаки буде те, що більш значущі сайти, швидше за все, отримають більше посилань з інших сайтів. Це відіграє важливу роль у виявленні фішингових сайтів, оцінюючи сайт від 0 до 1. Веб-сайт є важливим або вважається з найкращою якістю, якщо:

$$f(URL) = \begin{cases} Phishing, \text{якщо PageRank} < 0{,}5 \\ Legitimate, \text{ інакше} \end{cases} \quad (3)$$

4) DoubleSlashRedirecting (x_4)

Коли в URL-адресі з'являється послідовність "//", вона зазвичай позначає початок компонента шляху URL-адреси. У більшості випадків це за своєю суттю не означає фішинг. Однак під час спроб фішингу зловмисники можуть неправильно використовувати послідовність "//" або маніпулювати нею, щоб створити оманливі URL-адреси, які на перший погляд здаються законними. Ось як зловмисники можуть цим скористатися:

Підробка легітимних веб-сайтів полягає в тому, що зловмисники можуть створювати фішингові URL-адреси, які починаються з "//", за якою йде законне доменне ім'я. Наприклад, вони можуть створити URL-адресу на зразок "//example.com", щоб вона виглядала так, ніби вона є частиною домену "example.com". Ця техніка має на меті змусити користувачів ввести в оману, що вони відвідують законний веб-сайт, тоді як насправді вони переспрямовуються на фішинговий сайт.

Обфускація та переспрямування полягає в тому, що зловмисники можуть використовувати "//" у поєднанні з методами переспрямування URL-адреси, щоб приховати справжнє призначення URL-адреси. Використовуючи засоби скорочення URL-адрес або інші методи перенаправлення, зловмисники можуть приховувати шкідливі URL-адреси за, здавалося б, нешкідливими, що ускладнює користувачам виявлення спроб фішингу.

Наявність символу "//" всередині URL-адреси означає, що користувач буде перенаправлений на іншу сторінку або веб-сайт. Наприклад, URL має вигляд: "http:/www.legitimate.com//http://www.phishing.com." Ми перевіряємо позицію символу "//" в URL-адресі. Допустимо, щоб символ "//" знаходився на 8-й позиції URL-адреси

$$f(URL) = \begin{cases} Phishing, \text{якщо позиція DoubleSlash} > 8 \\ Legitimate, \text{ інакше} \end{cases} \quad (4)$$

5) HTTPs TOKEN (x_5)

Використання "HTTPS TOKEN" як терміна в контексті фішингу може бути дещо неоднозначним, оскільки воно не має прямого відношення до поширеної техніки фішингу. Однак можна інтерпретувати деякі поширені тактики фішингу. Підроблений HTTPS полягає в тому, що деякі спроби фішингу можуть використовувати HTTPS, щоб створити видимість легітимності. Зловмисники можуть отримати SSL-сертифікати для своїх фішингових веб-сайтів, завдяки чому вони виглядають безпечними за допомогою протоколу HTTPS. Це може змусити користувачів подумати, що сайт безпечний, хоча насправді він шкідливий. У цьому випадку частина "HTTPS" може бути використана, щоб ввести користувачів в оману, щоб переконати, що сайт безпечний і надійний. Приклад: http://https-www-pay-it-apps-mpp-home.solar.com/

Неправильне використання термінології дозволяє фішерам створити відчуття достовірності під час своїх спроб фішингу. Використання таких термінів, як "HTTPS" або "токен", може бути спробою зробити фішингове повідомлення більш законним або переконливим для одержувача, особливо якщо одержувач не знайомий із технічними деталями веб-безпеки:

$$f(URL) = \begin{cases} Phishing, \text{якщо "HTTPS" використовується в домені} \\ Legitimate, \text{ інакше} \end{cases} \quad (5)$$

Застосування методів машинного навчання до отриманих ознак фішингових атак

Кожна ознака з вилучених даних складе перелік вхідних даних, для яких буде застосовано алгоритм машинного навчання. Дані для дослідження були взяті з [9]. Набір даних містить 48 ознак, отриманих із 5000 фішингових веб-сторінок і 5000 законних веб-сторінок. Перші п’ять рядків датасету зображено на рис. 3 [9].

	id	NumDots	SubdomainLevel	PathLevel	UrlLength	NumDash	NumDashInHostname	AtSymbol	TildeSymbol
0	1	3	1	5	72	0	0	0	0
1	2	3	1	3	144	0	0	0	0
2	3	3	1	2	58	0	0	0	0
3	4	3	1	6	79	1	0	0	0
4	5	3	0	4	46	0	0	0	0

Рис. 3. Набір даних фішингових сторінок [9]

Для виявлення статистичної залежності між ознаками у наборі даних застосовано кореляцію Спірмена. Отримана матриця кореляції за тепловою шкалою (рис. 4), дозволяє виявити та оцінити статистичну залежність між ознаками в нашому наборі даних.

