

УДК 004.02

DOI: 10.18372/2073-4751.76.18241

Міщенко Л.Д.,  
orcid.org/0000-0002-5101-2517,  
Клименко І.А., д.т.н.,  
orcid.org/0000-0001-5345-8806

## РОЗПІЗНАВАННЯ ФЕЙКОВИХ НОВИН ІЗ ВИКОРИСТАННЯМ ОБРОБКИ ПРИРОДНОЇ МОВИ Й АРХІТЕКТУРИ З НИЗЬКИМ ЕНЕРГОСПОЖИВАННЯМ ДЛЯ ПЕРИФЕРІЙНИХ ОБЧИСЛЕНЬ

Національний технічний університет України  
“Київський політехнічний інститут імені Ігоря Сікорського”

liudamishchenko@gmail.com,  
ikliryna@gmail.com

### **Вступ**

Сучасний інформаційний простір заповнений новинами, що створює виклики у виявленні та розрізненні правдивих та фейкових новин. Проблема їх поширення стає все актуальнішою в суспільстві, вимагаючи вдосконалення методів розпізнавання на основі обробки природної мови та інтеграції з низькопотужними обчислювальними архітектурами для периферійних обчислень.

У роботі пропонується методика виявлення фейкових новин, що базується на механізмі уваги та контекстуальному аналізі в текстах. Відмінність у використанні цих механізмів полягає у здатності моделі фокусуватися на певних лінгвістичних аспектах текстів, забезпечуючи ефективну класифікацію без великих витрат обчислювальних ресурсів.

Результати експериментальних досліджень свідчать про успішність запропонованого методу в розпізнаванні фейкових новин через визначення відмінностей у стилі письма та мовних маркерах достовірності у текстах. Аналіз показує, що фейкові новини мають вищу частоту вживання емоційно забарвлених слів, імовірніших висловлювань та нецензурної лексики у порівнянні з правдивими новинами. Отримані результати також підтверджують відмінність у співвідношенні слів, спрямованих на повідомлення, у текстах фейкових та правдивих новин.

Інтеграція механізму уваги та контекстуального аналізу виявилася відмінною

для підвищення точності та ефективності розпізнавання фейкових новин у текстах. Оцінка моделі показала 81% точність у бінарній класифікації, що свідчить про спроможність моделі відрізняти справжні новини від фейкових.

Ця робота відкриває можливості подальших досліджень у галузі розпізнавання фейкових новин, спрямованих на вдосконалення методів обробки мовних особливостей та розуміння складних лінгвістичних нюансів у текстах новин.

### **Аналіз літературних даних та постановка проблеми**

Проблема фейкових новин в епоху цифрових медіа та швидким обігом інформацією, що характеризується навмисним поширенням оманливої або неправдивої інформації, привернула значну увагу багатьох дослідників. Виявлення фейкових новин є складним завданням, і численні наукові праці заглиблюються в тонкощі цієї проблеми, багато з яких через призму аналізу стилю письма.

У дослідженні [1] описано проблему виявлення фейкових новин на платформах соціальних медіа та запроваджено структуру інтелектуального аналізу даних. У ньому наголошується на важливості мовних і стилістичних особливостей у виявленні оманливого контенту, окремо зроблено аналіз стилю письма. Однак праця підкреслює необхідність багатовимірного підходу, який включає лінгвістичний аналіз і приділяє більше уваги поведінці користувачів і надійності

джерела та домену, що недоступно в рамках запропонованого способу.

Вирішенню цього питання присвячена наукова робота [2], яка використовує теоретичний підхід до раннього виявлення фейкових новин. Тут опрацьовано лінгвістичні ознаки та стиль написання як найважливіші показники, зосереджуючись на відмінностях у використанні мови між правдивими та оманливими новинами. Проте дослідники запропонували теоретичну модель аналізу, яка веде до практичної реалізації та проведення реальних експериментів. З іншого боку, вони покладаються на шаблонне написання статей, яке вже не є актуальним у сьогоденні.

Варіантом подолання цих труднощів може бути комплексний підхід [3]. У дослідженні обговорюються фундаментальні теорії та методи виявлення фейкових новин. Підкреслено роль мовних особливостей і стилю написання в розрізненні справжніх новин від фейкових. Проте в праці розглянуто лише деякі особливості стилю письма, такі як емоційність, лайка.

