

*Ю.В. Крак, В.О. Дідур, М.О. Молчанова, О.В. Мазурець, О.В. Собко, О.О. Залуцька, О.В. Бармак*

## МЕТОД ВИЯВЛЕННЯ ПОЛІТИЧНОЇ ПРОПАГАНДИ В ІНТЕРНЕТ-КОНТЕНТІ НЕЙРОМЕРЕЖЕВИМИ ЗАСОБАМИ ОБРОБКИ ПРИРОДНОЇ МОВИ

Автоматизація процесів виявлення пропаганди в текстовому інтернет-контенті засобами обробки природної мови є надзвичайно актуальною в сучасних умовах і здатна забезпечити швидке і своєчасне цільове виявлення ворожого маніпулятивного впливу в масштабних обсягах інтернет-контенту. В статті запропоновано метод автоматизованого виявлення пропаганди, який працює з українською мовою. Метод виявлення політичної пропаганди в інтернет-контенті нейромережевими засобами обробки природної мови призначений для виявлення та аналізу потенційно пропагандистського або маніпулятивного контенту, що поширюється в мережі Інтернет. Вхідними даними методу є ансамбль навчених моделей рекурентних нейронних мереж з токенизаторами і текстовий допис для аналізу. Вихідними даними є рівень і відсоткова оцінка наявності пропаганди як за кожною нейромережевою моделлю ансамбля, так і узагальнено.

Для дослідження ефективності розробленого методу виявлення політичної пропаганди в текстовому інтернет-контенті засобами обробки природної мови, що містить ансамблеве використання рекурентних нейромережових моделей архітектур BiLSTM та GRU, було створено програмну реалізацію методу засобами мови Python. Програмна реалізація дозволяє здійснювати навчання моделей нейронних мереж та використовувати їх для виявлення політичної пропаганди в текстовому інтернет-контенті. Виконано підготовку навчального україномовного набору даних.

Проведено тестове навчання ансамблю класифікаторів із нейромережових архітектур BiLSTM та GRU. Запропонований підхід, здатний визначати політичну пропаганду ансамблем моделей RNN з показниками Accuracy 0.97, Precision 0.973, Recall 0.981 і F1 0.976 за дискретного підходу (беггінг), та Accuracy 0.95, Precision 0.977, Recall 0.987 та F1 0.981 при бінарному підході (стекінг). Розроблений метод має обмеження: працює з текстовими дописами довжиною від 200 до 6300 символів. Для коротших та довших текстів спостерігається погіршення продуктивності.

Ключові слова: пропаганда, рекурентні нейронні мережі, ансамбль нейронних мереж, обробка природної мови.

*I. Krak, V. Didur, M. Molchanova, O. Mazurets, O. Sobko, O. Zalutska, O. Barmak*

## METHOD FOR POLITICAL PROPAGANDA DETECTION IN INTERNET CONTENT USING NEURAL NETWORK NATURAL LANGUAGE PROCESSING TOOLS

The automation of propaganda detection processes in textual Internet content using natural language processing is extremely relevant in modern conditions and can provide fast and well-timed targeted detection of hostile manipulative influence in largescale amounts of Internet content. The paper proposes a method of automated propaganda detection that operates in the Ukrainian language. The method for detecting political propaganda in Internet content using neural network natural language processing tools is intended to identify and analyze potentially propagandistic or manipulative content spread on the Internet. The input data of the method is an ensemble of trained models of recurrent neural networks with tokenizers and a text message for analysis. The output data are the level and percentage of propaganda presence for each neural network model of ensemble and in general.

To examine the effectiveness of developed method for detecting political propaganda in textual Internet content using natural language processing, which includes the ensemble use of recurrent neural network models of the BiLSTM and GRU architectures, a software implementation of the method was created using Python. The software implementation allows training neural network models and using them to detect political propaganda in textual Internet content. The training data set in Ukrainian was prepared.