Особливість кореляції Спірмена полягає в тому, що вона вимірює ступінь монотонної залежності між двома змінними, не обов'язково лінійну. Така матриця зручна для аналізу та виявлення нестандартних зв'язків, які можуть бути присутніми в даних. Аналіз отриманої

матриці показує, що між отриманими даними є відносно мала кількість висококорельованих даних, які б можна було б зменшити.

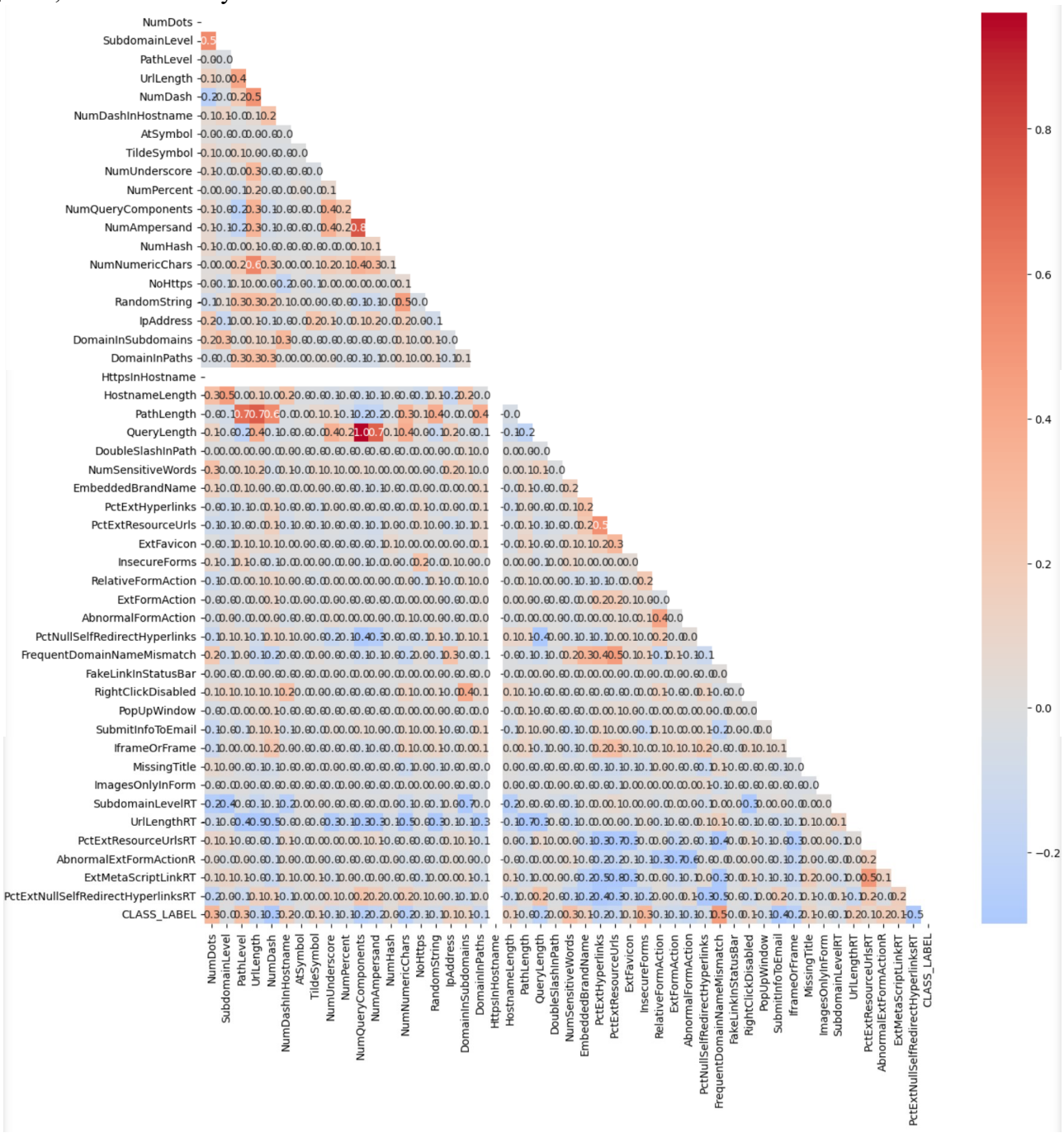


Рис. 4. Матриця кореляції Спірмена

Розуміння, яка кількість фішингових даних є в отриманому в [9] зображено на рис. 5. Дані фішингового датасету мають відповідні *Label i* поділені 50 на 50 з легітимними даними.

