

УДК 004.85:327.8

## **МОДЕЛЬ МАШИННОГО НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВОЇ СКЛАДОВОЇ У НОВИННИХ ТЕКСТАХ**

Панкратов Є. О.

Науковий керівник – доц. Чала Л. Е.

Харківський національний університет радіоелектроніки, каф. ШІ  
м. Харків, Україна

тел.: +38(095) 195-31-62, e-mail: yevhenii.pankratov@nure.ua

In today's world, information technologies play an increasingly important role in people's lives. One of the most important tasks in the field of information technology is the fight against fake information.

Now, when a variety of news can be easily distributed through social networks and Internet portals, the problem of detecting fake news is becoming more and more urgent. The rapid spread of these news can lead to serious consequences for society, so there is a need to create effective methods for detecting fake components in news texts.

The purpose of this study is to create a machine learning model for detecting fake content in news texts.

На сьогоднішній день існує багато методів виявлення фейкових новин, але більшість з них базуються на ручному аналізі тексту, що уповільнює процес виявлення та не дає можливості його автоматизувати. Одним з найбільш перспективних методів є використання моделі машинного навчання, яка може бути впроваджена в системи моніторингу новин та інформаційні портали для автоматичного виявлення фейкових новин у текстах. Така модель може суттєво зменшити поширення фейкових новин, підвищити якість інформації для потенційних користувачів та зміцнити довіру до засобів масової інформації.

Таким чином, доцільним є проведення досліджень щодо розвитку наукових підходів до аналізу новинних текстів та етичних питань, пов'язаних з застосуванням технологій машинного навчання в цій сфері.

У доповіді пропонується підхід до використання методів машинного навчання для створення моделі, яка буде автоматично виявляти фейкову складову у новинних текстах. В цій моделі використовується алгоритм класифікації текстів на основі нейронних мереж зі зворотнім поширенням помилки.

Для побудови моделі машинного навчання використано набір даних, який містить реальні та фейкові новинні тексти. Початковим етапом практичного застосування моделі є попередня обробка даних, що передбачає необхідність проведення токенизації, видалення стоп-слів та лематизації.

Нейромережева архітектура пропонованої моделі містить вхідний, прихований та вихідний шари. Для прихованого шару моделі

використовується функція активації *ReLU*, а для вихідного шару – функція активації *sigmoid*. Як функції втрат використовується бінарна перехресна ентропія:

$$h_j = \text{ReLU} \left( \sum_{i=1}^n w_{ij} x_i + b_j \right).$$

Навчання моделі з використанням алгоритму зворотного поширення помилки передбачало її тренування на навчальному наборі даних та оцінювання її ефективності на тестовому наборі даних.

Ефективність моделі оцінювалася на основі метрики точності (accuracy):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

де *TP* – кількість правильно класифікованих позитивних зразків; *TN* – кількість правильно класифікованих негативних зразків; *FP* – кількість неправильно класифікованих позитивних зразків, *FN* – кількість неправильно класифікованих негативних зразків.

Після навчання моделі та її перевірки на тестових даних точність виявлення фейкових новин складала від 80% до 95% в залежності від конфігурації нейронної мережі та параметрів векторизації тексту.

Таким чином, результати тестування запропонованої моделі машинного навчання підтвердили її працездатність та високу точність виявлення фейкових складових у новинних текстах. Результати цієї роботи можуть бути використані для боротьби з поширенням фейкових новин в інтернеті.

Доцільним слід вважати подальший розвиток проведених досліджень, спрямований на удосконалення розробленої моделі з використанням додаткових механізмів машинного навчання та аналізу актуальних наукових напрямків в галузі автоматизованого виявлення фейкових новин.

Список використаних джерел:

1. Kumar N., Aggarwal N., Kumar P., Yadav. V. (2020). Detection of Fake News in Social Media Networks. 2020 IEEE 6th International Conference for Convergence in Technology (I2CT). 545-549.
2. Zhang, W., Gao, Y., Fan, Y. (2019). Fake News Detection on Social Media: A Data Mining Perspective. ACM Transactions on Management Information Systems (TMIS), 1-22.