

**Медведєв Д.Г.**

<https://orcid.org/0000-0002-3747-1717>

Державний університет економіки і технологій

**Хоцкіна В.Б.**

<https://orcid.org/0000-0001-8963-4189>

Державний університет економіки і технологій

**Лисункін Д.Д.**

<https://orcid.org/0009-0004-9515-3380>

Державний університет економіки і технологій

## ДЕТЕКЦІЯ ФЕЙКОВИХ НОВИН З ВИКОРИСТАННЯМ NLP ТА ВІЗУАЛІЗАЦІЯ ЇЇ РЕЗУЛЬТАТІВ

Сучасний інформаційний простір характеризується стрімким зростанням обсягів цифрового контенту та високою швидкістю його поширення через мережу Інтернет і соціальні медіа. У таких умовах значно зростає ризик розповсюдження недостовірної інформації, зокрема фейкових новин, які можуть впливати на суспільну думку, політичні процеси, економічну ситуацію та інформаційну безпеку держави. Виявлення та протидія дезінформації є актуальною науково-практичною задачею, що потребує застосування сучасних методів обробки даних.

У даній статті розглянуто підхід до створення програмного забезпечення для автоматичної детекції фейкових новин із використанням методів обробки природної мови (NLP) та машинного навчання. Запропонована система базується на модульній архітектурі, що включає етапи попередньої обробки тексту, векторизації даних за допомогою TF-IDF, класифікації текстів із використанням алгоритму логістичної регресії та подальшої візуалізації результатів із застосуванням методу головних компонент (PCA).

У статті детально описано процес підготовки текстових даних, включаючи очищення, токенизацію та видалення стоп-слів. Особливу увагу приділено перетворенню тексту у числове представлення, що дозволяє ефективно застосовувати алгоритми машинного навчання. Проведено експериментальне дослідження на корпусі новинних текстів, яке показало, що запропонований підхід забезпечує точність класифікації понад 85 %.

Результати дослідження підтверджують ефективність використання класичних методів машинного навчання для задачі детекції фейкових новин. Запропоноване програмне забезпечення може бути використане у системах моніторингу інформаційного простору, новинних агрегаторах та дослідницьких проєктах. Також визначено перспективи розвитку системи, зокрема використання трансформерних моделей та інтеграція додаткових джерел даних.

**Ключові слова:** фейкові новини, машинне навчання, NLP, класифікація текстів, TF-IDF, логістична регресія.

**Постановка проблеми.** Поширення фейкових новин є однією з ключових проблем сучасного цифрового суспільства. Завдяки соціальним мережам та онлайн-платформам інформація розповсюджується миттєво, що значно ускладнює її перевірку. Внаслідок цього користувачі можуть отримувати недостовірні або маніпулятивні повідомлення, які впливають на їхні рішення та поведінку.

Автоматизація процесу перевірки достовірності новин дозволяє мінімізувати вплив людського фактора та підвищити ефективність ана-

лізу великих обсягів текстових даних. Для цього застосовуються методи машинного навчання, що здатні визначати закономірності в текстах та класифікувати їх за визначеними категоріями.

**Аналіз останніх досліджень і публікацій.** Існуючі підходи до виявлення фейкових новин можна поділити на кілька основних груп: лінгвістичні, статистичні та методи машинного навчання [4].

Лінгвістичні підходи базуються на аналізі стилістичних, синтаксичних та семантичних характеристик тексту. Зокрема, досліджуються емоційна

забарвленість, використання оціночної лексики, частота займенників та інші мовні ознаки, характерні для маніпулятивного контенту [1]. Однак такі методи мають обмеження через залежність від мови та предметної області.

Статистичні методи передбачають використання кількісних характеристик тексту, таких як частота слів, довжина речень та інші метрики. Вони є відносно простими у реалізації, але не враховують глибинний контекст повідомлення [5].

Сучасні дослідження приділяють значну увагу методам машинного навчання, таким як найвний байєсівський класифікатор, метод опорних векторів, логістична регресія та ансамблеві методи [2]. Вони дозволяють досягати високої точності за умови наявності якісно розмічених даних.

Останнім часом особливу популярність набули глибокі нейронні мережі та трансформерні моделі, такі як BERT, які здатні враховувати контекст слів у реченні та значно підвищують якість класифікації [4]. Також досліджуються підходи, що враховують соціальні сигнали, метадані та поведінкові характеристики користувачів.