The test training of an ensemble of classifiers based on the BiLSTM and GRU neural network architectures was conducted. The proposed approach is capable of detecting political propaganda by an ensemble of RNN models with Accuracy 0.97, Precision 0.973, Recall 0.981, and F1 0.976 in the discrete approach (bagging),

© Ю.В. Крак, В.О. Дідур, М.О. Молчанова, О.В. Мазурець, О.В. Собко, О.О. Залуцька, О.В. Бармак, 2024

and Accuracy 0.95, Precision 0.977, Recall 0.987, and F1 0.981 in the binary approach (stacking). The developed method has a limitation: it works with text posts from 200 to 6300 characters long. For shorter and longer texts, performance degradation is observed.

Keywords: propaganda, recurrent neural networks, ensemble of neural networks, natural language processing.

## Вступ

Пропаганда є невід'ємним складником інформаційних маніпуляцій і включає різноманітні форми, методи і засоби впливу на людей з метою зміни їхніх психологічних настроїв у бажаному напрямку. Тому її своєчасне виявлення є актуальною задачею інформаційних технологій. Такі маніпуляції часто використовуються для зміни психологічного клімату в суспільстві, мобілізації підтримки або дискредитації опонентів [1].

Пропаганда, яка розповсюджується в мережі Інтернет, представляє масштабну загрозу для національної безпеки країни. Несвоєчасне запобігання їй може призвести до руйнівних наслідків [2]. Тому автоматизація процесів виявлення пропаганди у текстовому інтернет-контенті засобами обробки природної мови є надзвичайно актуальною у сучасних умовах і здатна забезпечити швидко та своєчасне цільове виявлення ворожого маніпулятивного впливу в масштабних обсягах інтернет-контенту.

### 1. Аналіз останніх досліджень і публікацій

Наукові публікації вітчизняних вчених різнобічно висвітлюють актуальність проблеми автоматизованого виявлення пропаганди в текстовому інтернет-контенті. Особливо актуальними наразі є напрямки досліджень, присвячені інтегруалізації процесів виявлення пропаганди, що дозволяє уникнути низки технологічних проблем щодо моніторингу медіаджерел [3] та проблемі відділення проявів технік пропаганди від інших маніпулятивних впливів [4]. Відзначається, що елементи моделі пропаганди включають суб'єкт, зміст, форми і методи, а також засоби або канали передачі інформації.

Суб'єктом пропаганди є соціальна група, яка прагне впливати на аудиторію. Зміст пропаганди визначається соціальни-

ми інтересами суб'єкта та їх відношенням до інтересів суспільства загалом. Форми і методи пропаганди вибираються залежно від цілей та аудиторії, на яку має здійснюватись вплив. Засоби передачі інформації включають друковані видання, радіо, телебачення тощо. Об'єктом пропаганди є аудиторія або соціальні групи, які є метою впливу. Соціальні інтереси суб'єкта пропаганди впливають на її зміст і вибір форм, методів і засобів передачі інформації [5].

Виявлення пропаганди за допомогою NLP у тексті є складною задачею через використання в пропаганді тонких методів маніпулювання та контекстуальних залежностей. Для вирішення цієї проблеми, автори [6] досліджували ефективність сучасних великих мовних моделей, таких як GPT-3 і GPT-4 для виявлення пропаганди. Експерименти здійснювалися із використанням набору даних SemEval-2020 task 11, який містить статті новин, позначені 14 пропагандистськими техніками. Отримані результати показують, що GPT-4 досягає результатів, порівняних із поточним сучасним рівнем розвитку технологій.

Автори [7] відзначають, що існуючі методи виявлення пропаганди зосереджені насамперед на виявленні мовних особливостей її змісту. Однак ці методи зазвичай пропускають інформацію, представлену в зовнішньому новинному середовищі, з якого виникли та поширилися пропагандистські новини. Авторами [8] аналізується, як ЗМІ вплинули та відобразили громадську думку протягом першого місяця російського вторгнення за допомогою статей і каналів новин у Telegram українською, російською, румунською, французькою та англійською мовами. Було запропоновано та порівняно два методи багатомовної автоматизованої ідентифікації прокремлівської пропаганди, засновані на трансформерах (BERT) і лінгвістичних ознаках (SVM).

Метою статті є створення методу виявлення політичної пропаганди в текстовому інтернет-контенті нейромережеви ми засобами обробки природної мови, який буде працювати з українською мовою, а також його апробація.

