ЄВРОПЕЙСЬКІ ОРІЄНТИРИ РОЗВИТКУ УКРАЇНИ: НОВІ ВИМІРИ У ВОЄННИЙ ЧАС
Матеріали Міжнародної науково-практичної конференції
2

ЄВРОПЕЙСЬКІ ОРІЄНТИРИ РОЗВИТКУ УКРАЇНИ: НОВІ ВИМІРИ У ВОЄННИЙ ЧАС

МАТЕРІАЛИ
Міжнародної науково-практичної конференції

Одеса, 16 травня 2025 року

Том 2

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Одеська юридична академія»

ЄВРОПЕЙСЬКІ ОРІЄНТИРИ РОЗВИТКУ УКРАЇНИ: НОВІ ВИМІРИ У ВОЄННИЙ ЧАС

МАТЕРІАЛИ
Міжнародної науково-практичної конференції

Одеса, 16 травня 2025 року

Том 2

Одеса
2025

УДК 005.332.2(4):316.42(477)“364”“20”(062.552)
Є 24

Рекомендовано до друку вченою радою
Національного університету «Одеська юридична академія»
(протокол № 8 від 16.05.2025 р.)

За загальною редакцією ***С. В. Ківалова***

Є 24 **Європейські орієнтири розвитку України: нові виміри у воєнний час** : у 2 т. : матеріали Міжнар.наук.-практ. конф. (Одеса, 16 травня 2025 р.) / за заг. ред. С. В. Ківалова. – Одеса : Фенікс, 2025. – Т. 2. – 716 с.
ISBN 978-617-8430-61-0

Матеріали Міжнародної науково-практичної конференції «Європейські орієнтири України: нові виміри у воєнний час» містять результати наукових досліджень і практичних напрацювань вчених, фахівців і військовослужбовців, здійснені в умовах повномасштабної війни. У збірнику представлено як теоретичні, так і емпіричні дослідження у галузях філософії, загальної теорії держави і права, історії права, сучасних соціально-політичних процесів, соціології та психології.

Особливу увагу приділено аналізу загроз національній безпеці в конституційному, міжнародному та європейському правовому контекстах, а також питанням трудового, соціального, земельного, аграрного та екологічного права. Окремі розділи присвячено функціонуванню економіки та підприємництва в умовах євроінтеграції.

У збірнику також розглянуто проблематику цифрової трансформації, захисту інформації та кібербезпеки в умовах воєнного часу, методику викладання іноземних мов, перекладознавство, лінгвістику і журналістику. Висвітлено наукові здобутки в адміністративному, фінансовому, морському та митному праві, реформуванні судової системи, діяльності прокуратури, правоохоронних органів та адвокатури. Окремо проаналізовано питання кримінального права, кримінального процесу, кримінології, криміналістики, судової експертизи, медико-психологічного забезпечення правосуддя, а також аспекти цивільного, сімейного, господарського права, інтелектуальної власності та патентного судочинства. Розглянуто також актуальні питання діяльності бібліотек у закладах вищої освіти та фізичної підготовки здобувачів вищої освіти.

Збірник адресовано науковцям, викладачам, здобувачам вищої освіти, а також практикам у галузях права, економіки, соціології, політології, психології, філології, журналістики та кібербезпеки.

УДК 005.332.2(4):316.42(477)“364”“20”(062.552)

Відповідальний за випуск ***М. Р. Аракелян***

ISBN 978-617-8430-61-0

СЕКЦІЯ 22. ІНФОРМАТИЗАЦІЯ ТА ЦИФРОВІЗАЦІЯ СУСПІЛЬСТВА В ЕПОХУ СТРІМКОГО РОЗВИТКУ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ

СЕКЦІЯ 23. КІБЕРБЕЗПЕКА ОСОБИ, СУСПІЛЬСТВА ТА ДЕРЖАВИ ЯК ФУНДАМЕНТАЛЬНИЙ ЕЛЕМЕНТ ЗАБЕЗПЕЧЕННЯ ОБОРОНОЗДАТНОСТІ УКРАЇНИ

Дика Анастасія Іванівна
Національний університет «Одеська юридична академія»,
аспірантка кафедри інформаційних технологій

