



Наукові перспективи
Видавнича група

№ 5 (59)

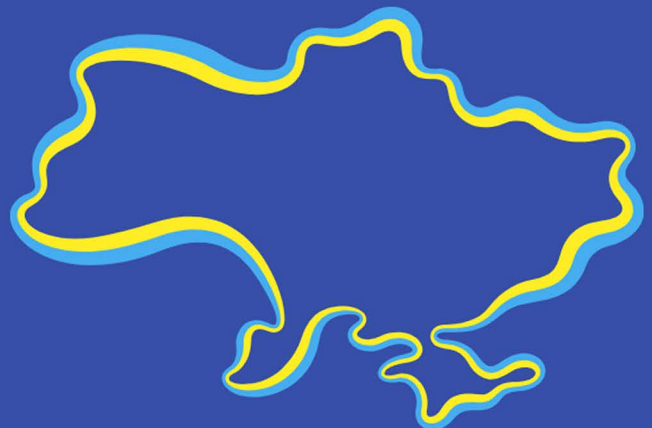
2026

ІТ НАУКА ТЕХНІКА

СЬОГОДНІ



З Україною
в серці!



Видавнича група «Наукові перспективи»

Всеукраїнська Асамблея докторів наук із державного управління

«Наука і техніка сьогодні»

Випуск № 5(59) 2026

Київ – 2026

Publishing Group «Scientific Perspectives»

Ukrainian Assembly of Doctors of Sciences in Public Administration

"Science and technology today"

Issue № 5(59) 2026

Kyiv – 2026

ISSN 2786-6025 Online

УДК 001.32:1 /3](477)(02)

R40-05553

DOI:  Crossref
we use DOIs

[https://doi.org/10.52058/2786-6025-2026-5\(59\)](https://doi.org/10.52058/2786-6025-2026-5(59))

**«Наука і техніка сьогодні» (Серія «Педагогіка», Серія «Право», Серія «Економіка»,
Серія «Фізико-математичні науки», Серія «Техніка»):
журнал. 2026. № 5(59) 2026. С. 6514**



*Згідно наказу Міністерства освіти і науки України від 07.04.2022 № 320 журналу
присвоєно категорію "Б" із економіки та педагогіки (спеціальності – 015 -
Педагогічні науки; 076 - Економічні науки)*

*Згідно наказу Міністерства освіти і науки України від 06.06.2022 № 530 журналу
присвоєно категорію "Б" із права (спеціальність – 081 Юридичні науки)*

*Згідно наказу Міністерства освіти і науки України від 10.10.2022 № 894 журналу присвоєно
категорію "Б" із техніки (спеціальність - 122 Комп'ютерні науки)*

*Журнал видається за підтримки Міждержавної гільдії інженерів консультантів, Інституту філософії та
соціології Національної Академії Наук Азербайджану (Баку, Азербайджан), громадської організації «Християнська
академія педагогічних наук України» та громадської організації «Всеукраїнська асоціація педагогів і психологів з
духовно-морального виховання»*

*Рекомендовано до видавництва Президією Всеукраїнської Асамблеї докторів наук з державного управління
(Рішення від 25.05.2026, № 9/3-1/26)*



Журнал включено до міжнародної наукометричної бази Index
Copernicus (IC), міжнародної пошукової системи Google Scholar та до
міжнародної наукометричної бази даних Research Bible

Згідно Порядку формування Переліку наукових фахових видань України, затвердженого наказом МОН
України від 15.01.2018 № 32, повнотекстовий доступ до наукових статей журналу представлений на платформі
«Наукова періодика України» в Національній бібліотеці України імені В.І. Вернадського НАН України та в
Національному репозитарії академічних текстів

Головний редактор:



Коренева Інна Миколаївна - доктор педагогічних наук, професор, декан факультету природничої і фізико-математичної освіти Глухівського національного педагогічного університету імені Олександра Довженка; професор кафедри теорії і методики викладання природничих дисциплін Глухівського національного педагогічного університету імені Олександра Довженка (Україна)

Редакційна колегія:

1. **Біляковська Ольга Орестівна** доктор педагогічних наук, професор, завідувачка кафедри загальної педагогіки та педагогіки вищої школи Львівського національного університету імені Івана Франка (Україна)
2. **Воровка Маргарита Іванівна** – докторка педагогічних наук, професорка, професорка кафедри освітології та педагогіки мистецтва Мелітопольського державного педагогічного університету імені Богдана Хмельницького (Україна)

