користувачів та подальшого навчання синтетичних користувачів у задачах автоматизованого UX-тестування.

Отримані результати є перспективними для навчання моделей штучного інтелекту, здатних імітувати логіку прийняття рішень користувача, виявляти когнітивні бар'єри та прогнозувати потенційні UX/UI-проблеми під час взаємодії з цифровими банківськими системами. Водночас встановлено, що метод має певні обмеження, серед яких вплив процесу вербалізації на природність поведінки користувача, збільшення часу виконання сценаріїв та складність масштабування досліджень [4]. У зв'язку з цим доцільним є використання методу вербалізації мислення як складової комплексного підходу, що поєднує аналіз користувацьких логів, UX-метрик, записів сесій користувачів та автоматизованих методів оцінювання інтерфейсів.

1. Garrett J. J. The Elements of User Experience: User-Centered Design for the Web and Beyond. Berkeley: New Riders, 2011. 192 p.
2. Nielsen J. Usability Engineering. San Francisco: Morgan Kaufmann, 1994. 362 p.
3. Дослідження розвитку цифрового банкінгу в Україні. International Science Group. – 2024. – URL: https://isg-journal.com.
4. Ericsson K. A., Simon H. A. Protocol Analysis: Verbal Reports as Data. Cambridge: MIT Press, 1993. 434 p.
5. Barnum C. M. Usability Testing Essentials: Ready, Set...Test!. Cambridge: Morgan Kaufmann, 2020. 384 p.
6. Krug S. Don't Make Me Think: A Common Sense Approach to Web Usability. 3rd ed. Berkeley: New Riders, 2014. 216 p.

## Розроблення та дослідження методів виявлення фішингових веб-ресурсів на основі машинного навчання

УДК 004.056

Євген Бакаляр[1], Олександр Сиропятов[2]

*Національний університет «Одеська політехніка»,*
*[1]10328112@stud.op.edu.ua, [2]o.a.syropiatov@op.edu.ua*

*Вступ.* Фішингові веб-ресурси становлять значну загрозу для корпоративної та особистої інформації. Традиційні методи протидії (сигнатурний аналіз, статичні чорні списки) малоефективні через динамічність фішингових кампаній та короткий цикл життя злоякісних доменів (кілька годин–24 години). Зловмисники використовують HTTPS та методи імітації легітимних сайтів, що утруднює виявлення класичними засобами. Актуальним рішенням є застосування машинного навчання для автоматизованого аналізу URL-адрес на основі лексичних ознак, що дозволяє виявляти нові фішингові ресурси в режимі реального часу.

*Мета.* Розроблення та дослідження автоматизованого виявлення фішингових веб-ресурсів шляхом виявлення та аналізу їх лексичних ознак із застосуванням алгоритмів машинного навчання.

*Основна частина.* Розроблена концепція детекції включає модулі: збір даних, попередню обробку URL-адрес, екстракцію лексичних ознак, формування векторного представлення. Технологічний стек: Python, Pandas, Scikit-learn, XGBoost. Ключові показники: обсяг неалфавітно-цифрових знаків, використання IP-адресації, кількість піддоменів, скорочені URL, релевантні терміни. Класифікацію реалізовано на Random Forest, XGBoost, Logistic Regression та Decision Tree. Random Forest забезпечує мінімізацію перенавчання; XGBoost опрацьовує нелінійні залежності; Logistic Regression — базовий класифікатор; Decision Tree — оцінювання окремих дерев. Програмний продукт може бути імплементований у корпоративні шлюзи, системи моніторингу трафіку, Web Application Firewall та endpoint-рішення.

*Експериментальні дослідження та результати.* Апробація здійснювалася на датасету 10 000 URL-адрес (легітимні та фішингові). Дані очищено та нормалізовано. Перевірка ґрунтувалася на перехресній перевірці з показниками Accuracy, Recall, F1-score. Результати: Random Forest — Accuracy 96.4%, Precision 95.8%, Recall 96.9%, F1-score 96.3%; XGBoost — 95.7%, 95.1%, 95.9%, 95.5% (з підвищеними витратами ресурсів); Logistic Regression — 91.8%; Decision Tree — 89.6%. Нижчі показники останніх двох можуть використовуватись у системах із обмеженими ресурсами. Результати обґрунтовують перспективність ML-алгоритмів у системах автоматизованого моніторингу.

*Висновок.* Дослідження підтвердило ефективність ML-методів для виявлення фішингових веб-ресурсів на основі лексичного аналізу URL-адрес. Модель Random Forest продемонструвала найкращу точність (96.4%) у порівнянні з іншими класифікаторами. Система забезпечує детекцію в режимі реального часу без аналізу вмісту сторінок. Результати обґрунтовують доцільність інтеграції розробленого підходу в корпоративні шлюзи, WAF та SIEM-системи для оперативного моніторингу та блокування фішингових загроз. Система готова до практичного впровадження у системи інформаційної безпеки.

1. Мельник О.В. Методи машинного навчання у задачах виявлення кіберзагроз. Кібербезпека та захист інформації. 2023. №4. С. 12–19.
2. Anti-Phishing Working Group (APWG). Phishing Activity Trends Report, 4th Quarter 2025.
3. Федоренко С.М. Аналіз лексичних характеристик URL для протидії фішингу. Сучасні інформаційні системи. 2024. №1. С. 74–80.
4. Грищенко М.О. Оцінка ефективності бінарної класифікації у завданнях інформаційної безпеки. Сучасні проблеми захисту інформації. Львів: НУ "Львівська політехніка", 2023. С. 88–95.
5. Бойко В.І. Застосування алгоритмів машинного навчання для аналізу шкідливих URL-адрес. Київ: КПІ ім. Ігоря Сікорського, 2022. 112 с.