У рамках дослідження також було виконано такі завдання: підготовка навчального україномовного набору даних; розробка програмного забезпечення, що реалізовує створений метод; навчання ансамблю нейромережевих класифікаторів; дослідження ефективності методу з використанням розробленого програмного забезпечення.

Основним внеском статті є розробка дієвого методу автоматизованого виявлення політичної пропаганди в україномовних текстах.

## 2. Метод та матеріали

Враховуючи недостатню кількість україномовних даних, є потреба у створенні власного розміченого набору даних, що буде використовуватись для навчання нейромереж.

### 1.1. Підготовка набору даних.

Для навчання моделей рекурентних нейронних мереж було сформовано набір даних з понад 25000 дописів, які були розмічені відповідно приналежності категоріям «Пропаганда» та «Не пропаганда». Переліки пропагандистських та верифікованих джерел було сформовано згідно офіційних каналів Президента й Верховної Ради України, а також за даними аналітичних міжнародних авторитетних досліджень [9] та аналітичних зведень [10].

Для нормалізації вхідних даних було відкинуто записи довжиною менше 200 і більше 6300 символів. У результаті фільтрації даних отримано набір, що складається із 21 222 елементів, де 10737 записів належать класу «пропагандистський допис» та 10485 записів класу «допис без пропаганди». Отож, кількість записів, які не місять пропаганди і знаходяться в діапазоні довжини 200..800 символів становить більшу половину набору. Це може негативно вплинути на якість класифікації у майбутньому. Водночас набір пропаганди-

стських текстів рівномірніше розподілений. Всі багатомовні фрагменти було автоматично перекладено українською мовою.

1.2. **Схема методу виявлення політичної пропаганди** наведена на рис. 1. Вхідними даними методу є ансамбль навчених моделей рекурентних нейронних мереж з токенизаторами і текстовий допис для аналізу. На кроці 1 відбувається вибір і завантаження ансамблю RNN-моделей, а також їхніх токенизаторів.

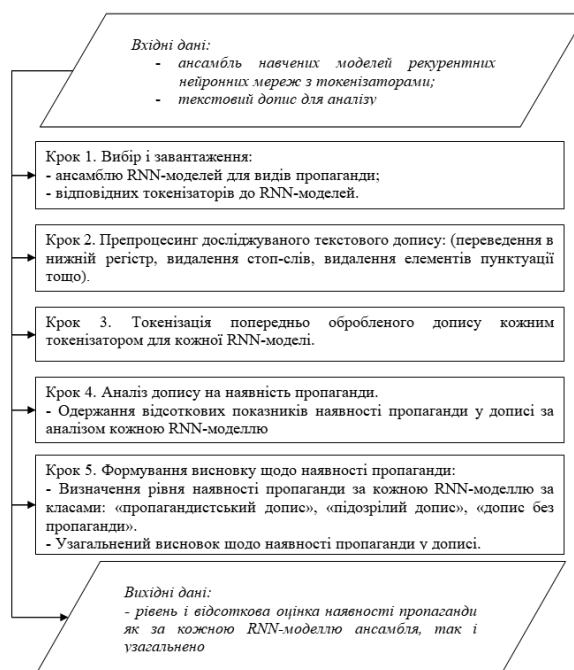


Рис. 1. Схема методу виявлення пропаганди

Наступним кроком 2 здійснюється попередня обробка користувацького допису для аналізу, що включає в себе перетворення тексту у нижній регістр, видалення стоп-слів та елементів пунктуації тощо.

На кроці 3 попередньо оброблений текст перетворюється у числові послідовності, які будуть подані нейронним мережам на вхід для подальшої бінарної класифікації. Кроком 4 є аналіз допису на наявність пропаганди, що включає в себе одержання відсоткових показників наявності пропаганди у дописі за аналізом кожної RNN-моделлю.

На кроці 5 здійснюється формування висновку щодо наявності пропаганди. Пропонується використати два підходи –

бінарний (стекінг) та дискретний (беггінг). Для бінарного підходу для визначення рівня пропаганди для нейромереж ансамблю отримуються бінарні оцінки, де оцінка 0 – не містить пропаганди, 1 – містить пропаганду. У дискретному підході оцінка нейромереж береться як дискретна величина з проміжку від 0 до 1, де 1 – максимальний прояв пропаганди, а 0 – її відсутність.

