Секція 1

МЕТОД АНАЛІЗУ ДОСТОВІРНОСТІ ТЕКСТОВИХ ПОВІДОМЛЕНЬ ДЛЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ КІБЕРБЕЗПЕКИ

Шевчук П.О.
Хмельницький національний університет
Науковий керівник: Мазурець О.В.

Актуальність. Виявлення достовірності даних є важливою задачею в сфері кібербезпеки, адже в останній час в мережі Інтернет має місце розповсюдження неправдивої інформації [1], тому розробка прикладних рішень, які здатні перевіряти достовірність текстів є актуальною задачею розвитку інтелектуальних систем кібербезпеки [2].

Мета полягає в розробці методу аналізу достовірності текстових повідомлень для інтелектуальних систем кібербезпеки.

Основні положення. У дослідженні розроблено метод аналізу достовірності текстових повідомлень для інтелектуальних систем кібербезпеки на основі ансамблевого підходу, що поєднує логістичну регресію, дерево рішень, градієнтне посилення та випадковий ліс.

У якості вхідних даних представлені текст для аналізу та навчений ансамбль класифікаторів. Першим кроком система визначає мову тексту та при необхідності автоматично перекладає на англійську. Другим кроком увесь текст приводиться до нижнього регістру та видаляються стоп-символи. Отриманий результат проходить етапи токенізації, лематизації та векторизації. На третьому етапі дані аналізуються на наявність логістичної регресії та визначається дерево рішень. Окрім цього проводиться градієнтний бустінг та визначення випадкового лісу. Четвертим кроком формуються результати оцінки у вигляді зваженої оцінки. На виході отримується оцінка на принадлежність тексту для аналізу до категорії достовірності тексту.

Поданий метод був апробований шляхом розробки програмного забезпечення та продемонстрував високу ефективність, використовуючи зважений показник достовірності, обчислений на основі виходів кожної моделі.

Висновки. Запропонований метод аналізу достовірності текстових повідомлень для інтелектуальних систем кібербезпеки на основі ансамблевого підходу показав ефективність на рівні 92%, що є високим показником порівняно із сучасними рішеннями у сфері кібербезпеки.

Реалізація тестового програмного забезпечення була виконана у вигляді веб-додатка за допомогою технологій Scikit-Learn та Flask. На основі ансамблевих моделей було сформовано зважений показник достовірності тексту, який обчислюється як сума впливових коефіцієнтів кожної моделі, помножених на вихід відповідної моделі класифікатора.

Для навчання класифікаторів використовувався збалансований англомовний набір даних, що складався з 44 898 зразків, зокрема 23 481 зразків фейкової інформації та 21 417 зразків дійсної інформації.

Список літератури

1. Дослідження на поширення використання українцями соцмереж. *Детектор Медіа.* URL: https://detector.media/infospace/article/213998/2023-07-10-opora-osnovnym-dzherelom-informatsii-mayzhe-80-ukraintsiv-ie-sotsialni-merezhi (дата звернення 11.11.24).
2. Звідки українці беруть інформацію в умовах війни? *Українська Правда.* URL: https://life.pravda.com.ua/society/2022/06/2/248923 (дата звернення 11.11.24).

Відомості про авторів

Шевчук Павло Олександрович, студент кафедри комп'ютерних наук, Хмельницький національний університет, shevchuk12072005@gmail.com
Мазурець Олександр Вікторович, доцент кафедри комп'ютерних наук, Хмельницький національний університет, exechong@gmail.com