ЗАСТОСУВАННЯ NLP-МЕТОДІВ ДЛЯ ВИЯВЛЕННЯ XSS-АТАК

Cross-Site Scripting (XSS) – це один з найбільш поширених і небезпечних типів вразливостей у веборієнтованому програмному забезпеченні. Сутність XSS-атак полягає у впровадженні шкідливого скриптового коду, переважно JavaScript, у вебсторінки, які надалі відображаються іншим користувачам. Унаслідок таких атак зловмисники можуть здійснювати викрадення куків, перенаправлення, зміну вмісту сторінок або здійснювати фішингові атаки [1]. Традиційні методи боротьби з XSS, серед яких апаратні та програмні фаєрволи, стосуються використання правил фільтрації символів, екранування небезпечних конструкцій та ручного складання чорних списків. Проте кіберзлочинці можуть обійти ці методи шляхом обфускації шкідливих навантажень, що робить їх неефективними проти XSS-атак, спеціально розроблених для обходу WAF [2].

У цьому контексті постає задача пошуку нових, більш гнучких способів виявлення атак. Застосування методів обробки природної мови (Natural language processing, NLP) до задачі виявлення XSS є перспективним напрямом, оскільки введення користувача можна розглядати як текстовий потік, що містить семантичні й синтаксичні ознаки потенційної загрози. Це зумовлює задачу класифікації текстових повідомлень на безпечні та потенційно шкідливі (тобто такі, що містять XSS). На відміну від традиційних систем, які працюють на основі суворих правил, NLP-моделі здатні навчатися на реальних прикладах і виявляти складні шаблони, притаманні атакам.

Підготовка даних є критично важливою складовою в побудові моделі виявлення XSS. Для цього можна сформувати набір коротких повідомлень, частина з яких є легітимним текстом, а інші – шкідливими скриптами XSS. Моделі оброблення тексту, зокрема TF-IDF (Term Frequency-Inverse Document Frequency) векторизація, дозволяє визначити важливість термів (слів або фраз) у документі відносно загальної кількості документів.

TF-IDF у задачі виявлення шкідливого вводу застосовується для перетворення тексту користувача у числові ознаки, які можна аналізувати автоматично. На першому етапі формується набір текстових прикладів, які маркують як «шкідливі» або «безпечні». Це дозволяє моделі розрізняти шаблони, притаманні шкідливим фрагментам XSS [3].

Далі кожне введення розбивається на терми – це можуть бути слова, символи або *n*-грами (послідовності символів довжини *n*). Після цього для кожного терму в кожному введенні обчислюється його TF-IDF-значення. TF (Term Frequency) відображає, наскільки часто терм з'являється у конкретному введенні. IDF (Inverse Document Frequency) зменшує вагу тих термів, які зустрічаються дуже часто в усьому корпусі, тобто є «звичними», і підвищує вагу рідкісних, потенційно більш

значущих термів. Тим самим, TF-IDF дозволяє виявити специфічні символи або фрази, що часто зустрічаються саме у шкідливих введеннях, але не у звичайних.

Внаслідок векторизації кожен текст перетворюється у вектор чисел, який відображає важливість термів. Ці вектори можуть бути використані як вхідні ознаки для класифікаційних моделей машинного навчання. В процесі аналізу можна побачити, що, наприклад, терми на зразок “<script>“, “onerror=“, “javascript:” часто мають високі TF-IDF-значення саме у шкідливих прикладах. Це дозволяє не лише автоматизувати виявлення небезпечних введень, а й інтерпретувати, які саме фрагменти тексту вважаються підозрілими.

Результати роботи такої системи можна оцінити за допомогою метрик точності (precision), повноти (recall) та F1-міри, які дозволяють визначити ефективність виявлення атак за умов, коли класи є незбалансованими (більшість повідомлень – безпечні). Для цього автором було проведено дослідження обробки тексту та навчання класифікатора з набором даних Kaggle, а саме: XSS Detection Dataset [4] шляхом векторизації TF-IDF.