Таким чином, сучасні дослідження демонструють перехід від простих статистичних методів до складних гібридних моделей, що поєднують NLP, машинне навчання та аналіз контексту [3].

**Постановка завдання.** Метою даного дослідження є розробка програмного забезпечення для автоматичної детекції фейкових новин із використанням методів обробки природної мови та машинного навчання.

Для досягнення поставленої мети необхідно вирішити такі задачі:

1. Проаналізувати існуючі підходи до класифікації текстових даних.
2. Реалізувати модуль попередньої обробки текстової інформації.
3. Виконати перетворення текстів у числове представлення.
4. Реалізувати алгоритм класифікації на основі машинного навчання.
5. Забезпечити можливість візуалізації результатів класифікації.
6. Провести експериментальну перевірку ефективності системи.

**Виклад основного матеріалу. Архітектура програмного забезпечення.** Розроблене програмне забезпечення має модульну структуру, що забезпечує гнучкість та можливість подальшого розширення функціоналу.

Загальна архітектура системи представлена на рис. 1.

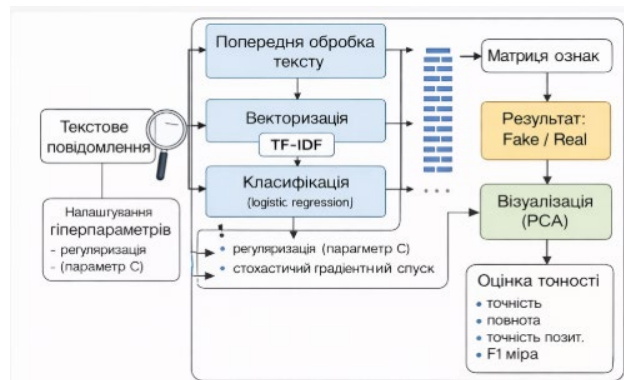


Рис. 1. Архітектура програмного забезпечення для детекції фейкових новин

Система складається з таких основних модулів:

- модуль попередньої обробки тексту;
- модуль векторизації;
- модуль класифікації;
- модуль візуалізації результатів.

**Модуль попередньої обробки тексту.** Попередня обробка тексту є важливим етапом підготовки даних для подальшого аналізу. Вона дозволяє зменшити шум у даних та підвищити якість навчання моделі.

До основних етапів попередньої обробки належать:

- перетворення тексту до нижнього регістру;
- видалення спеціальних символів;
- видалення стоп-слів;
- токенізація;
- лематизація (за потреби).

Після виконання цих операцій текст стає придатним для подальшого перетворення у числове представлення.

**Модуль векторизації.** Для перетворення текстових даних у числову форму використовується метод TF-IDF (Term Frequency – Inverse Document Frequency).

Метод TF-IDF дозволяє оцінити важливість кожного слова у тексті відносно всього корпусу документів. Значення TF-IDF визначається як добуток двох показників:

- TF – частота терміна в документі;
  - IDF – обернена частота документа.
- Формально:

$$TF\text{-}IDF(t, d) = TF(t, d) \times IDF(t)$$

де: t – термін;  
d – документ.

Використання TF-IDF дозволяє виділити найбільш інформативні слова та зменшити вплив поширених термінів.

**Модуль класифікації.** Для класифікації текстів використано алгоритм логістичної регресії. Даний метод є одним із базових алгоритмів бінарної класифікації та забезпечує достатньо високу точність при невеликих обчислювальних витратах.

Логістична регресія оцінює ймовірність належності тексту до одного з класів — «fake» або «real». Функція прийняття рішення базується на сигмоїдній функції:

$$P(y=1|x) = 1 / (1 + e^{-z})$$

де  $z$  – лінійна комбінація ознак.

Перевагами даного методу є:

- інтерпретованість результатів;
- швидкість навчання;
- стійкість до перенавчання при достатній кількості даних.

**Модуль візуалізації.** Для підвищення наочності результатів класифікації реалізовано модуль візуалізації з використанням методу головних компонент (PCA).

Метод PCA дозволяє зменшити розмірність простору ознак та відобразити дані у двовимірному або тривимірному просторі.

Результат роботи модуля візуалізації представлено на рис. 2 та рис. 3.