У випадку стекінгу відбувається одержання бінарної оцінки, й висновок щодо класу допису формується за правилами: «пропагандистський допис», якщо понад 50% моделей отримали бінарні оцінки 1; «допис без пропаганди», якщо понад 50% моделей отримали бінарні оцінки 0; «підозрілий допис», якщо нейромережеві моделі мають паритетні результати голосування (близько половини з оцінками 0 і половини з оцінками 1).

Для визначення рівня пропаганди у випадку дискретної оцінки, експертним шляхом устанавлюються межі трьох класів: верхня межа класу «допис без пропаганди» та нижня межа класу «пропагандистський допис».

Після чого здійснюється розрахунок загальної дискретної оцінки приналежності допису до вказаних класів (1):

$$Eval = k_1 \cdot RNN_1 + k_2 \cdot RNN_2 + \dots + k_n \cdot RNN_n \quad (1)$$

де  $k_1, k_2, \dots, k_n$  – коефіцієнти впливу дискретних оцінок, отриманих нейромережами  $RNN_1, RNN_2, \dots, RNN_n$  відповідно.

Коефіцієнти впливу дискретних нейромережевих оцінок  $k_1, k_2, \dots, k_n$  обираються емпіричним шляхом в залежності від фокусування процесу на виявленні пропаганди відповідних видів.

Відповідно до вищевикладеного матеріалу, результатом роботи запропонованого методу є рівень і відсоткова оцінка наявності пропаганди за кожною RNN-моделлю ансамбля, а також узагальнені рівень і відсоткова оцінка наявності пропаганди у досліджуваному дописі.

### 3. Експерименти

Для проведення експерименту з дослідження ефективності розробленого ме-

тоду виявлення пропаганди в інтернет-контенті було сформовано ансамбль із двох нейромережевих моделей. Зокрема, було використано рекурентні нейронні мережі архітектур BiLSTM та GRU. Підбір різних нейромережевих моделей обумовлений їхніми специфічними можливостями з аналізу текстових послідовностей.

BiLSTM шляхом використання прихованого стану дозволяє аналізувати текстові послідовності у прямому та зворотному напрямках, що допомагає усунути бар'єри традиційних RNN. GRU має механізми воріт, які дозволяють ефективніше управляти градієнтами в часі, що робить її більш стійкою до проблеми зникнення градієнтів порівняно з класичними RNN.

У випадку використання архітектур BiLSTM та GRU для проведення експерименту, формула (1) набуде вигляду:

$$Eval = k_1 \cdot BiLSTM_r + k_2 \cdot GRU_r, \quad (2)$$

де  $k_1$  – коефіцієнт впливу дискретної оцінки отриманої нейромережею BiLSTM,  $k_2$  – коефіцієнт впливу дискретної оцінки отриманої нейромережею GRU,  $BiLSTM_r$  та  $GRU_r$  – дискретні оцінки виявлення пропаганди нейромережами BiLSTM та GRU відповідно.

Під час дослідів нейромережі навчались із різними параметрами (batch, epoch), результати порівняння найкращих моделей наведені у Табл.1.

Таблиця 1

Залежність метрик від параметрів нейромереж

| Параметри: | GRU  |      | BiLSTM |      |
|------------|------|------|--------|------|
| batch      | 32   | 64   | 32     | 64   |
| epoch      | 20   | 20   | 20     | 20   |
| Метрики:   |      |      |        |      |
| Точність   | 0.97 | 0.96 | 0.96   | 0.95 |
| Втрати     | 0.04 | 0.06 | 0.04   | 0.07 |

Як видно з Табл. 1, GRU має вищу точність, ніж BiLSTM, за однакових параметрів. На Рис. 2 наведено розподіл коректно класифікованих текстів нейромережею GRU, а на Рис. 3 – розподіл некоректно класифікованих текстів.

За валідаційні дані було взято 3573 записи, з яких до класу «пропагандистсь-



кий допис» належало 1951, а до класу «допис без пропаганди» – 1622. З них коректно було класифіковано 1912 класу «пропагандистський допис» та 1565 текстів класу «допис без пропаганди». 57 текстів класу «допис без пропаганди» нейромережею хибно класифіковано як пропаганда, а 39 текстів класу «пропагандистський допис» хибно класифіковані як не пропагандистські. Загальна точність на валідаційних даних становить 0.97. Як видно з числових даних, клас «допис без пропаганди» класифікується дещо гірше, ніж клас «пропагандистський допис».