Рішення цієї ж проблеми автори опублікували в дослідженні [4], за результатами якого мова ворожнечі часто зустрічається у фейкових новинах. Ця праця присвячена виявленню ворожих висловлювань на основі аналізу стилю написання, емоційного тону та лінгвістичних маркерів, пов'язаних із шкідливим вмістом, що сприяє повному розумінню виявлення фейкових новин. Хоча деякі частини даної роботи ще потребують вирішення. Наприклад, досягти більш детальної класифікації та повнішого діапазону категорій при заповненні бази даних.

Для подолання цих труднощів вчені розробляють автоматичні методи виявлення фейкових новин на основі стилю написання в дослідженні [5]. Вони створили два класифікатори: нейронну мережу та модель на основі стилOMETричних характеристик. Аналіз стилOMETричної моделі показує, що вона зосереджена на сенсаційній та афективній лексиці, яка характерна для фейкових новин. Однак у роботі не

враховано ненормативну лексику, слова, що описують ймовірність події, емоційне забарвлення тексту. Ці показники істотно впливають на результат і розширюють можливості аналізу новин.

Проблема обмежених можливостей обчислень розглядається у статті [6], де виявлення фейкових новин відбувається в режимі реального часу. Запропонований підхід використовує методи вилучення подій і тем у поєднанні з механізмом об'єднання тем для обробки даних новин і зменшення їх кількості. Результати дослідження свідчать про те, що описаний метод демонструє менше часу для обробки масиву новин, а отже й обчислювальних ресурсів. Проте експерименти не проводилися з архітектурою з низьким електроспоживанням.

Рішення з механізмом уваги для аналізу тексту було запропоновано в дослідженні [7]. В основі цього рішення є механізм уваги для архітектури з низьким електроспоживанням. Проте специфіка тексту не відповідає новинам, тому ця проблема залишається відкритою.

Обробка природної мови передбачає створення алгоритмів і моделей, призначених для автономного визначення та розрізнення справжніх статей новин від оманливого або сфабрикованого вмісту. Як обговорювалося раніше в статті [8], у центрі уваги було використання технології обробки природної мови для виявлення фейкових новин. Усі наступні розробки спрямовані на вдосконалення, таке як точність класифікації, або прискорення вже існуючих методів.

### **Мета та задачі дослідження**

Метою дослідження є вдосконалення методу розпізнавання фейкових новин на основі обробки природної мови з використанням архітектури з низьким електроспоживанням для периферійних обчислень за рахунок механізму уваги та контекстуального аналізу. Це дасть можливість виконати перевірку новин на периферійних пристроях з обмеженими обчислювальними ресурсами без втрати швидкості.

Для досягнення мети поставлено такі задачі:

- удосконалити метод обробки природної мови для аналізу новин на периферійних пристроях із застосуванням механізму уваги;
- розробити спосіб розпізнавання мовних ознак або маркерів достовірності в письмовому стилі, який дозволяє вдосконалити метод розпізнавання фейкових новин з архітектурою з низьким електроспоживанням.

### **Матеріали та методи досліджень**

Беручи до уваги швидкість поширення новин, стандартний аналіз обробки природної мови на периферійних пристроях не задовольняє рівнем ефективності. Він потребує вдосконалення та різних методів зменшення обсягів обчислень. Тому об'єктом запропонованого дослідження є метод обробки природної мови для автоматичного розпізнавання фейкових новин. Основною гіпотезою дослідження є оптимізація методу аналізу новин з використанням архітектури з низьким електроспоживанням за допомогою застосування механізму уваги та контекстуального аналізу.

### **Модифікований метод виявлення фейкових новин**

Відомий метод для ідентифікації фейкових новин за допомогою обробки природної мови описаний у роботі [9]. Опрацювання вхідних даних за описаним методом базується на загальноприйнятих способах збору та форматуванню даних для їх подальшого аналізу та моделювання. Етапи опрацювання даних за наведеним методом ідентифікації фейкових новин наведено далі:

1. Збір даних: перший крок полягає у виборі різноманітних, надійних джерел для набору даних. Вхідні набори даних узяті із бази даних Європейського Союзу, яка містить надійні новинні джерела, тексти новин та публікацій користувачів соціальних мережах, а також веб-сайти, які зазвичай публікують неправдиву інформацію. Це забезпечує різноманітність вхідних

даних, дозволяє моделі працювати з різним контентом, включає посилання на перевірені джерела та факти. Такий підхід до формування масиву даних допомагає визначити справжні або фейкові новини, порівняти отримані результати аналізу та зробити оцінку правильності роботи алгоритму.