- Краснюк М.Т., Кулаженко В.В., Савінков Н.Д.** 4865
ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕВРИСТИЧНИХ ТА МЕТАЕВРИСТИЧНИХ ТЕХНОЛОГІЙ МАТЕМАТИЧНОЇ ОПТИМІЗАЦІЇ ДЛЯ ПІДТРИМКИ АКСЕЛERAЦІЇ ІННОВАЦІЙНИХ БІЗНЕС-ПРОЄКТІВ ПІД ЧАС ПЕРЕХОДУ ВІД VUCA- ДО BANI-СВІТУ
- Криворучко О.В., Шкарупило В.В., Кулініч О.М., Шестак Я.І.** 4883
АДАПТИВНІ МЕТОДИ УПРАВЛІННЯ ВІРТУАЛІЗОВАНИМИ РЕСУРСАМИ В МЕРЕЖАХ ЗАКЛАДІВ ВИЩОЇ ОСВІТИ НА ОСНОВІ АРХІТЕКТУРИ SDN/NFV
- Крижанівська І.В., Ткачук А.Г., Пуховський Є.С., Біленький І.О.** 4898
ІНТЕЛЕКТУАЛЬНІ МЕХАТРОННІ СИСТЕМИ ПРОГНОЗНОГО ОБСЛУГОВУВАННЯ ПРОМИСЛОВИХ РОБОТІВ НА ОСНОВІ АНАЛІЗУ ПАРАМЕТРІВ ВІБРАЦІЇ, СТРУМУ ТА ТЕМПЕРАТУРИ
- Крилик Л.В.** 4913
ОСОБЛИВОСТІ ЗАСТОСУВАННЯ ЧИСЕЛЬНИХ МЕТОДІВ В ОБЧИСЛЕНІ ВИЗНАЧЕНИХ ІНТЕГРАЛІВ
- Крилик Л.В., Романець К.В.** 4926
РОЗРОБКА TELEGRAM-БОТА ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН ІЗ ВИКОРИСТАННЯМ NLP І МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ
- Кубицький В.О., Божок А.В.** 4941
ЗБЕРЕЖЕННЯ ІНФОРМАЦІЇ В ПРОМІЖНИХ ШАРАХ ГЛИБОКИХ МЕРЕЖ: ВПЛИВ АРХІТЕКТУРНИХ РІШЕНЬ НА ЯКІСТЬ БАГАТОРІВНЕВИХ ПРЕДСТАВЛЕНЬ ДЛЯ ЗАДАЧ ПОДІБНОСТІ
- Кукла О.С.** 4953
МЕТОДИ ЗАБЕЗПЕЧЕННЯ КІБЕРСТІЙКОСТІ ТА ВІДМОВОСТІЙКОСТІ ВЕБ-ІНФРАСТРУКТУРИ В УМОВАХ МАСОВАНИХ КІБЕРАТАК: ДОСВІД УКРАЇНСЬКИХ ХОСТИНГ-ПРОВАЙДЕРІВ
- Кучерук С.В.** 4965
ФОТОКАТАЛІТИЧНА ДЕСТРУКЦІЯ АНІОННИХ БАРВНИКІВ: МОДЕЛЮВАННЯ Й ОПТИМІЗАЦІЯ ПРОЦЕСІВ ІЗ ВИКОРИСТАННЯМ ШТУЧНОГО ІНТЕЛЕКТУ

Крилик Людмила Вікторівна кандидат технічних наук, доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, м. Вінниця, <https://orcid.org/0000-0001-6642-754X>

Романець Катерина Вікторівна бакалавриня факультету інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, м. Вінниця

РОЗРОБКА TELEGRAM-БОТА ДЛЯ ВИЯВЛЕННЯ ФЕЙКОВИХ НОВИН ІЗ ВИКОРИСТАННЯМ NLP І МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. У статті представлено розробку Telegram-бота для виявлення фейкових новин із використанням NLP і моделей штучного інтелекту. Метою розробки є розширення функціональних можливостей Telegram-бота за рахунок поєднання ML-класифікації новинного тексту, аналізу заголовків, профільного аналізу результату, обробки статей за URL, перевірки відкритих джерел та освітнього модуля з медіаграмотності. Розроблений бот має клієнт-серверну логіку роботи: взаємодія з користувачем відбувається через Telegram-інтерфейс, а серверна частина виконує попередню обробку тексту, визначення мови, переклад за потреби, класифікацію та формування відповіді. Для реалізації використано Python, фреймворк aiogram, методи TF-IDF і Bag-of-Words, а також моделі Logistic Regression та Multinomial Naive Bayes. Для навчання моделей застосовано відкритий набір даних Fake and Real News Dataset із платформи Kaggle, що містить понад 40 тис. англійських записів. Модель класифікації новин Logistic Regression + TF-IDF з діапазоном n-грам (1,2) показала $F1=0,9804$, а модель аналізу заголовків Multinomial Naive Bayes + Bag-of-Words з діапазоном n-грам (1,1) продемонструвала $F1=0,9382$. Порівняння з аналогами показало, що розроблений бот розширює наявні рішення за сімома функціональними характеристиками: підтримує окремий аналіз заголовків, обробляє статті за URL, формує профіль результату з показниками емоційності, ризику маніпуляції, конкретності фактів і наявності джерел, виконує пошук схожих матеріалів у відкритих джерелах із fallback-переходом до Google, містить освітній модуль, інтерактивне тестування користувача та двомовну взаємодію з мовною адаптацією. При цьому швидкість класифікації становить близько 1,08 мс для текстів новин і $0,03 \div 0,04$

мс для заголовків. Розробка не замінює професійний фактчекінг, але є інструментом для швидкої попередньої перевірки новинного контенту в середовищі Telegram, допомагаючи користувачу орієнтуватися в медіапросторі та звертати увагу на потенційно сумнівні повідомлення.

Ключові слова: фейкові новини; Telegram-бот; NLP; машинне навчання; TF-IDF; Bag-of-Words; клікбейт; медіаграмотність.

Krylik Lyudmila Viktorivna Candidate of Technical Sciences, Associate Professor, Associate Professor at the Department for Computer Science, Vinnytsia National Technical University, Vinnytsia, <https://orcid.org/0000-0001-6642-754X>

Romanets Kateryna Viktorivna Bachelor of the Department for Computer Science, Vinnytsia National Technical University, Vinnytsia