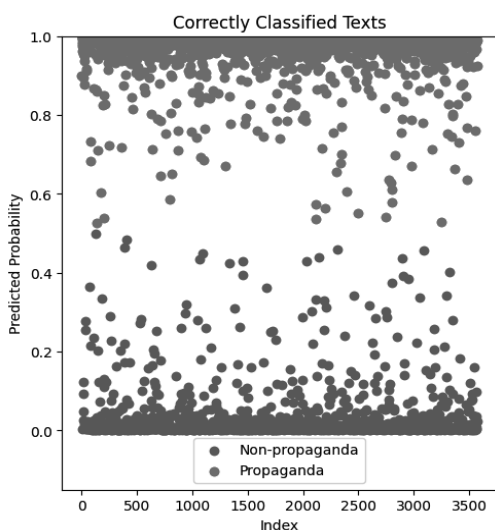


Рис. 2. Розподіл коректно класифікованих текстів нейромережею GRU

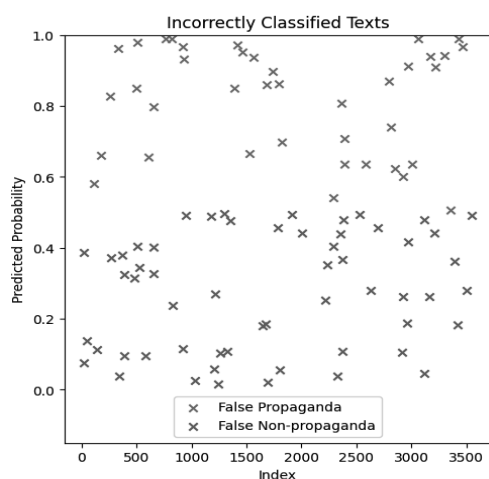


Рис. 3. Розподіл некоректно класифікованих текстів нейромережею GRU

На Рис. 4 наведено розподіл коректно класифікованих текстів нейромережею BiLSTM, а на Рис. 5 – розподіл некоректно класифікованих текстів.

З 3573 валідаційних дописів коректно класифіковано 1883 дописи класу «пропагандистський допис» та 1572 тексти класу «допис без пропаганди». 86 текстів класу «допис без пропаганди» нейромережею хибно класифіковано як пропаганда, а 32 тексти класу «пропагандистський допис» хибно класифіковані як не пропагандистські. Загальна точність на валідаційних даних становить 0.967.

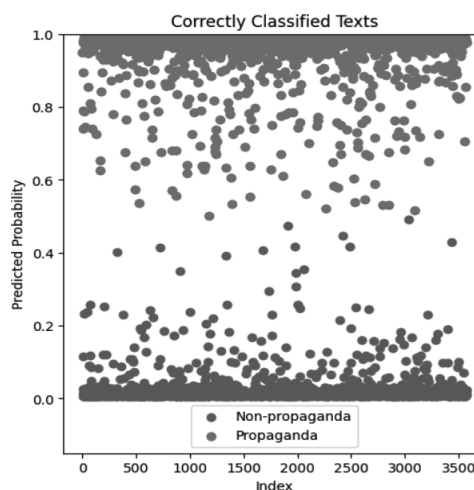


Рис. 4. Розподіл коректно класифікованих текстів нейромережею BiLSTM

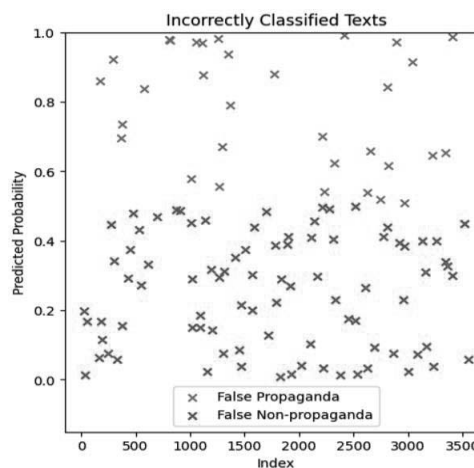


Рис. 5. Розподіл некоректно класифікованих текстів нейромережею BiLSTM

Як бачимо з Рис.3 та Рис.5, тексти мають достатньо високий рівень міжкласової роздільності, тоді, як видно на Рис.4 та Рис.6, некоректно класифіковані дані зосереджені ближче до центральної частини графіків, що свідчить про доцільність підходу з розбиттям на 3 класи: «пропагандистський допис», «допис без пропаганди», «підозрілий допис».