2. Маркування: точне маркування або присвоєння лейблів є важливим для навчання та оцінки моделі. Кожна стаття в наборі даних автоматично позначається як «справжня» або «фейк» на основі оцінки фактчекерів з бази даних перевірених новин.

3. Балансування: забезпечення збалансованого набору даних з майже рівною кількістю справжніх та фейкових статей необхідно для запобігання упередженості в моделі та оптимізації продуктивності.

4. Попередня обробка даних: починається з очищення тексту. Елементи, такі як *HTML*-теги, *URL*-адреси та спеціальні символи, видаляються для фокусування моделі *NLP* на смисловому наповненні тексту. Далі відбуваються перетворення усіх на нижній регістр, токенизація, вилучення стоп-слів, стемінг, лематизація, обробка числових даних, видалення пустих значень із бази даних, робота з незбалансованими класами та кодування міток.

5. Після завершення збору даних та попередньої обробки отримуємо очищений та структурований набір даних, готовий для виокремлення особливостей тексту та моделювання.

6. Вилучення ознак: попередньо оброблені текстові дані перетворюються в числові значення, які використовуються для машинного навчання. Текст перетворюється у векторне представлення високої розмірності, зберігаючи семантичну інформацію.

7. Лінгвістична та контекстуальна інженерія ознак: на цьому етапі виділяються додаткові лінгвістичні ознаки (теги частин мови, синтаксичні залежності, іменовані сутності) відповідають за детальні лінгвістичні характеристики. Контекстуа-

льні ознаки (оцінки аналізу настроїв) – емоційне забарвлення тексту.

8. Класифікація та оцінка важливості слів у тексті: використовується алгоритми класифікації, такі як Наївний Байєс, метод опорних векторів. *TF-IDF* є найбільш розповсюдженим серед них.

9. Навчання моделі: на основі векторного представлення токенів справжніх та фейкових новин, обрані моделі проходять етап навчання. Також відбувається налаштування гіперпараметрів для оптимізації продуктивності.

10. Перехресна перевірка та оцінка результатів: оцінка продуктивності моделі за допомогою метрик, таких як точність, швидкість та спроможність запам'ятовування, вимірюють її ефективність у відділенні справжніх та фейкових новин.

11. Тестування стійкості: моделі піддаються тестуванню стійкості на наборах новин, не тренуваних раніше. Важливо, щоб набори даних відрізнялися наповненням та лінгвістичними особливостями.

В роботі запропоновано вдосконалення методу розпізнавання фейкових новин на етапі лінгвістичною та контекстуальної інженерія ознак. Зроблено припущення, що використання механізму уваги та контекстуального аналізу дозволить опрацювання масиву вхідних даних на периферійних пристроях без впливу на швидкість роботи алгоритму.

### **Особливості способу реалізації механізму уваги**

Механізми уваги в обробці природної мови дозволяють моделям зосереджуватися на певних частинах тексту під час обробки інформації. Метод розпізнавання фейкових новин з аналізом стилю письма запропоновано вдосконалити за допомогою механізму уваги. Він позначає потенційно оманливий або сфабрикований вміст, розпізнаючи відхилення від очікуваних норм стилю письма. В основі механізму уваги закладено виявлення аномалій. Тобто, звертається особлива увага на певні слова, фрази чи структури речень. Так у дослідженні зроблено акцент на слова для

опису репортажів або звітів та розрізняються як “аномалії”. Вони мають вирішальне значення в написанні розповідей, діалогів, інтерв'ю та журналістських дописах, оскільки допомагають приписувати заяви та дії конкретним особам, збагачуючи текст і роблячи його більш привабливим та інформативним. Приклади: сказати, сказати, запитати, прошепотіти, стверджувати, запропонувати, повідомити, пояснити, погодитися, заперечити.