Приклад застосування методу опорних векторів

Прикладом алгоритму машинного навчання, як інструменту виявлення аномалій в даній роботі розглядається метод опорних векторів (SVM – Support Vector Machines). Даний алгоритм машинного навчання використовується для класифікації та регресії. SVM є інструментом який має свої переваги. Векторні представлення вхідних даних у представлені у високорозмірному просторі функцій, що дозволяє SVM ефективно розділяти дані, що мають

складні зв'язки. Це особливо є важливим при аналізі великих обсягів вхідних даних, що часто відбувається в області кібербезпеки.

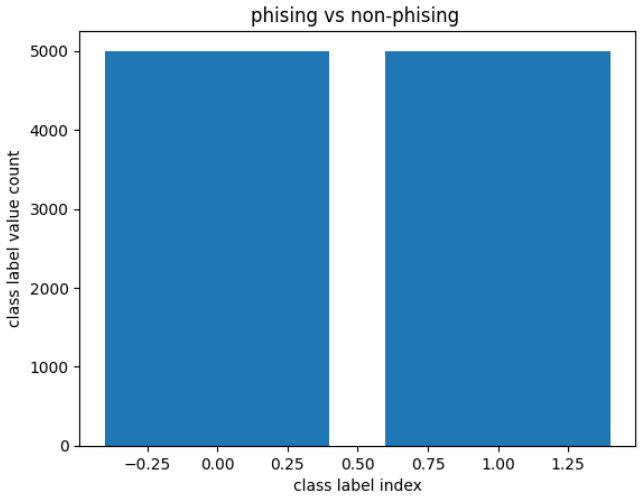


Рис. 5. Розподіл фішингових та легітимних даних

Фішингові сайти можуть мати складні залежності між характеристиками, що потребують моделі, здатної розділити класи в не лінійних формах. SVM може використовувати ядро для перетворення даних в високорозмірний простір, де вони можуть бути краще розділені. SVM є ефективним інструментом для виявлення аномалій, таких як фішинг, завдяки своїм властивостям, які дозволяють працювати з високорозмірними даними, нелінійними залежностями та невеликими вибірками.

У даному дослідженні метод опорних векторів (SVM) використано для аналізу даних щодо виявлення фішингу і розглядається у розрізі застосування різних типів ядр. Лінійне ядро та радіально-базисна функція (RBF) використано в SVM. Основною відмінністю між застосуванням лінійного та RBF ядра полягає в їхніх властивостях розділення даних. Лінійне ядро використовує лінійні границі рішення, тоді як RBF може розділяти дані, використовуючи не лінійні границі рішення, що робить його корисним для нелінійних вибірок даних. Застосування використання даних моделей для даних з фішингом полягає у визначені алгоритму, який дозволить отримати вищі результаті точності моделі для класифікації вхідних даних. Дослідження виконувалось в середовищі Jupyter Notebook з використання бібліотек машинного навчання на мові Python.

Аналіз результатів отриманих даних для обчислення точності виявлення фішингових атак проводився з використання матриці плутанин. Матриця плутанини - це інструмент, який зручно використовувати для оцінки продуктивності моделі класифікації. Візуалізація процесу прогнозування дозволяє порівнювати фактичні значення класів з передбаченими моделлю. Результати матриці плутанини показано для SVM – linear та SVM – RBF відповідно на рис.6. Отриманні данні необхідні для обчислення метрик точності та чутливості моделі.