#### 4. Практичне застосування

Для дослідження ефективності розробленого методу, що містить ансамблеве використання RNN-моделей архітектур BiLSTM та GRU, було створено програмну реалізацію методу засобами мови Python. Інтерфейс програмної частини, що відповідає за процес виявлення пропаганди розробленим методом, наведено на Рис. 6.

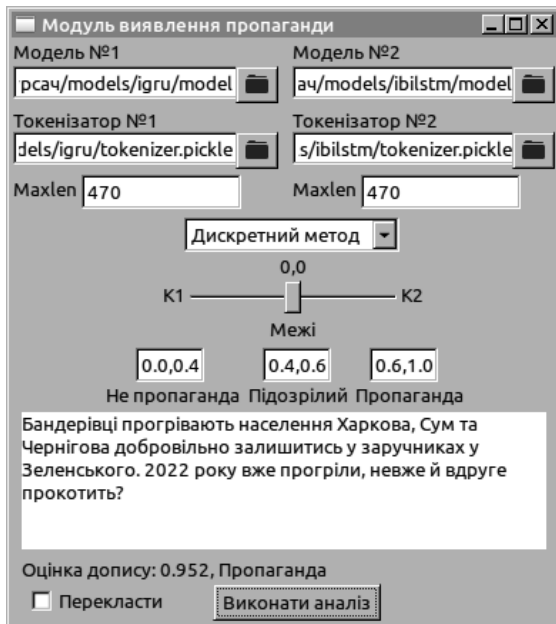


Рис. 6. Модуль виявлення політичної пропаганди

Із запровадженням категорії «підозрілий допис», відсоток помилок першого та другого роду зменшився. У разі використання бінарного підходу, некоректно класифікованими виявилось 178 зразків із 3573. Однак із 178 зразків 71 зразок є хибним, 107 визначено підозрілими. З 71 хибного зразка 26 текстів хибно віднесено до класу «допис без пропаганди», і 45 текстів хибно віднесено до класу «пропагандистський допис». Щодо результатів дискретного підходу, то некоректно класифікованими виявились 130 зразків, що в цілому не погіршило статистику показників нейромережі GRU, але зі 130 зразків 37 текстів хибно віднесено до класу «допис без пропаганди» і 52 тексти хибно віднесено до класу «пропагандистський допис».

#### 5. Результати та дискусія

Для дослідження ефективності виявлення політичної пропаганди в тексто-

вому інтернет-контенті розробленим методом було використано метрики Accuracy, Precision, Recall та F1 [11]. Значення метрик для дискретного та бінарного варіацій методу наведено в Табл.2. Хоча бінарний підхід дав гірші результати за метрикою Accuracy, однак дав кращі показники метрик Precision, Recall та F1, тоді як дискретний підхід Accuracy практично не погіршив, але водночас метрики Precision, Recall та F1 дещо йому поступаються.

Таблиця 2

Значення метрик за беггінгу та стекінгу

| Підхід  | Accuracy | Precision | Recall | F1    |
|---------|----------|-----------|--------|-------|
| Беггінг | 0.97     | 0.973     | 0.981  | 0.976 |
| Стекінг | 0.95     | 0.977     | 0.987  | 0.981 |

Для експерименту параметри дискретного підходу були такими:  $k_1=0.5$ ,  $k_2=0.5$ ,  $l_2 = 0.45$ ,  $l_4 = 0.55$ .

#### Висновки

Запропоновано метод виявлення політичної пропаганди в текстовому інтернет-контенті нейромережевими засобами обробки природної мови, який працює з українською мовою, а також здійснена його апробація.