Механізм, застосований до аналізу стилю написання, допомагає ідентифікувати потенційні ознаки фейкових новин. Завдяки цьому, модель *NLP* покращує здатність виявляти відхилення в стилі написання та вказує на фейковість новини.

### **Виокремлення мовних ознак або маркерів достовірності в стилі написання**

Виявлення маркерів достовірності передбачає розпізнавання мовних і контекстуальних ознак у тексті, які вказують на достовірність або надійність вхідної інформації. Ці маркери відіграють вирішальну роль у розрізненні справжніх новин від фейкових. Вони дозволяють запропонованому методу звертати увагу на конкретні лінгвістичні аспекти та контекстуальні підказки, які сигналізують про оманливі нарративи в тексті.

Розглянуто розпізнавання мовних ознак або маркерів достовірності в стилі написання, таких як емоційно забарвлена лексика, «легка лайка» або імовірнісні слова. Детальне пояснення до кожного виділеного лексичного напрямку:

- емоційно забарвлена лексика, також відома як емоційно насичена або емоційна мова, – це слова та фрази, які мають сильний емоційний відтінок, часто викликаючи певні почуття, реакції та настрої у читача чи слухача. Ці слова можуть викликати емоційну реакцію, як позитивну, так і негативну, і часто використовуються, щоб переконати, вплинути або залучити аудиторію на емоційному рівні. Приклади: «сердечний», «жахаючий», «чарівний», «трагічний», «надихаючий»;

- імовірнісні слова – це слова або терміни, які передають невизначеність, вірогідність або ймовірність. Ці слова та фрази часто виражають можливість події чи стану, створюючи відчуття ймовірності чи випадковості. У різних контекстах імовірнісні слова допомагають передати ступінь достовірності чи невизначеності. Приклади: ймовірно, можливо, ймовірно, можливо, можливо, здається, може;

- «легка лайка» – зазвичай стосуються виразів або фраз, які вважаються відносно м'якими або менш образливими порівняно з сильнішими формами лайки чи лайки. Ці вирази все ще можуть виражати розчарування, роздратування чи акцент, але вони, як правило, менш вульгарні чи соціально неприйнятні. Люди часто використовують легку лайку, щоб висловити свої почуття, не використовуючи відвертих чи грубих слів. Приклади м'якої лайки включають такі слова, як «дідько», «бовдур», «гімно», «нахаба», «потвора», «шеллепа».

Вищевказані приклади слів виокремлюються як текстова одиниця, тоді як легка лайка та ненормативна лексика зазвичай є фразами чи висловлюваннями. Це ставить додаткову задачу, яка потребує зосередження на тонкощах синтаксичної складності, семантичних нюансів, намагаючись досягнути не тільки значення слова, а й суть висловлювання.

Уже існує подібна модель класифікації, заснована на лінгвістичних ознаках. Вона описана в статті [10] і може бути застосована як основа для розробки. Цей підхід виявився ефективним, але автори відзначають, що набір даних необхідно розширити, щоб включити нові класифікації. Такі розширення запропоновано в цьому дослідженні.

### **Методи оцінки ефективності моделі**

Для оцінки ефективності моделі в розрізненні справжніх і фейкових новинних статей, з урахуванням поставленої мети роботи, обрано методи оцінки *Precision* і *Recall*.

Назва методів перекладається як Точність і Запам'ятовування. Вони є важливими в задачах бінарної класифікації, де один клас вважається позитивним (наприклад, правдиві новини), а інший – негативним (наприклад, фейкові новини).

*Precision* (1) вимірює точність позитивних прогнозів, зроблених моделлю. Це відношення правильно передбачених позитивних спостережень до загальної кількості прогнозованих позитивних спостережень. Високий бал *Precision* вказує на низький рівень помилкових позитивних результатів.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

*Recall* (2), також відомо як чутливість, вимірює здатність моделі знаходити всі позитивні випадки. Це відношення правильно прогнозованих позитивних спостережень до фактичних позитивних спостережень.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

У контексті виявлення фейкових новин *Precision* допоможе оцінити частку правильно ідентифікованих справжніх новинних статей серед усіх статей, які вважаються справжніми. У той же час *Recall* оцінить, скільки справжніх новинних статей було правильно визначено з усіх справжніх новинних статей.