Обчислення метрик класифікації дозволило отримати значення загальної точності (Accuracy), точності (Precision), яка вказує на відсоток правильно виявлених позитивних прогнозів серед всіх позитивних прогнозів та чутливості (Recall), яка вказує на відсоток правильно виявлених позитивних класів [10]. Розрахунки проводились за формулами (6), (7), (8) та показано в таблиці 2.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (6)$$

$$Precision = TP / (TP + FP) \quad (7)$$

$$Recall = TP / (TP + FN) \quad (8)$$

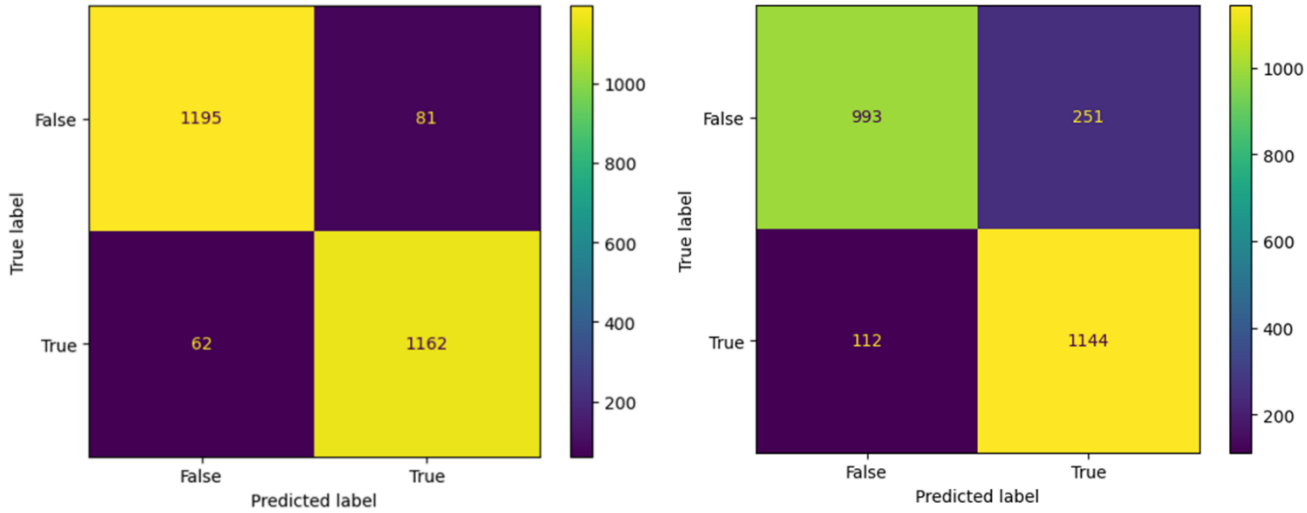


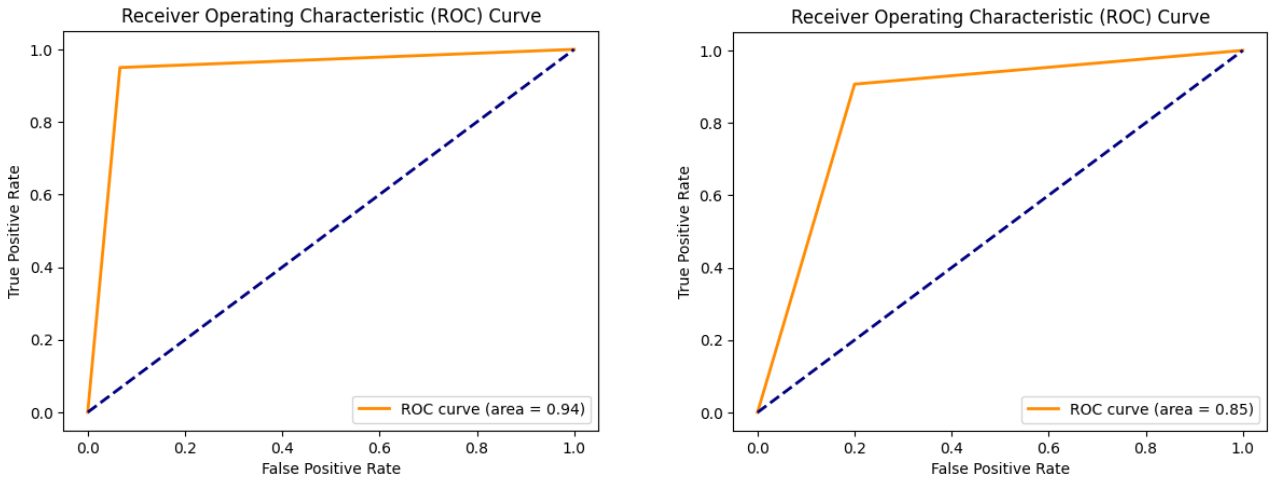
Рис. 6. Матриця плутанини: а) SVM – linear; б) SVM – RBF