DEVELOPMENT OF A TELEGRAM BOT FOR DETECTING FAKE NEWS USING NLP AND ARTIFICIAL INTELLIGENCE MODELS

Abstract. This article presents the development of a Telegram bot for detecting fake news using NLP and artificial intelligence models. The aim of the development is to expand the functional capabilities of the Telegram bot by combining machine learning-based classification of news text, headline analysis, contextual analysis of results, processing of articles by URL, verification of open sources, and an educational module on media literacy. The developed bot operates on a client-server architecture: interaction with the user takes place via the Telegram interface, whilst the server-side component performs text pre-processing, language detection, translation where necessary, classification and response generation. Python, the aiogram framework, TF-IDF and Bag-of-Words methods, as well as Logistic Regression and Multinomial Naive Bayes models were used for implementation. The Fake and Real News Dataset from the Kaggle platform, containing over 40,000 English-language records, was used to train the models. The Logistic Regression + TF-IDF news classification model with an n-gram range of (1,2) achieved an F1 score of 0.9804, whilst the Multinomial Naive Bayes + Bag-of-Words headline analysis model with an n-gram range of (1,1) achieved an F1 score of 0.9382. A comparison with similar tools has shown that the developed bot extends existing solutions in seven key areas: it supports separate analysis of headlines, processes articles by URL, generates a results profile with indicators of emotionality, risk of manipulation, factual accuracy and the presence of sources, performs a search for similar materials in open sources with a fallback to Google, and includes an educational module, interactive user testing and bilingual interaction with language adaptation. The classification speed is approximately 1.08 ms for news texts and 0.03÷0.04 ms for

ISSN 2786-6025 Online

headlines. The development does not replace professional fact-checking, but serves as a tool for rapid preliminary verification of news content within the Telegram environment, helping users navigate the media landscape and draw attention to potentially dubious reports.

Keywords: fake news; Telegram bot; NLP; machine learning; TF-IDF; Bag-of-Words; clickbait; media literacy.

Постановка проблеми. Сучасний інформаційний простір перенасичений новинним контентом, який поширюється швидше, ніж проходить перевірку. У Telegram користувач часто отримує новини не як повноцінні журналістські матеріали, а як короткі дописи, емоційні заголовки, переслані повідомлення або посилання без достатнього контексту.

Під час такого поширення може втрачатися першоджерело, дата, автор або обставини події, тому повідомлення нерідко оцінюється за першим враженням, а не за змістом і доказами.

Проблема не обмежується очевидними фейками. Маніпулятивний контент може містити частину правдивої інформації, але подавати її через перебільшення, емоційний тиск або замовчування суттєвих деталей. Ручна верифікація таких повідомлень потребує часу й навичок роботи із джерелами, що ускладнює її виконання під час швидкого перегляду новинного рядка. Через це виникає потреба в інструменті, який працює безпосередньо в Telegram, забезпечує автоматизовану первинну оцінку новинного контенту.

Аналіз останніх досліджень і публікацій. У межах теми виявлення фейкових новин розглянуто наукові підходи на основі NLP і машинного навчання, експертні фактчекінгові платформи, механізми модерації соціальних мереж та спеціалізовані Telegram-боти. У роботі К. Shu та співавторів розглянуто проблему виявлення фейкових новин у соціальних медіа, особливості даних, ознак і методів класифікації тексту [1]. R. Oshikawa та співавтори систематизують NLP-підходи до автоматизованого аналізу фейкових новин, зокрема використання класичних моделей машинного навчання та нейронних методів [2].

Серед практичних рішень розглянуто StopFake [3], Snopes, PolitiFact, МедіаГвардія, Bazilik, а також Telegram-боти «Перевірка» від Gwara Media [4], StopFake Bot, Fake True Bot і News Checking Bot. Соціальні платформи, зокрема X і Facebook, використовують модерацію, алгоритмічну фільтрацію та позначення сумнівного контенту, проте ці механізми працюють у межах самих платформ і недоступні користувачу як окремий інструмент перевірки. Аналіз показав, що більшість розглянутих рішень орієнтовані на експертну перевірку, готову базу матеріалів або самостійну оцінку новини користувачем. Через це залишається потреба в універсальному Telegram-інструменті, який поєднує

автоматизовану первинну оцінку новинного контенту, роботу з різними типами вхідних даних і допоміжні засоби для самостійної перевірки інформації [5].

Отже, **метою статті** є розширення функціональних можливостей Telegram-бота для виявлення фейкових новин за рахунок поєднання ML-класифікації новинного тексту, аналізу заголовків, профільного аналізу результату, перевірки відкритих джерел та освітнього модуля з медіаграмотності в межах єдиного діалогового інтерфейсу.

Виклад основного матеріалу. Розроблювана система проєктується як Telegram-бот із клієнт-серверною архітектурою. Клієнтську частину формує месенджер Telegram, через який користувач надсилає текст новини, заголовок або посилання на статтю та отримує результат аналізу. Серверна частина приймає повідомлення, виконує попередню підготовку тексту, класифікацію, формує короткий результат і надає додаткову інформацію за запитом користувача. Взаємодія між частинами здійснюється через Telegram Bot API [6].

Серверна частина складається з набору функціонально незалежних компонентів, кожен із яких відповідає за окремий етап обробки запиту. Основна логіка побудована як послідовний конвеєр: отримання повідомлення → визначення типу запиту → вилучення тексту за потреби → попередня обробка → класифікація → формування відповіді → додаткові дії за запитом користувача (профіль аналізу, перевірка відкритих джерел, освітній модуль). Загальну структурну схему компонентів системи наведено на рис. 1.

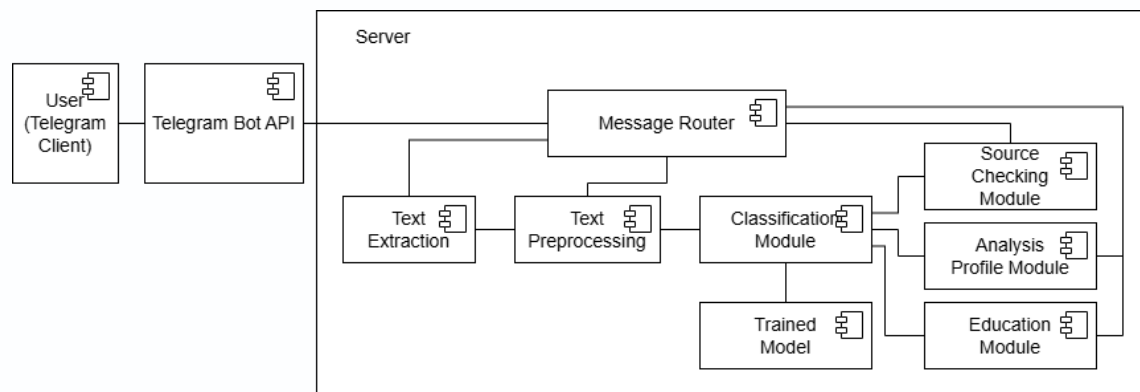


Рис. 1. Загальна структурна схема компонентів Telegram-бота для виявлення фейкових новин