### **Результати дослідження розпізнавання фейкових новин на периферійних пристроях з використанням механізму уваги та контекстуального аналізу**

#### **1. Результати удосконаленого методу розпізнавання фейкових новин із застосуванням механізму уваги**

Для підтвердження результатів удосконалення методу розпізнавання фейкових новин на основі обробки природної мови з використанням архітектури з низьким електроспоживанням для периферійних обчислень за рахунок механізму уваги проведено ряд експериментальних досліджень. Вони зосереджені на виявленні частоти “аномалій” у тексті.

Середньостатистичне відсоткове наповнення груп слів для опису репортажів або звітів у текстах новин показано в табл. 1.

Таблиця 1. Наявність слів для опису репортажів або звітів у правдивих та фейкових новинах у відсотках

Група слів	Правдиві новини, %	Фейкові новини, %
Слова репортажів	32	18

За отриманими результатами кількість слів для опису репортажів становить меншу відсоткову частину тексту у фейкових новинах, порівнюючи з правдивими.

## 2. Результати удосконаленого методу розпізнавання фейкових новин із виявленням маркерів достовірності

Для обґрунтування удосконалення методу розпізнавання фейкових новин на основі обробки природної мови з використанням архітектури з низьким електроспоживанням для периферійних обчислень за рахунок виявлення маркерів достовірності проведено ряд експериментальних досліджень. Вони зосереджені на виявленні частоти легкої лайки та нецензурних висловлювань, емоційно забарвлених слів, слів імовірності у тексті. Середньостатистичне

відсоткове наповнення зазначених груп слів у текстах новин показано в таблиці 2.

Таблиця 2. Наявність певних груп слів у правдивих та фейкових новинах у відсотках

Група слів	Правдиві новини, %	Фейкові новини, %
Легка лайка	5	14
Емоційні слова	34	61
Нецензурні висловлювання	2	9
Слова ймовірності	8	25

Згідно результатів експериментів, кількість легкої лайки та нецензурних висловлювань, емоційно забарвлених слів, слів імовірності, становить більшу відсоткову частину у фейкових новинах порівнюючи з правдивими.

Результати впливу емоційно забарвлених слів на контекстне наповнення тексту фейкових новин показано на рис. 1. Графік відсоткового показує відсоткове співвідношення груп слів механізму уваги та контекстуального аналізу в неправдивих та правдивих текстах.

The dependence graph of the use of emotional coloring of the text on fake and real news

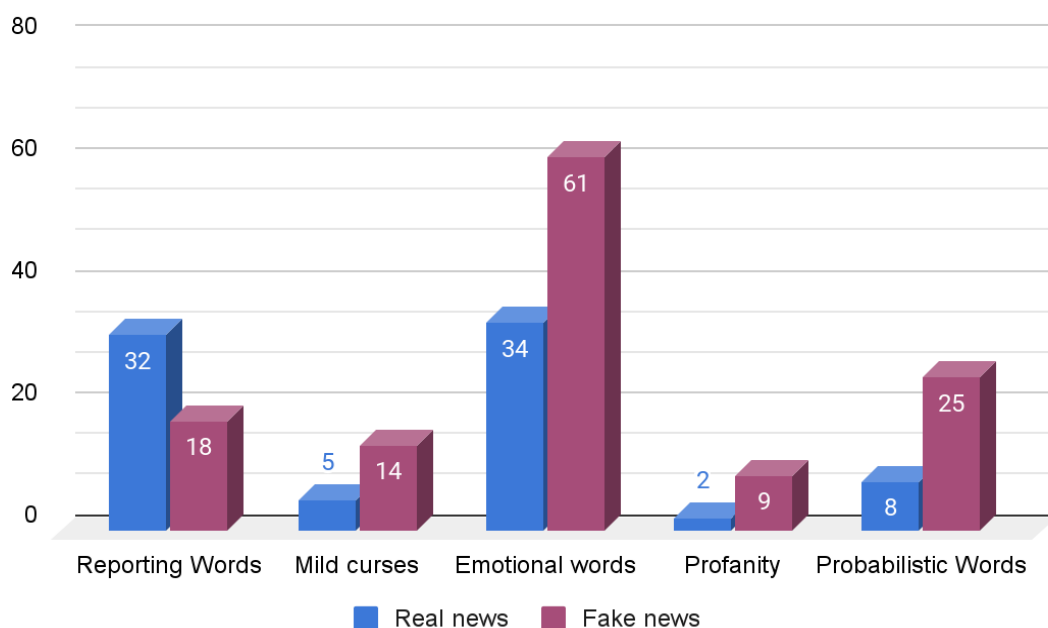


Рис. 1. Відсоткове наповнення текстів новин емоційно забарвленими словами

### **Обговорення результатів дослідження розпізнавання фейкових новин на основі NLP з використанням архітектури з низьким електроспоживанням для периферійних обчислень**