Рис. 2. Блок-схема алгоритму детекції фейкових новин

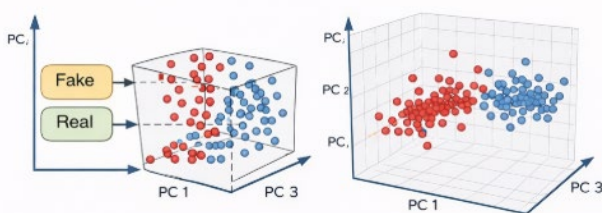


Рис. 3. Візуалізація результатів класифікації після зниження розмірності

Візуалізація дозволяє оцінити розподіл текстів у просторі ознак та проаналізувати якість розділення класів.

**Реалізація програмного забезпечення.** Розроблене програмне забезпечення реалізовано мовою програмування Python із використанням сучасних бібліотек для обробки текстових даних, машин-

ного навчання та комп'ютерної графіки. Основним інструментом для побудови моделі класифікації було обрано бібліотеку Scikit-learn, яка надає зручні засоби для векторизації тексту, навчання моделей та оцінювання їхньої ефективності.

Для попередньої обробки текстових даних використовуються стандартні методи очищення тексту, зокрема видалення HTML-тегів, спеціальних символів, посилань та приведення тексту до нижнього регістру. Після цього виконується токенізація тексту та видалення стоп-слів, що дозволяє зменшити кількість несуттєвих ознак.

Перетворення текстових даних у числове представлення здійснюється за допомогою методу TF-IDF. У результаті кожен текст представлений у вигляді вектору ознак, що характеризує важливість окремих слів у документі.

Для класифікації текстів використовується алгоритм логістичної регресії. Навчена модель дозволяє визначити ймовірність того, що текст належить до класу «fake» або «real». Після навчання модель зберігається у вигляді файлу та може бути використана для аналізу нових текстових повідомлень.

Для візуалізації результатів класифікації використовується метод зниження розмірності PCA. Отримані координати документів у двовимірному або тривимірному просторі передаються до модуля візуалізації, який реалізовано з використанням бібліотек ModernGL та GLFW. Це дозволяє відображати результати класифікації у вигляді точок у просторі, де кожна точка відповідає окремому тексту новини.

**Експериментальні результати.** Для оцінки ефективності розробленого програмного забезпечення було проведено експериментальні дослідження на корпусі новинних текстів, що містив повідомлення двох категорій: достовірні новини (real) та фейкові новини (fake).

Перед навчанням моделі дані було розділено на навчальну та тестову вибірки у співвідношенні 80 % до 20 %. Навчальна вибірка використовувалася для побудови моделі логістичної регресії, а тестова – для оцінювання якості класифікації.

Оцінювання проводилося за такими метриками:

- точність (accuracy);
- повнота (recall);
- точність позитивних передбачень (precision);
- F1-міра.

Отримані результати свідчать про достатньо високу ефективність застосованого підходу. Значення точності класифікації перевищує 0,85, що

підтверджує доцільність використання методу TF-IDF у поєднанні з логістичною регресією для задачі детекції фейкових новин.

Аналіз матриці помилок показав, що більшість хибних спрацювань виникає у випадках текстів зі змішаними ознаками або нейтральним стилем викладення. Це свідчить про необхідність подальшого вдосконалення моделі шляхом використання більш складних методів обробки контексту.

**Приклад роботи системи.** Для демонстрації ефективності розробленого програмного забезпечення було проведено тестування системи на реальних текстових повідомленнях. Метою експерименту було перевірити здатність моделі коректно визначати достовірність новинних повідомлень.

Розглянемо приклад обробки новинного тексту системою.

Приклад 1.

Новина: “Nasa confirms discovery of water traces on Mars based on new satellite data.”

Після виконання етапів попередньої обробки тексту та векторизації за допомогою методу TF-IDF модель класифікації виконала аналіз повідомлення. У результаті система визначила новину як достовірну.

Результат класифікації: Real (достовірна новина).

Отриманий результат пояснюється тим, що текст має нейтральний стиль викладення, містить посилання на наукові джерела та не має ознак емоційної маніпуляції.

Приклад 2.

Новина: “Scientists proved that drinking coffee cures all diseases instantly.”

Після обробки тексту системою було отримано наступний результат:

Результат класифікації: Fake (фейкова новина).

Алгоритм визначив новину як фейкову через використання перебільшених тверджень, відсутність посилань на наукові джерела та наявність маніпулятивних формулювань. Такі лінгвістичні особливості є характерними для дезінформаційних повідомлень.

Отримані результати демонструють, що запропонований підхід дозволяє ефективно визначати фейкові новини на основі аналізу текстових характеристик. Використання методів машинного навчання разом із попередньою обробкою тексту забезпечує достатній рівень точності для практичного застосування системи.