Компонент маршрутизації приймає вхідний запит і визначає, до якого компонента системи потрібно передати дані: перевірка новини, перевірка заголовка або звернення до освітнього модуля. Компонент вилучення тексту використовується у випадках, коли вхідне повідомлення містить URL-адресу: він отримує текстовий вміст WEB-сторінки та передає виділений матеріал на попередню обробку. На етапі попередньої обробки виконується очищення

ISSN 2786-6025 Online

тексту: приведення до нижнього регістру, видалення службових символів, посилань та зайвих пробілів, а також перевірка мінімальної довжини та змістовності.

Загальна схема конвеєра обробки тексту та класифікації наведена на рис.

2.

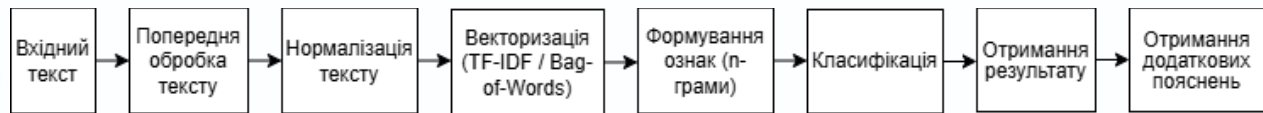


Рис. 2. Загальна схема конвеєра обробки тексту та класифікації

Після підготовки тексту виконується його векторизація – перетворення у числове представлення. Розглянуто два методи: TF-IDF та Bag-of-Words. TF-IDF зважає слова залежно від їхньої частоти в документі відносно всього корпусу – слово, що часто зустрічається в документі, але рідко у всьому наборі даних, отримує вищу вагу. Bag-of-Words базується на підрахунку частот без урахування позиції слова. Додатково задіяні n -грами – послідовності з одного або кількох слів, що враховують словосполучення, типові для маніпулятивних текстів.

Для класифікації порівнювались три алгоритми. Multinomial Naive Bayes – ймовірнісна модель, яка визначає клас на основі частот слів [7]:

$$C^* = \arg \max_C (P(C) \prod_{i=1}^n P(x_i | C)), \quad (1)$$

де C^* – клас, до якого відноситься текстове повідомлення,

n – кількість ознак у тексті,

x_i – окреме слово або ознака тексту,

$P(C)$ – апіорна ймовірність класу,

$P(x_i | C)$ – умовна ймовірність появи ознаки x_i у класі C .

Логістична регресія моделює лінійну залежність між ознаками тексту та ймовірністю належності до класу:

$$P(y = 1 | x) = \frac{1}{1 + e^{-(w^t x + b)}}, \quad (2)$$

де x вектор ознак тексту (наприклад, TF-IDF представлення),

w – вектор вагових коефіцієнтів моделі,

b – зсув (bias),

$w^t x$ – скалярний добуток, що відображає внесок ознак у результат,

$P(y = 1 | x)$ – ймовірність належності тексту до позитивного класу [8].

На відміну від Naive Bayes, логістична регресія не спирається на припущення про незалежність ознак, що краще підходить для довших текстів. Random Forest підвищує точність за рахунок комбінування кількох незалежних дерев рішень, фінальний прогноз формується за більшістю голосів:

$$y = \text{mode}\left(\{f_t(x)\}_{t=1}^T\right), \quad (3)$$

де $f_t(x)$ – прогноз t -го дерева рішень,

T – загальна кількість дерев в ансамблі,

mode – функція вибору найчастішого значення,

y – фінальний прогноз моделі.

Для навчання та тестування моделей використано відкритий набір даних Fake and Real News Dataset із платформи Kaggle, що містить понад 40 тис. англomовних записів (файли True.csv та Fake.csv). Розподіл класів збалансований: реальні новини – 54 %, фейкові – 46 %. Для оцінювання якості застосовано метрики Accuracy, Precision, Recall і F1.

Результати серії експериментів із класифікації текстів новин наведено в табл. 1.

Таблиця 1

Результати експериментів з класифікації текстів новин

№	Модель	Accuracy	Precision	Recall	F1	Час навчання, с	Час аналізу, мс
1	LogReg + TF-IDF (1,2)	0,9821	0,9826	0,9782	0,9804	50,44	1,0818
2	LogReg + TF-IDF (1,1)	0,977	0,9783	0,9712	0,9748	23,12	0,7364
3	RF + BoW (1,2)	0,9715	0,9858	0,9514	0,9683	81,97	1,1523
4	NB + BoW (1,2), $\alpha=0,5$	0,9348	0,9429	0,9129	0,9277	49,79	1,0802

ISSN 2786-6025 Online

Як видно з табл. 1, найвищий показник $F1=0,9804$ продемонструвала конфігурація Logistic Regression + TF-IDF з діапазоном n-грам (1,2) при часі аналізу $\sim 1,08$ мс. Використання біграм підвищує F1 логістичної регресії з 0,9748 (1,1) до 0,9804 (1,2). Random Forest дає схожу точність, але навчається у 1,8 раза довше (90 с проти 50 с). Multinomial Naive Bayes суттєво поступається за F1 (0,9277 при оптимальних параметрах). Порівняння F1-міри моделей аналізу текстів новин наведено на рис. 3.