Дослідження, присвячене розпізнаванню фейкових новин за допомогою *Natural Language Processing* з використанням архітектури з низьким електроспоживанням для периферійних обчислень за рахунок механізму уваги та контекстуального аналізу представляє комплексний підхід до боротьби з дезінформацією.

Для проведення дослідження було навчено модель та на її основі проаналізовано 4 різні набори новин із різних джерел: соціальні мережі *Facebook* та *X*, бази перевірених новин *EUvsDisinfo*, *ZaboronaMedia*. Отримані результати вказують на наявність визначених груп слів у текстах новин, їх представлено у відсотковому співвідношенні та показано на рис. 1. Вони підтверджують, що фейкові новини більш наповнені емоційно забарвленими словами, слова ймовірності та нецензурною лексикою. При цьому, слова для звітів чи репортажів використовуються рідше. Це пояснюється тим, що зазвичай автори фейкових новин намагаються рідше писати висновки. Таким чином, вони спонукають читача робити суб'єктивні або далекі від істини висновки.

Інтеграція механізму уваги та контекстуального аналізу в метод розпізнавання фейкових новин показала свою ефективність та надійність для обробки тексту новин. Точність запропонованого підходу бінарної класифікації з використанням архітектури з низьким електроспоживанням становить 81%, що є помітною перевагою. Це пояснюється тим, що аналіз тільки чотирьох груп слів, а не всього масиву текстів, не потребує значних обчислювальних ресурсів. На відміну від рішення [11], де низьку складність обчислень вдається досягнути шляхом попередньої обробки даних та двох специфічних мов в методі розпізнавання фейкових новин. Цей підхід

дозволяє виокремити такі групи слів, при яких висока точність класифікації новин значним чином не впливає на витрати обчислювальних ресурсів роботи методу та на особливу мову вхідних текстів.

Проведена робота сприяє постійним зусиллям зі зменшення впливу фейкових новин у соціумі, базуючись на розумінні лінгвістичних нюансів. Під час проведення дослідження вдалося досягти високої точності розпізнавання фейкових новин з обмеженими обчислювальними можливостями архітектури з низьким електроспоживанням.

Дослідження підкреслює важливість тонкого розуміння стилів письма як потужного інструменту для вирішення проблеми розпізнавання фейкових новин за допомогою обробки природної мови з використанням архітектури з низьким електроспоживанням для периферійних обчислень. Результати дослідження створюють основу для подальших досліджень у галузі аналізу новин на правдивість на пристроях із обмеженими обчислювальними ресурсами. А саме, для напрямку підвищення надійності та точності у обробці різноманітних мовних особливостей і складних лінгвістичних нюансів, властивих розпізнаванню фейкових новин.

Майбутні дослідження можуть зосереджуватися на порівнянні механізму уваги та контекстуального аналізу з іншими найсучаснішими алгоритмами. Також вивченні впливу емоційної забарвленості тексту у сценаріях, що включають мультимедійний контент або еволюцію мовних шаблонів для підвищення його застосовності у соціальних мережах.

### **Література**

1. Shu K., Sliva A., Wang S., Tang J., Liu H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*. 2017. V. 19(1). P. 22–36.
2. Zhou X., Jain A., Phoha V.V., Zafarani R. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice*. 2020. V. 1(2). P. 1–25.