**Аналіз результатів роботи системи.** Проведені експериментальні дослідження показали, що запропонована система здатна ефективно

виконувати класифікацію новинних повідомлень на достовірні та фейкові. Використання методу TF-IDF дозволяє виділяти ключові слова та фрази, які характерні для різних типів новин.

Аналіз результатів класифікації показав, що модель демонструє стабільні результати при роботі з текстами різної тематики. Найбільш точні результати отримуються для текстів, які мають яскраво виражені лінгвістичні ознаки маніпуляції, зокрема емоційно забарвлену лексику, перебільшені твердження або відсутність джерел інформації.

У процесі експериментів було встановлено, що деякі помилки класифікації виникають у випадках, коли текст містить змішані ознаки або нейтральний стиль викладення. Наприклад, новини з наукової тематики або короткі інформаційні повідомлення можуть мати подібну структуру, що ускладнює їх класифікацію.

Незважаючи на це, результати дослідження підтверджують ефективність використання класичних методів машинного навчання для задачі детекції фейкових новин. Поєднання попередньої обробки тексту, TF-IDF векторизації та логістичної регресії забезпечує достатній рівень точності для практичного використання системи.

**Порівняльний аналіз методів.** Для підтвердження ефективності обраного підходу було проведено порівняльний аналіз із альтернативними алгоритмами класифікації, зокрема:

- наївним байєсівським класифікатором;
- методом опорних векторів;
- k-найближчих сусідів.

Результати експериментів показали, що логістична регресія забезпечує збалансоване співвідношення між точністю та швидкістю. Метод опорних векторів продемонстрував дещо вищу точність, однак потребує більших обчислювальних ресурсів. Наївний байєсівський класифікатор характеризується високою швидкістю роботи, проте має нижчу точність при роботі з великим словниковим простором.

Таким чином, логістична регресія є оптимальним вибором для реалізації програмного забезпечення з урахуванням обмежень ресурсів та вимог до швидкодії.

**Практичне застосування системи.** Розроблене програмне забезпечення може бути використане:

- у системах моніторингу інформаційного простору;
- у новинних агрегаторах;
- у дослідницьких проєктах з аналізу медіа-контенту;

– у навчальному процесі для демонстрації методів NLP.

Завдяки модульній архітектурі система може бути розширена шляхом інтеграції нейронних мереж або попередньо навчених мовних моделей.

**Перспективи розвитку системи.** Розроблена система детекції фейкових новин може бути вдосконалена у кількох напрямках, що дозволить підвищити точність класифікації та розширити функціональні можливості програмного забезпечення.

Одним із перспективних напрямів є використання більш складних моделей обробки природної мови, зокрема нейронних мереж та трансформерних архітектур. Сучасні моделі, такі як BERT або DistilBERT, здатні враховувати контекст слів у реченні та аналізувати складні семантичні залежності між ними. Це дозволяє більш ефективно визначати маніпулятивні або недостовірні повідомлення, які можуть бути складними для класифікації за допомогою класичних статистичних методів.

Іншим напрямом розвитку є інтеграція додаткових джерел інформації для аналізу новин. Наприклад, система може враховувати метадані повідомлень, такі як джерело новини, дата публікації або кількість поширень у соціальних мережах. Аналіз таких характеристик може допомогти більш точно оцінювати достовірність інформації.

Також перспективним є розширення можливостей модуля візуалізації. У поточній версії системи результати класифікації відображаються у вигляді точок у двовимірному або тривимірному просторі після зниження розмірності. У подальшому можна реалізувати більш інтерактивні інструменти аналізу, зокрема динамічні фільтри, підсвічування кластерів або можливість детального перегляду текстів, що знаходяться у певних областях простору ознак.

Крім того, система може бути інтегрована з веб-платформами або новинними агрегаторами для автоматичного аналізу новин у режимі реального часу. Такий підхід дозволить використовувати розроблене програмне забезпечення для моніторингу інформаційного простору та швидкого виявлення потенційно недостовірних повідомлень.

Таким чином, подальший розвиток системи може бути спрямований на підвищення точності алгоритмів, розширення функціональності програмного забезпечення та інтеграцію з сучасними інформаційними платформами. Це дозволить зро-

бити систему більш ефективним інструментом для аналізу новинних повідомлень та боротьби з поширенням дезінформації.