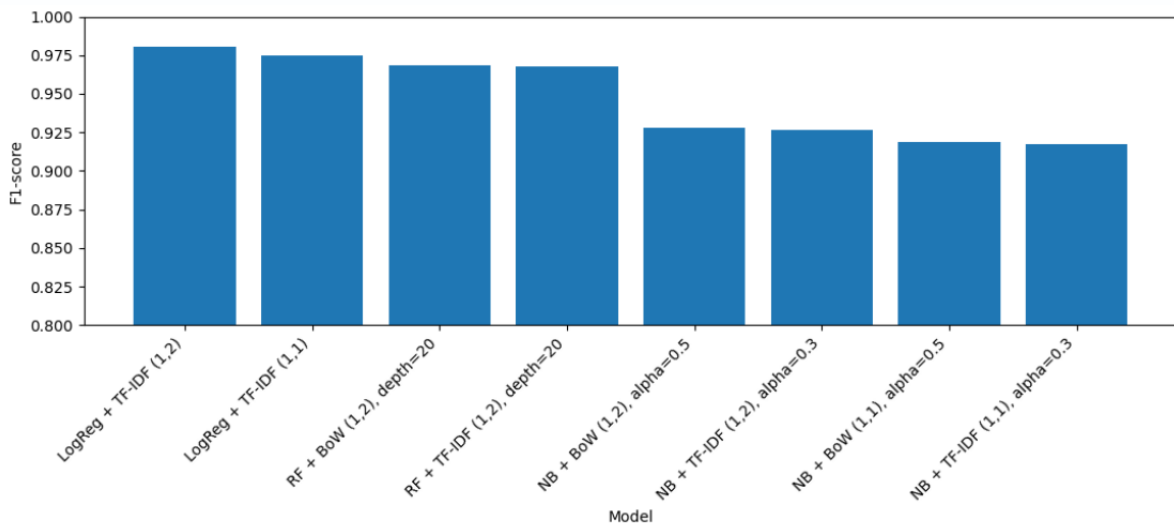


Рис. 3. Порівняння моделей аналізу текстів новин за значенням F1-міри

Результати серії експериментів з аналізу заголовків наведено в табл. 2.

Таблиця 2

Результати експериментів з аналізу заголовків

№	Модель	Accuracy	Precision	Recall	F1	Час навчання, с	Час аналізу, мс
1	LogReg + TF-IDF (1,1)	0,9486	0,9535	0,9333	0,9433	2,13	0,0384
2	NB + BoW (1,1), $\alpha=1,0$	0,9428	0,9291	0,9475	0,9382	1,06	0,0317
3	NB + TF-IDF (1,1), $\alpha=0,5$	0,9348	0,925	0,9333	0,9291	1,03	0,0323
4	RF + TF-IDF (1,2)	0,9072	0,9438	0,8479	0,8932	4,33	0,0591

Згідно з даними табл. 2 встановлено, що для заголовків логістична регресія формально показує найвищий $F1=0,9433$ (конфігурація (1,1)), однак різниця з Multinomial Naive Bayes + BoW (1,1) мінімальна – 0,0051 (0,54 %). При цьому NB навчається вдвічі швидше (1,06 с проти 2,19 с) і класифікує заголовки за $\sim 0,03$ мс. Використання біграм для заголовків не дає покращення: $F1$ NB знижується з 0,9382 до 0,9338 при переході з (1,1) на (1,2). Обраний діапазон n -грам (1,1) оптимальний для коротких текстів.

Порівняння $F1$ -міри моделей аналізу заголовків наведено на рис. 4.

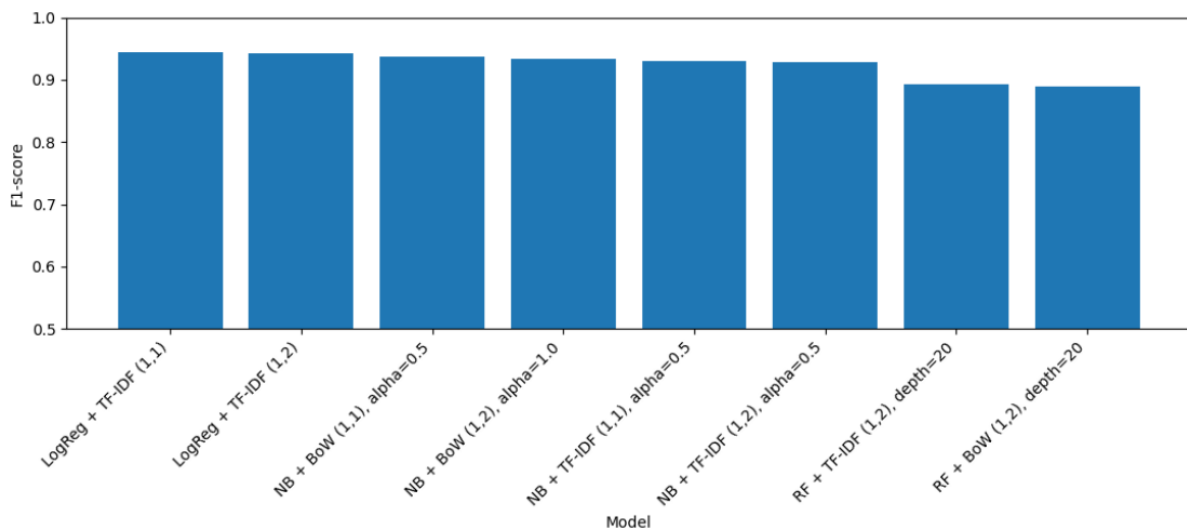


Рис. 4. Порівняння моделей аналізу заголовків за значенням $F1$ -міри

За результатами досліджень сформовано фінальні конфігурації: для режиму верифікації новин – Logistic Regression + TF-IDF, n -грами (1,2), $F1=0,9804$; для аналізу заголовків – Multinomial Naive Bayes + Bag-of-Words, n -грами (1,1), $F1=0,9382$. Обидві моделі збережено у форматі `.joblib` [9].

