УДК 004.8

В. Д. Яловінський

РОЗРОБКА СИСТЕМИ АНАЛІЗУ СОЦІАЛЬНИХ МЕРЕЖ ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН

Вінницький національний технічний університет;

Анотація

Досліджено теоретичні аспекти створення систем для автоматизованого аналізу контенту в соціальних мережах. Запропоновано методи виявлення фейкових новин, які ґрунтуються на застосуванні алгоритмів машинного навчання, обробки текстових і мультимедійних даних, а також графового аналізу. Наведено можливі архітектурні рішення та обґрунтовано доцільність їх використання.

Ключові слова: фейкові новини, соціальні мережі, графовий аналіз, машинне навчання, мультимедійний аналіз.

Abstract

The theoretical aspects of creating systems for automated analysis of content on social networks are investigated. Methods for detecting fake news based on the use of machine learning algorithms, text and multimedia data processing, and graph analysis are proposed. Possible architectural solutions are presented and the feasibility of their use is substantiated.

Keywords: fake news, social networks, graph analysis, machine learning, multimedia analysis.

Розробка системи аналізу соціальних мереж для виявлення фейкових новин базується на інтеграції сучасних технологій та алгоритмів. Ключовою особливістю такої системи є її здатність працювати з великими обсягами даних у реальному часі, забезпечуючи аналіз текстового, мультимедійного контенту і соціальних взаємозв'язків.

Першим етапом функціонування системи є автоматизований збір даних із соціальних мереж. Для цього використовується технологія API соціальних платформ, яка дозволяє отримувати доступ до відкритих постів, коментарів, медіафайлів і метаданих [1]. Якщо API недоступний або має обмеження, можна застосовувати методи веб-скрапінгу, наприклад, бібліотеки BeautifulSoup чи Scrapy, які забезпечують збір інформації із соціальних платформ. Для потокової обробки зібраних даних використовується Apache Kafka, що забезпечує передачу даних між модулями системи та дозволяє працювати з потоками в реальному часі [2].

Обробка текстового контенту реалізована із застосуванням моделей обробки природної мови (NLP). На цьому етапі текст проходить через процедури попередньої обробки, які включають токенізацію, видалення стоп-слів, лематизацію та визначення ключових слів. Використання моделей машинного навчання, таких як BERT або GPT, дозволяє класифікувати повідомлення за категоріями, виявляючи маніпулятивний або неправдивий контент. Додаткові моделі аналізу емоційного забарвлення тексту можуть використовуватись для виявлення тенденцій до поляризації або підбурювання [3].

Для мультимедійного контенту використовується модуль, побудований на основі бібліотеки OpenCV та моделей глибокого навчання, таких як Convolutional Neural Networks (CNN). Він виконує аналіз зображень і відео, спрямований на виявлення маніпуляцій, включаючи deepfake [4]. Процедура аналізу включає детекцію обличчя, оцінку змінних характеристик та розпізнавання текстур, що є характерними для підроблених файлів. Обробка мультимедійного контенту проводиться паралельно із текстовим аналізом, що дозволяє обробляти публікації комплексно.

Графовий аналіз є невід'ємною складовою системи, оскільки взаємозв'язки між користувачами мереж створюють механізми для виявлення бот-мереж і кампаній дезінформації [5]. Соціальний граф моделюється за допомогою графових алгоритмів, таких як алгоритм PageRank або Label Propagation, що забезпечують оцінку впливу окремих вузлів та груп у мережі. Для візуалізації та роботи із графовими даними використовується бібліотека NetworkX або спеціалізовані інструменти, як Gephi.

Система використовує мову програмування Python як основний інструмент для розробки та реалізації алгоритмів. Платформи TensorFlow і PyTorch забезпечують можливість створення та навчання моделей машинного навчання, що використовуються для класифікації контенту та аналізу медіафайлів [6]. Архітектура системи дозволяє масштабування за рахунок використання контейнеризації (Docker) і розподілених обчислень, наприклад, у хмарному середовищі.

Ефективність системи забезпечується за рахунок адаптації моделей до специфіки аналізованого контенту. Наприклад, для конкретної соціальної мережі можуть бути налаштовані унікальні параметри NLP-моделей або додані додаткові моделі, спрямовані на ідентифікацію регіональних аномалій. Постійне навчання на оновлених наборах даних дозволяє системі швидко адаптуватися до нових викликів, включаючи розробку нових схем маніпуляції та дезінформації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Adiwardana D., Luong M., Socher R. Towards Better NLP: State-of-the-Art Language Models – 2020 – 45 с.
2. Brown T., Mann B., Ryder N. Language Models Are Few-Shot Learners – OpenAI, 2020 – 87 с.
3. Devlin J., Chang M., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding – 2019 – 14 с.
4. Jurafsky D., Martin J. Speech and Language Processing (3rd Edition) – Prentice Hall, 2022 – 1024 с.
5. Kaggle – Platform for Data Science Competitions [Електронний ресурс] – Режим доступу до ресурсу: https://kaggle.com
6. HubSpot Blog on Fake News Analysis [Електронний ресурс] – Режим доступу до ресурсу: https://blog.hubspot.com

Яловінський Віталій Дмитрович – студент групи 2ПІ-23м, факультету інформаційних технологій та комп'ютерної інженерії Вінницький національний технічний університет, Вінниця, vitalikvega@gmail.com.

Yalovinskyi V. D. – student of group 2PI-23m, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, vitalikvega@gmail.com.