3. Zhou X., Zafarani R. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*. 2020. V. 53(5). P. 1–40.
4. Schmidt A., Wiegand M. A survey on hate speech detection using natural language processing. *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media* / Valencia, Spain, 2017. P. 1–10.
5. Przybyla P. Capturing the style of fake news. *Proceedings of the AAAI conference on artificial intelligence*. V. 34, No. 01. 2020. P. 490–497.
6. Zhang C., Gupta A., Qin X., Zhou Y. A computational approach for real-time detection of fake news. *Expert Systems with Applications*. 2023. V. 221. 119656.
7. Vaswani A., et al. Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS 2017)* / Long Beach, USA, 2017. 11 p.
8. Mishchenko L., Klymenko I. Method for Detecting Fake News Based on Natural Language Processing. *The VI International Scientific and Practical Conference «Modern ways of solving the problems of science in the world»* / Warsaw, Poland, 2023. P. 375–378.
9. Alammery, A.S. Arabic questions classification using modified TF-IDF. *IEEE Access*. 2021. V. 9. P. 95109–95122.
10. Pathak A., Srihari R.K. BREAKING! presenting fake news corpus for automated fact checking. *Proceedings of the 57th annual meeting of the association for computational linguistics: student research workshop* / Florence, Italy, 2019. P. 357–362.
11. Almandouh M.E., Alrahmawy M. F., Eisa M., Tolba A.S. Ensemble Based Low Complexity Arabic Fake News Detection. *International Journal of Intelligent Systems and Applications in Engineering*. 2023. V. 11(2). P. 1022–1031.

**Мищенко Л.Д., Клименко І.А.**

## **РОЗПІЗНАВАННЯ ФЕЙКОВИХ НОВИН ІЗ ВИКОРИСТАННЯМ ОБРОБКИ ПРИРОДНОЇ МОВИ Й АРХІТЕКТУРИ З НИЗЬКИМ ЕНЕРГОСПОЖИВАННЯМ ДЛЯ ПЕРИФЕРІЙНИХ ОБЧИСЛЕНЬ**

*Робота присвячена розробці та застосуванню удосконаленого методу розпізнавання фейкових новин з використанням обробки природної мови (NLP) та архітектури з низьким електроспоживанням для периферійних обчислень. Впроваджено механізм уваги та контекстуальний аналіз для виявлення лінгвістичних ознак достовірності та стилістичних відмінностей між фейковими та правдивими новинами. Такий підхід має на меті забезпечити перевірку новин на периферійних пристроях з обмеженими обчислювальними ресурсами без шкоди для швидкості.*

*Експериментальні дослідження, проведені на різних джерелах новин, підтверджують високу ефективність запропонованого методу. Фейкові новини виявлені за вищим використанням емоційно забарвлених слів, імовірнісних висловлювань та нецензурної лексики порівняно з правдивими новинами.*

*Інтеграція механізму уваги та контекстуального аналізу дозволяє досягти точності виявлення фейкових новин на рівні 81%. Результати цієї роботи відкривають шлях для подальших досліджень у напрямку розуміння та виявлення складних лінгвістичних особливостей у текстах новин, спрямованих на боротьбу з дезінформацією у суспільстві.*

**Ключові слова:** *фейкові новини, обробка природної мови (NLP), контекстуальний аналіз, механізм уваги, малопотужні пристрої, архітектура з низьким електроспоживанням.*



**Mishchenko L.D., Klymenko I.A.**

**RECOGNITION OF FAKE NEWS USING NATURAL LANGUAGE PROCESSING AND A LOW-POWER ARCHITECTURE FOR EDGE COMPUTING**

*The research focuses on enhancing the method of fake news detection based on natural language processing (NLP) utilising low-power edge computing architecture through attention mechanisms and contextual analysis. An attention mechanism and contextual analysis are implemented to detect linguistic signs of credibility and stylistic differences between fake and real news. This approach aims to enable news verification on peripheral devices with limited computational resources without compromising speed.*

*Experimental studies validate the efficiency of the proposed method in identifying anomalies in text frequency and credibility markers. Fake news is identified by the higher use of emotionally charged words, probative statements and profanity compared to real news.*

*The integration of attention mechanisms and contextual analysis showcases a notable improvement in identifying linguistic anomalies typical of fake news, achieving a classification accuracy of 81%. The findings contribute to combating misinformation by leveraging linguistic nuances and signify potential advancements in news veracity assessment on resource-constrained devices.*

**Keywords:** *fake news, natural language processing (NLP), contextual analysis, attention mechanism, edge computing, low-power architecture.*