Ключовим програмним рішенням, що реалізує інтелектуальне ядро системи, є NLP-конвеєр аналізу тексту, реалізований у класі `AnalysisPipeline`. Він містить окремі методи для аналізу повного тексту новини та заголовка. Обидва сценарії виконують очищення тексту, визначення мови, переклад англійською за потреби, класифікацію відповідною ML-моделлю та формування профілю аналізу. Профіль поєднує результат ML-класифікації та NLP-евристики і містить чотири показники: емоційність тексту (за наявністю емоційно забарвлених слів і знаків оклику), ризик маніпуляції, конкретність фактів (за кількістю чисел, дат і цитат) і наявність джерел (за характерними мовними конструкціями та посиланнями). Ризик маніпуляції обчислюється як зважена комбінація відповідно до формули (4):

$$\begin{aligned} \text{manipulation} = & \text{fake}_{\text{percent}} \cdot 0,65 + \text{emotionality} \cdot 0,25 + \\ & +(100 - \text{source_presence}) \cdot 0,10, \end{aligned} \quad (4)$$

де $\text{fake}_{\text{percent}}$ – відсоток ризику фейковості від ML-моделі;

emotionality – показник емоційності тексту (%);

source_presence – показник наявності джерел (%).

Для заголовків використовується аналогічна логіка, але основну вагу має відсоток клікбейтності, оскільки заголовок є коротким текстом і містить менше фактологічних ознак. Усі показники нормалізуються в діапазон 0÷100 % і відображаються через inline-кнопку після короткого результату аналізу.

Фрагменти програмної реалізації двох режимів аналізу наведено на рис. 5 та рис. 6.

```
#Аналіз новини
class AnalysisPipeline:
    async def analyze_news(self, text: str):
        original_text = clean_text(text)
        detected_lang = self._detect_lang(original_text)
        processed_text = original_text
        if detected_lang != "en":
            processed_text = await self.translator.to_english(
                original_text,
                source_lang=detected_lang,
            )
            processed_text = clean_text(processed_text)
        prediction = self.news_classifier.predict(processed_text)
        prediction = self._adjust_news_result(prediction, original_text)
        prediction.profile = self.build_news_profile(
            original_text=original_text,
            prediction=prediction,
        )
        return prediction, detected_lang, processed_text
    def build_news_profile(self, original_text: str, prediction: Prediction):
        emotionality = self._emotionality_score(original_text)
        fact_specificity = self._fact_specificity_score(original_text)
        source_presence = self._source_presence_score(original_text)

        manipulation_risk = round(
            prediction.fake_percent * 0.65
            + emotionality * 0.25
            + (100 - source_presence) * 0.10
        )
        return AnalysisProfile(
            emotionality=self._clamp(emotionality),
            manipulation_risk=self._clamp(manipulation_risk),
            fact_specificity=self._clamp(fact_specificity),
            source_presence=self._clamp(source_presence),
        )
```

Рис. 5. Фрагмент програмної реалізації режиму аналізу новини

```
#Аналіз заголовка
class AnalysisPipeline:
    async def analyze_headline(self, text: str):
        original_text = clean_text(text)
        detected_lang = self._detect_lang(original_text)
        processed_text = original_text
        if detected_lang != "en":
            processed_text = await self.translator.to_english(
                original_text,
                source_lang=detected_lang,
            )
            processed_text = clean_text(processed_text)
        prediction = self.headline_classifier.predict(processed_text)
        prediction.profile = self.build_headline_profile(
            original_text=original_text,
            prediction=prediction,
        )
        return prediction, detected_lang, processed_text
    def build_headline_profile(self, original_text: str, prediction: HeadlinePrediction):
        emotionality = self._emotionality_score(original_text)
        fact_specificity = self._fact_specificity_score(original_text)
        source_presence = self._source_presence_score(original_text)
        manipulation_risk = round(
            prediction.clickbait_percent * 0.70
            + emotionality * 0.25
            + (100 - fact_specificity) * 0.05
        )
        return AnalysisProfile(
            emotionality=self._clamp(emotionality),
            manipulation_risk=self._clamp(manipulation_risk),
            fact_specificity=self._clamp(fact_specificity),
            source_presence=self._clamp(source_presence),
        )
```

Рис. 6. Фрагмент програмної реалізації режиму аналізу новини та заголовка

Як видно з рис. 5 та 6, обидва режими мають спільну NLP-частину, але використовують різні моделі класифікації. Для новини формується оцінка ризику фейковості, а для заголовка – рівень клікбейтності.

Окремою складовою реалізації є освітній модуль із трьома тематичними блоками (розпізнавання клікбейту, специфіка поширення новин у Telegram, алгоритм самостійної перевірки) та інтерактивним тестом із 10 запитань, які випадково обираються з бази 100 питань. Після проходження тесту бот підраховує кількість правильних відповідей і виводить користувачу підсумковий результат.

Зміна мови інтерфейсу між українською та англійською реалізована через функцію `t(lang, key)` і команду `/language`; бот також автоматично перекладає неангломовний вхідний текст перед подачею до ML-моделі.

Перевірка відкритих джерел через GDELT API виконується після натискання кнопки «Перевірити за джерелами»; якщо результатів немає – система формує кнопку переходу до Google з готовим пошуковим запитом.

Приклад результату аналізу текстового повідомлення та відповідного профілю наведено на рис. 7.

