УДК 004.056.5:004.8

Ігор РУДЬКО, Юрій ДОРОФЕЄВ

Національний університет «Одеська політехніка»

РОЗРОБЛЕННЯ ЗАСТОСУНКУ ДЛЯ НАВЧАННЯ РОЗПІЗНАВАННЮ ВІРУСНИХ ЕЛЕКТРОННИХ ЛИСТІВ

Вступ. В умовах інтенсивного розвитку цифрових технологій електронна пошта є одним із ключових засобів професійної та особистої комунікації. Однак саме через електронну пошту щорічно здійснюється понад половина усіх кібератак, спрямованих на викрадення даних, встановлення шкідливого програмного забезпечення або компрометацію корпоративних мереж[1].

Особливої актуальності набуває питання навчання користувачів методам виявлення підозрілих повідомлень та формування стійких навичок кібергігієни.

Мета. Метою роботи є розроблення інтерактивного навчального застосунку, який моделює процеси отримання, аналізу та оцінювання електронних листів із можливими ознаками фішингу або вірусної активності. Основна ідея полягає у поєднанні генеративних мовних моделей і методів машинного навчання для створення навчального середовища, що імітує реальні сценарії атак.

1. Архітектура та функціональні модулі системи

Запропонована система складається з трьох основних модулів: генерації контенту, аналізу та оцінювання безпеки. Така модульна структура забезпечує гнучкість і можливість масштабування розробки, що дозволяє адаптувати застосунок для різних категорій користувачів – від студентів до працівників IT-компаній.

Модуль генерації відповідає за створення фішингових і вірусних листів, максимально наближених до реальних. Для цього використовується локальна велика мовна модель Ollama, здатна формувати тексти різного стилю - від офіційних бізнес-листів до коротких особистих повідомлень. Завдяки цьому користувач стикається з правдоподібними прикладами атак, що підвищує ефективність тренування.

Особливу увагу приділено реалістичності сценаріїв, приклади яких наведено у статті[2]. Модель імітує не лише структуру повідомлень, але й характерні помилки зловмисників: підроблені адреси відправників, незначні стилістичні відхилення, емоційно маніпулятивні фрази. Таким чином, користувач вчиться помічати деталі, які часто ігноруються у реальному житті.

Модуль аналізу реалізований за допомогою комбінації класичних і сучасних алгоритмів обробки тексту: Word2Vec, Doc2Vec, GloVe, а також байєсівського класифікатора SpamBayes. Вони дозволяють автоматично визначати потенційно небезпечні елементи, такі як аномалії лексики, підозрілі посилання чи приховані елементи HTML, що можуть бути використані для експлуатації вразливостей.

Аналіз здійснюється локально, без підключення до зовнішніх серверів. Це забезпечує конфіденційність персональних даних і виключає можливість витоку інформації, що робить систему придатною для використання в освітніх та корпоративних середовищах.

На рисунку 1 подано загальну схему роботи системи, що демонструє послідовність основних етапів – від створення листа до аналізу та оцінки безпеки.