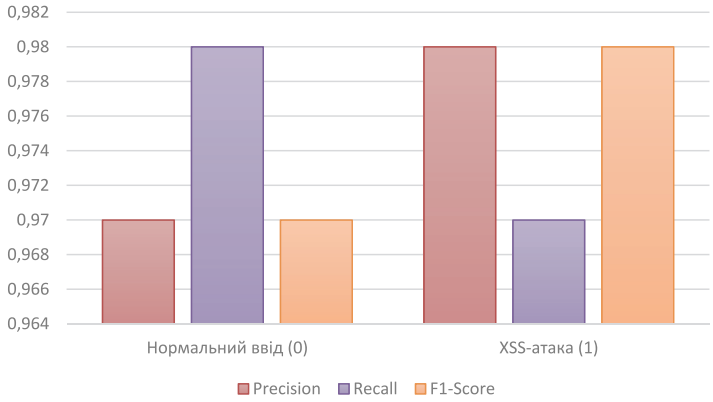

Рисунок 1 – Оцінка класифікації XSS через TF-IDF + Logistic Regression

Результати експерименту з виявлення XSS-атак за допомогою TF-IDF-векторизації та логістичної регресії демонструють високу ефективність моделі. Загальна точність класифікації становить приблизно 98%, що свідчить про здатність моделі правильно класифікувати як шкідливі, так і безпечні рядки. Значення precision для класу XSS досягає 98%, тобто, коли модель позначає введення як атаку, вона майже завжди має рацію. Recall для цього класу становить приблизно 97%, що означає, що більшість справжніх атак виявляються коректно. Тим самим, модель майже не пропускає шкідливі скрипти і рідко помиляється при виявленні. Високе значення F1-міри (98%) підкреслює збалансованість між

precision та recall. Отримані результати свідчать, що TF-IDF добре захоплює специфічні шаблони, характерні для XSS (“<script>“, “onerror=“, “alert()”), а логістична регресія як класифікатор ефективно розрізняє такі шаблони. Загалом, навіть базовий підхід без глибокого аналізу структури HTML/JS, дозволяє досягти високої якості класифікації, що є перспективним для створення легких і швидких систем виявлення на стороні бекенду вебзастосунків.

Графічне подання метрик класифікації (precision, recall, f1-score) для обох класів показує, що всі основні метрики – precision, recall і f1-score – для обох класів перебувають на рівні 97-98%. Це вказує на високий рівень збалансованості моделі та її здатність ефективно виявляти як XSS-атаки, так і безпечні введення. Така стабільність у показниках робить підхід придатним для практичного використання у вебзастосунках.

Однак, попри високу ефективність, TF-IDF має обмеження: він не враховує контекст символів, послідовність токенів та логіку виконання шкідливого коду. Це відкриває перспективи для подальших досліджень у напрямі використання контекстно-залежних моделей, зокрема глибоких нейронних мереж (LSTM, GRU, BERT), які здатні краще моделювати структуру і поведінку введеного коду.

Загалом TF-IDF є ефективним стартовим інструментом, але його поєднання з глибшими моделями дозволить покращити точність та стійкість систем виявлення атак у реальних умовах.

Список використаних джерел:

1. “Cross Site Scripting (XSS).” OWASP Cheat Sheet Series. *OWASP Foundation.* URL: https://owasp.org/www-community/attacks/xss/

2. Bronjon G., Tasiruddin A., Hemanta K. Detection of XSS Attacks in Web Applications: A Machine Learning Approach. *International Journal of Innovative Research in Computer Science and Technology (IJIRCST).* 2021. Vol. 9. No.1. P. 1-10. URL: https://doi.org/10.21276/ijircst.2021.9.1.1.

3. Almarashy A. H., Feizi-Derakhshi M. R., Salehpour P. Enhancing Fake News Detection by Multi-Feature Classification. *IEEE Access*. 2023. Vol. 11. P. 139601–139613. URL: https://doi.org/10.1109/ACCESS.2023.3339621.

4. Cross site scripting XSS dataset for Deep learning. URL: https://www.kaggle.com/datasets/syedsaqlainhussain/cross-site-scripting-xss-dataset-for-deep-learning.

Ключові слова: XSS-атаки, веббезпека, машинне навчання, обробка природної мови, виявлення вразливостей, векторизація тексту, класифікація введень.

Keywords: XSS attacks, web security, machine learning, natural language processing, vulnerability detection, text vectorization, input classification.