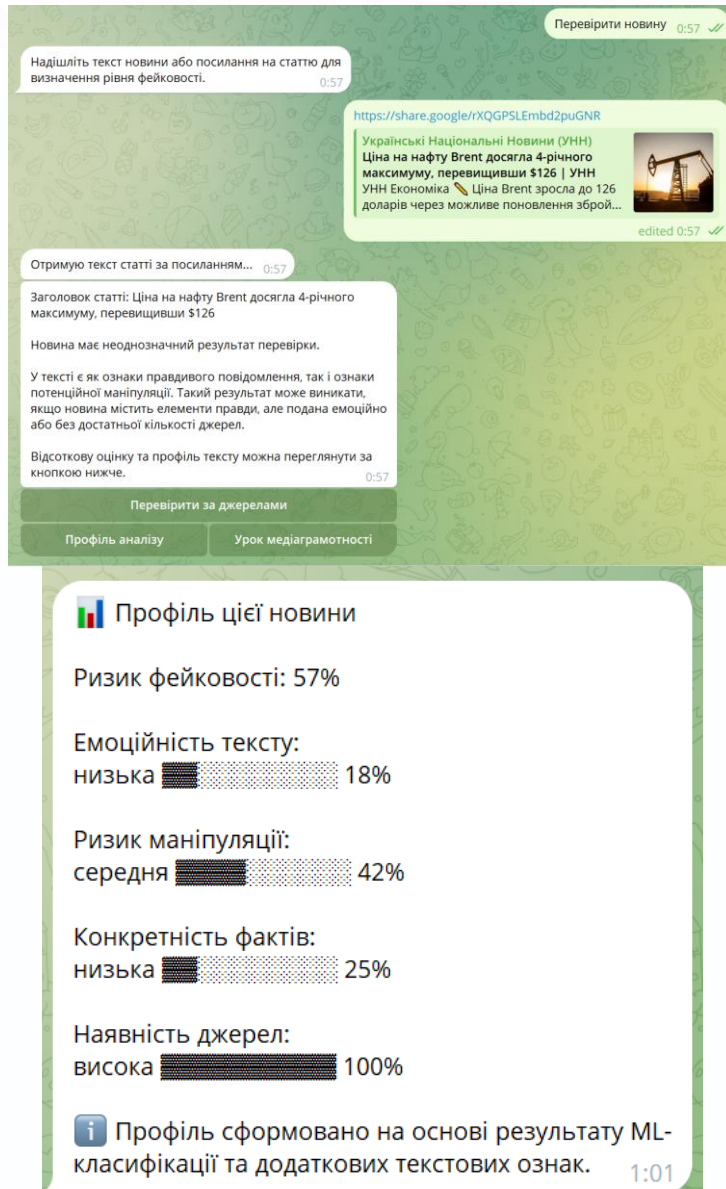


Рис. 7. Результат аналізу повідомлення та профіль аналізу

Тестування програмного продукту виконувалося вручну через інтерфейс Telegram: перевірено понад 100 сценаріїв, зокрема режими аналізу новин і заголовків, обробку URL, профіль аналізу, перевірку відкритих джерел, освітній модуль, мультимовність і реакцію на некоректне введення. Граничне тестування шкали 0÷100% підтвердило коректність переходів між діапазонами на межах 39÷40%, 59÷60% і 74÷75%. Час класифікації становить ~1,08 мс для текстів новин і 0,03÷0,04 мс для заголовків; затримка у сценаріях з URL та відкритими джерелами визначається швидкістю мережевих запитів.

Порівняння розробленого Telegram-бота з аналогами за функціональними можливостями наведено в табл. 3.

Таблиця 3

Порівняння функціональних можливостей розробленого Telegram-бота з аналогами

Критерій	Розроблений бот	StopFake Bot	GwaraMedia «Перевірка»	Fake True Bot
Тип перевірки	Автоматична ML-оцінка	Пошук у базі матеріалів	Передача запиту фактчекерам	Самостійна оцінка через запитання
Аналіз нового тексту	Є	Немає повного ML-аналізу	Переважно ручна обробка	Немає ML-аналізу
Аналіз заголовка	Є, окремий режим	Немає	Немає	Немає
Робота з URL	Є, із вилученням тексту статті	Частково, через матеріали сайту або базу	URL може бути надісланий на перевірку	Немає повноцінної обробки URL
Профіль аналізу	Є: 4 показники	Немає	Немає	Немає
Перевірка відкритих джерел	Є, із переходом до Google за готовим запитом	Через базу матеріалів	Через ручну перевірку фактчекерами	Немає
Освітній модуль	Є: 3 блоки та тест	Частково через матеріали сайту	Частково через медіаматеріали	Є елементи самоперевірки
Мультимовний інтерфейс	Українська / англійська	Обмежено	Обмежено	Обмежено

Як видно з табл. 3, розроблений бот розширює наявні рішення за сімома функціональними характеристиками: підтримує окремий аналіз заголовків; обробляє статті за URL, формує профіль аналізу з чотирма показниками, виконує пошук схожих матеріалів у відкритих джерелах із Google fallback, містить розширений освітній модуль із тематичними блоками та тестуванням, реалізує двомовний інтерфейс та автоматичний переклад вхідного тексту.

ISSN 2786-6025 Online

Висновки. Розроблено Telegram-бот для виявлення фейкових новин із використанням NLP і моделей штучного інтелекту. Реалізовано клієнт-серверну архітектуру, де клієнтська частина представлена Telegram-інтерфейсом, а серверна виконує маршрутизацію запитів, попередню обробку тексту, ML-класифікацію, формування профілю аналізу та перевірку відкритих джерел.

Для реалізації обрано мову Python, фреймворк aiogram та бібліотеки scikit-learn [10]. Навчання моделей виконано на відкритому датасеті Fake and Real News Dataset (понад 40 тис. записів, Kaggle). Модель класифікації новин Logistic Regression + TF-IDF з діапазоном n-грам (1,2) досягла $F1=0,9804$ при часі аналізу $\sim 1,08$ мс. Модель аналізу заголовків Multinomial Naive Bayes + Bag-of-Words з діапазоном n-грам (1,1) досягла $F1=0,9382$ при часі аналізу $0,03 \div 0,04$ мс.