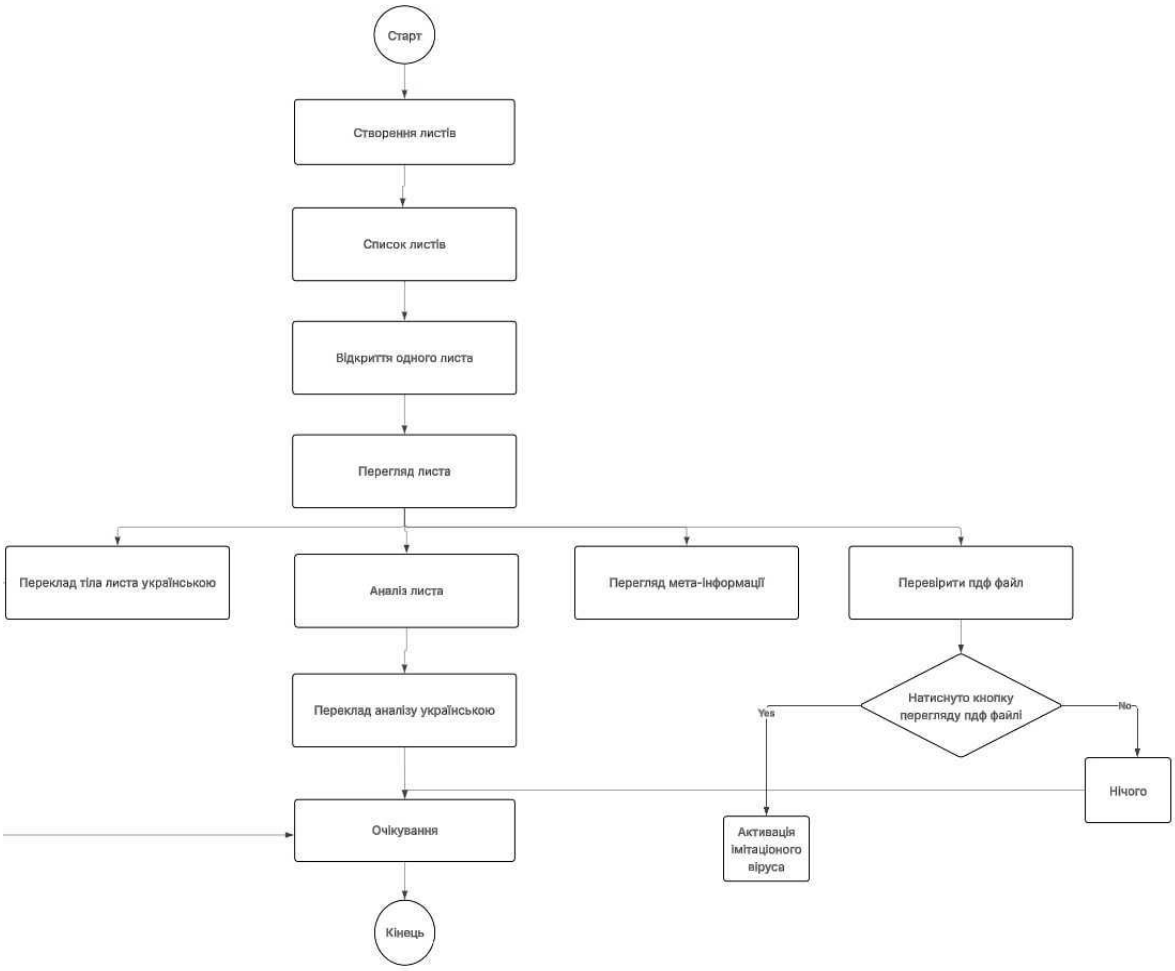

Рисунок 1 – Блок-схема інтерфейсу демонстрації:

Блок-схема інтерфейсу демонстрації: послідовність створення листів, перегляду списку та одного листа, дій користувача (переклад, аналіз, перегляд метаданих, перевірка PDF і активація імітаційного вірусу)

Модуль оцінювання безпеки відображає результати аналізу та надає зворотний зв'язок користувачу. Після кожного тренування програма формує статистику: кількість правильно ідентифікованих загроз, помилки та рекомендації щодо покращення уваги до деталей. Для реалізації інтерфейсу використано Python і бібліотеку Tkinter, що забезпечує кросплатформну сумісність і простоту використання. Інтерфейс інтуїтивний, з поділом на вкладки: «Пошта», «Аналіз», «Результати». Кожен лист супроводжується кнопкою для позначення як безпечного або небезпечного, після чого система миттєво відображає оцінку рішення з коротким поясненням.

2. Тестування та результати

Для оцінювання ефективності застосунку проведено експеримент серед студентів спеціальності «Кібербезпека» Національного університету «Одеська політехніка». Учасники проходили серію тренувальних сеансів, під час яких отримували згенеровані приклади листів із різним рівнем складності. Уже після трьох сесій середній рівень точності розпізнавання небезпечних повідомлень підвищився на 63 %, що підтвердило навчальну ефективність розробленої системи.

Отримані результати засвідчили, що користувачі почали приділяти більшу увагу дрібним деталям у тексті листів, таким як доменна адреса, стиль написання, логотипи або формулювання термінових закликів. Це вказує на підвищення рівня обізнаності й критичного мислення при взаємодії з електронними повідомленнями, що є ключовою метою навчання.

Крім того, учасники відзначили зручність інтерфейсу та високу реалістичність змодельованих сценаріїв, що сприяло зануренню в процес і підвищенню мотивації до навчання.

Подальший розвиток системи передбачає розширення функціональності додаванням мультимовної підтримки, інтеграцією з корпоративними поштовими клієнтами та впровадження адаптивного рівня складності завдань залежно від результатів користувача. Також планується створення адміністративної панелі для викладачів або керівників, які зможуть відстежувати прогрес групи в режимі реального часу.

Таким чином, розроблений застосунок може бути ефективним інструментом не лише для освітніх закладів, а й для корпоративних тренінгів із підвищення кіберграмотності працівників

Висновок. Розроблений застосунок є ефективним засобом інтерактивного навчання безпечній роботі з електронною поштою. Поєднання методів NLP, машинного навчання та генеративних мовних моделей дозволяє створити сучасну систему, що моделює реальні сценарії атак і формує стійкі навички протидії фішинговим і вірусним загрозам.

Перелік використаних джерел.

1. Newman M. E. J., Forrest S., Balthrop J. Email Networks and the Spread of Computer Viruses. Physical Review E. 2002. Vol. 66, No 3. P. 035101. DOI: 10.1103/PhysRevE.66.035101.

2. Singh P., Maravi Y. P. S., Sharma S. Phishing Websites Detection Through Supervised Learning Networks. International Conference on Computing and Communications Technologies (ICCCT). 2015. P. 61–65. DOI: 10.1109/ICCCT2.2015.7292720.