Таблиця 2

Результати обчислення метрик класифікації

	SVM – linear	SVM – RBF
Accuracy	0.942	0.8548
Precision	0.9374034003091191	0.8200716845878137
Recall	0.9498825371965545	0.910828025477707

Аналіз обчислених метрик класифікації показав, що кращі результати надає SVM – linear, в порівнянні з SVM – RBF. Це пов'язано з тим, що дані лінійно роздільні або кількість їх ознак невелика в порівнянні з обсягом даних. Отже, лінійне ядро підходить для задач, де взаємозв'язки між ознаками та вихідною змінною лінійні. Використання додаткових метрик, таких як крива ROC надають змогу визначити адекватність класифікатора та порівняти його з іншими отриманими моделями, як це показано на рис. 7.



а) б)

Рис. 7. Метрика ROC: а) SVM – linear; б) SVM – RBF

Чим більше площа під кривою ROC, тим краще модель робить прогнози щодо визначення класів. Отже, отримані в попередньому кроці результати підтверджуються результатами отриманих ROC кривих, з яких видно що площа SVM – linear більша і відображає частку правильно виявлених позитивних прикладів відносно загальної кількості позитивних прикладів.

Висновки

Запропонована концептуальна модель виявлення фішингових атак дозволяє за відповідними етапами проводити дослідження щодо застосування методів машинного навчання. Розуміння природи, тактик проведення фішингових атак дозволяє визначити основні ознаки, за якими буде проводитись класифікація даних. В подальших дослідженнях, відповідно до запропонованої концепції виявлення фішингових атак, є можливість застосувати інші методи SVM та розробити методики обробки вхідних даних моделі для підвищення точності моделей виявлення аномалій типу фішингу. Запропонована концептуальна модель дозволяє підвищити розуміння та важливість процесу виявлення кіберзагроз в забезпеченні кібербезпеки інформаційних систем організацій.

Дана робота виконувалась в рамках виконання науково-дослідної роботи за темою: «Методологія виявлення шкідливих процесів в інформаційних системах (реєстраційний номер НДР 0121U113613).

Перелік посилань

1. Про основні засади забезпечення кібербезпеки України, Закон України № 2163-VIII (2024) (Україна). https://zakon.rada.gov.ua/laws/show/2163-19#Text
2. Cost of a data breach 2023 | IBM. IBM in Deutschland, Österreich und der Schweiz. https://www.ibm.com/reports/data-breach
3. Phishing. NIST. https://www.nist.gov/itl/smallbusinesscyber/guidance-topic/phishing
4. Alanezi, M. (2021). Phishing Detection Methods: A Review. Technium: Romanian Journal of Applied Sciences and Technology, 3(9), 19–35. https://doi.org/10.47577/technium.v3i9.4973
5. Buchyk, S., Shutenko, D., & Toliupa, S. Phishing Attacks Detection. Information Technology and Implementation. (IT&I-2022), November 30 - December 02, 2022, Kyiv, Ukraine. pp.193-201. https://ceur-ws.org/Vol-3384
6. Indrasiri, P. L., Halgamuge, M. N., & Mohammad, A. (2021). Robust Ensemble Machine Learning Model for Filtering Phishing URLs: Expandable Random Gradient Stacked Voting Classifier (ERG-SVC). IEEE Access, 9, 150142–150161. https://doi.org/10.1109/access.2021.3124628
7. Rushton, J. (2024, 1 березня). 50+ Phishing Statistics You Need to Know – Where, Who & What is Targeted. Cyber Threats. https://www.techopedia.com/phishing-statistics
8. Chiew, K. L., Tan, C. L., Wong, K., Yong, K. S. C., & Tiong, W. K. (2019). A new hybrid ensemble feature selection framework for machine learning-based phishing detection system. Information Sciences, 484, 153–166. https://doi.org/10.1016/j.ins.2019.01.064
9. Phishing Dataset for Machine Learning. Kaggle: Your Machine Learning and Data Science Community. https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning

Haidur, G. I. (2021). Detection of traffic anomalies in the information systems of organizations using Machine Learning methods on the base of algorithms for forecasting category fields. Telecommunication and Information Technologies, 73(4). https://doi.org/10.31673/2412-4338.2021.044153

Надійшла 06.04.2024