**Висновки.** У результаті виконаної роботи було розроблено програмне забезпечення для автоматичної детекції фейкових новин із використанням методів обробки природної мови та машинного навчання.

Основні результати роботи:

1. Встановлено, що для автономної детекції найефективнішим є поєднання статистичного аналізу тексту (TF-IDF) та лінійних класифікаторів. Сформульовано вимоги до швидкодії та інтерпретованості системи.

2. Розроблено клас TextPreprocessor, який виконує очищення від HTML-тегів та URL, нормалізацію, токенизацію та фільтрацію стоп-слів. Забезпечено параметризацію під українську, англійську та російську мови.

3. Створено пайплайн на основі векторизації TF-IDF (з використанням уніграм та біграм) та логістичної регресії з балансуванням класів. Модель забезпечує високу точність класифікації та підтримує збереження ваг для подальшого використання.

4. Розроблено консольний застосунок для виконання повного циклу роботи з системою: підготовки даних (preprocess), навчання моделі (train), оцінювання метрик (eval) та прогнозування (predict).

5. Програмно реалізовано алгоритм перетворення багатовимірних векторів текстів у 2D або 3D простір за допомогою методу головних компонент (PCA). Це дозволило сформувати набір метаданих (координати, клас, впевненість) для візуалізації.

6. Розроблено інтерактивний застосунок на основі бібліотек ModernGL та GLFW. Система дозволяє відображати тисячі новин у вигляді точок у 3D-просторі, фільтрувати їх за класами («fake»/«real») та аналізувати розподіл даних.

7. Створено віконний застосунок на базі Tkinter для зручного введення тексту новин, миттєвого отримання результату класифікації з індикатором впевненості та запуску візуалізації.

Отримані результати підтверджують можливість використання класичних алгоритмів машинного навчання для задачі детекції фейкових новин. Подальші дослідження можуть бути спрямовані на інтеграцію глибоких нейронних мереж та розширення функціоналу системи.

**Список літератури:**

1. Jurafsky D., Martin J. *Speech and Language Processing*. Pearson, 2023.
2. Bishop C. *Pattern Recognition and Machine Learning*. Springer, 2006.
3. Pedregosa F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2011.
4. Zhou X., Zafarani R. Fake News Detection: A Survey. *ACM Computing Surveys*, 2020.
5. Aggarwal C. *Machine Learning for Text*. Springer, 2018.
6. Devlin J. et al. BERT: Pre-training of Deep Bidirectional Transformers. 2019.
7. Mikolov T. et al. Efficient Estimation of Word Representations. 2013.
8. Manning C., Schütze H. *Foundations of Statistical NLP*. MIT Press, 1999.
9. Goldberg Y. *Neural Network Methods in NLP*. 2017.
10. Shu K. et al. Fake News Detection on Social Media. *SIGKDD*, 2017.

**Medvediev D.H., Khotskina V.B., Lysunkin D.D. FAKE NEWS DETECTION USING NLP AND VISUALIZATION OF ITS RESULTS**

*The modern information space is characterized by a rapid increase in the volume of digital content and the high speed at which it spreads via the Internet and social media. Under such conditions, the risk of spreading misinformation—particularly fake news—significantly increases, as this can influence public opinion, political processes, the economic situation, and national information security. Detecting and countering disinformation is a pressing scientific and practical challenge that requires the application of modern data processing methods.*

*This paper presents an approach to developing software for the automatic detection of fake news using natural language processing (NLP) and machine learning methods. The proposed system is based on a modular architecture that includes stages of text preprocessing, data vectorization using TF-IDF, text classification using a logistic regression algorithm, and subsequent visualization of results using the principal component analysis (PCA) method.*

*The paper describes in detail the process of preparing text data, including cleaning, tokenization, and stop-word removal. Particular attention is paid to converting text into a numerical representation, which allows for the effective application of machine learning algorithms. An experimental study was conducted on a corpus of news texts, which showed that the proposed approach achieves a classification accuracy of over 85%.*

*The results of the study confirm the effectiveness of classical machine learning methods for fake news detection tasks. The developed software can be applied in information monitoring systems, news aggregators, and research projects. Future work includes the use of transformer-based models and the integration of additional data sources.*

**Keywords:** *Fake news, machine learning, natural language processing, text classification, TF-IDF, logistic regression.*

Дата першого надходження статті до видання: 22.03.2026

Дата прийняття статті до друку після рецензування: 17.04.2026

Дата публікації (оприлюднення) статті: 19.05.2026