Мета роботи – розширення функціональних можливостей Telegram-бота для виявлення фейкових новин досягнута за рахунок реалізації таких функціональних розширень: автоматичної ML-оцінки фейковості тексту; окремого режиму аналізу заголовків на клікбейтність; профілю аналізу з відсотковою оцінкою та чотирма показниками тексту (емоційність, ризик маніпуляції, конкретність фактів, наявність джерел); розширеного освітнього модуля з трьома тематичними блоками та інтерактивним тестуванням; автоматичного пошуку схожих матеріалів у відкритих джерелах із Google fallback; обробки статей за URL із вилученням тексту зі сторінки; двомовного інтерфейсу (українська/англійська) та автоматичного визначення мови й перекладу вхідного тексту. Розроблений бот не замінює професійний фактчекінг, але є інструментом швидкої попередньої перевірки новинного контенту в середовищі Telegram.

Література:

1. Shu K. Fake News Detection on Social Media: A Data Mining Perspective / K. Shu, A. Sliva, S. Wang, J. Tang, H Liu // ACM SIGKDD Explorations Newsletter. – 2017. – Vol. 19, Issue 1. – P. 22-36.
2. Oshikawa R., Qian J., Wang W. Y. A Survey on Natural Language Processing for Fake News Detection [Електронний ресурс] / R. Oshikawa, J. Qian, W. Y. Wang // Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018). – 2018. – Режим доступу: <https://aclanthology.org/L18-1107/>.
3. StopFake. Як розпізнати фейк [Електронний ресурс]. – Режим доступу: <https://www.stopfake.org/uk/yak-rozpoznaty-fejk/>.
4. GwaraMedia. «Перевірка» – бот, який виявляє фейкові новини [Електронний ресурс]. – Режим доступу: <https://gwaramedia.com/perevirka-razom-smozhemo/>.
5. Романець К. В. / К.В. Романець, Л.В. Крилик // Автоматизована перевірка новин у месенджері Telegram на основі поєднання методів NLP та освітнього підходу [Електронний ресурс] // LIV Всеукраїнська науково-технічна конференція факультету інтелектуальних

ISSN 2786-6025 Online

інформаційних технологій та автоматизації (ВНТУ-2026). м. Вінниця – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2026/paper/view/27320>.

6. Telegram Bot API [Електронний ресурс]. – Режим доступу: <https://core.telegram.org/bots/api>.

7. Scikit-learn. MultinomialNB [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html.

8. Scikit-learn. LogisticRegression [Електронний ресурс]. – Режим доступу: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html.

9. Kaggle. Training models for fake news and clickbait detect [Електронний ресурс]. – Режим доступу: <https://www.kaggle.com/code/katerynaromanets/training-models-for-fake-news-and-clickbait-detect>.

10. aiogram. aiogram 3.x Documentation [Електронний ресурс]. – Режим доступу: <https://docs.aiogram.dev/en/v3.15.0/>.

References:

1. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective [Fake News Detection on Social Media: A Data Mining Perspective]. *ACM SIGKDD Explorations Newsletter – ACM SIGKDD Explorations Newsletter*, 19, 1, 22-36.

2. Oshikawa, R., Qian, J., & Wang, W.Y. (2018). A Survey on Natural Language Processing for Fake News Detection [A Survey on Natural Language Processing for Fake News Detection]. Proceedings from LREC 2018: 11th International Conference on Language Resources and Evaluation – The 11th International Conference on Language Resources and Evaluation. Retrieved from <https://aclanthology.org/L18-1107/>.

3. StopFake. Yak rozpiznaty fejk [StopFake. How to recognize a fake]. www.stopfake.org/uk/yak-rozpiznaty-fejk/. Retrieved from <https://www.stopfake.org/uk/yak-rozpiznaty-fejk/> [in Ukrainian].

4. GwaraMedia. «Perevirka» – bot, yakyi vyiavliaie feikovi novyny [GwaraMedia. «Perevirka» – a bot that detects fake news]. gwaramedia.com/perevirka-razom-smozhemo/. Retrieved from <https://gwaramedia.com/perevirka-razom-smozhemo/> [in Ukrainian].

5. Romanets, K.V., & Krylyk, L.V. (2026). Avtomatyzovana perevirka novyn u mesendzheri Telegram na osnovi poiednannia metodiv NLP ta osvithnoho pidkhodu [Automated news verification in the Telegram messenger based on a combination of NLP methods and an educational approach]. Proceedings from VNTU '26: LIV Vseukrainska naukovo-tekhnichna konferentsiia fakultetu intelektualnykh informatsiinykh tekhnolohii ta avtomatyzatsii – The 54th All-Ukrainian Scientific and Technical Conference of the Faculty of Intelligent Information Technologies and Automation. Vinnitsia: VNTU. Retrieved from <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2026/paper/view/27320> [in Ukrainian].

6. Telegram Bot API. Retrieved from <https://core.telegram.org/bots/api>.

7. Scikit-learn. MultinomialNB. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.MultinomialNB.html.

8. Scikit-learn. LogisticRegression. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html.

9. Kaggle. Training models for fake news and clickbait detect. Retrieved from <https://www.kaggle.com/code/katerynaromanets/training-models-for-fake-news-and-clickbait-detect>.

№ 5(59)
2026

НАУКА i ТЕХНІКА

СЬОГОДНІ

ISSN 2786-6025 Online

10. aiogram. aiogram 3.x Documentation. Retrieved from <https://docs.aiogram.dev/en/v3.15.0/>.

Дата першого надходження статті до видання: 14.05.2026

Дата прийняття статті до друку після рецензування: 27.05.2026

4940