У межах дослідження виконано підготовку навчального україномовного набору даних, здійснено тестове навчання ансамблю класифікаторів із нейромережових архітектур BiLSTM та GRU, розроблено програмне забезпечення, що імплементує створений метод та проведено дослідження його ефективності.

Запропонований підхід здатний визначати політичну пропаганду ансамблем моделей RNN з показниками Accuracy 0.97, Precision 0.973, Recall 0.981 і F1 0.976 при дискретному підході (беггінг), та Accuracy 0.95, Precision 0.977, Recall 0.987 та F1 0.981 у разі бінарного підходу (стекінг).

Розроблений метод має обмеження: працює з текстовими дописами довжиною від 200 до 6300 символів. Для коротших та довших текстів спостерігається погіршення продуктивності.

Подальші дослідження будуть спрямовані на аналіз залежності розглянутих показників ефективності запропонованого

методу від ознак і параметрів аналізованого допису, таких як розмір, жанр та тематика. Перспективним напрямком продовження досліджень також є збільшення кількості RNN-моделей у ансамблі для покращення показників ефективності і спеціалізація моделей по окремих видах пропаганди.

### Література

1. В. В. Посмітна, Особливості сучасної антиукраїнської маніпулятивної пропаганди в текстах і заголовках російських інформаційних повідомлень, Сучасні умови. Науковий вісник Міжнародного гуманітарного університету, 2023. №59. С. 35–39.
2. І. А. Колеснікова, Проблеми протидії пропаганді та дезінформації у відкритих джерелах інформаційно-телекомунікаційної мережі інтернет. Актуальні проблеми протидії злочинності та корупції: збірник тез Всеукраїнської науково-практичної конференції, 2023. С. 88–93.
3. М. А. Польовий, Технологічні проблеми автоматичного виявлення проросійської пропаганди в мережі Facebook. Актуальні проблеми управління інформаційною безпекою держави: зб. тез наук. доп. наук.-практ. конф. Національна академія СБУ. Київ, 2021. С. 96-98.
4. Я. В. Тарасенко, Основні проблеми комп'ютеризованого виявлення відмінностей між інформаційним впливом та пропагандою. Сучасний рух науки: тези XII міжнародної науково-практичної конференції, 2021. С. 380-380.
5. Пропаганда. Accessed: 01.04.2024. <https://vue.gov.ua/Пропаганда>
6. K. Sprenkamp, D. J. Gordon, L. Zavolokina, Large Language Models for Propaganda Detection. *Computation and Language*, 2023.
7. X. Liu, K. Ma, K. Ji, Zh. Chen, B. Yang, Graph-based multi-information integration network with external news environment perception for Propaganda detection, *International*

*Journal of Web Information Systems*, 2024. Vol. 20, Issue 2. pp. 195–212.

8. V. Solopova, O. I. Popescu, C. Benzmüller, Automated Multilingual Detection of Pro-Kremlin Propaganda in Newspapers and Telegram Posts, *Datenbank Spektrum*, 2023. №23. pp. 5–14.
9. Russia Today's Disinformation Campaign. Accessed: 02.04.2024. <https://blogs.state.gov/stories/2014/04/29/russia-today-s-disinformation-campaign>.
10. Canada sanctions Russian propagandists, singers, actors, musicians, and Wagner Group media. Accessed: 02.04.2024. <https://english.nv.ua/life/canada-sanctions-russian-propagandists-singers-actors-musicians-and-wagner-group-media-50302091.html>.
11. O. Zalutska, M. Molchanova, O. Sobko, O. Mazurets, O. Pasichnyk, O. Barmak, I. Krak, Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network. *CEUR Workshop Proceedings*, 2023. pp. 561–571.

### References

1. V. V. Posmitna, Features of anti-ukrainian manipulative propaganda in: Modern conditions. *Scientific Bulletin of the International Humanitarian University*, 2023 [in Ukrainian].
2. I. A. Kolesnikova, Problems of countering propaganda and disinformation in open sources of the Internet information and telecommunications network, in: *Conference Organizing Committee 88*, 2023. [in Ukrainian].
3. M. A. Polevyi, Technological problems of automatic detection of pro-Russian propaganda in the Facebook network. *Actual problems of state information security management: coll. theses of sciences add. science and practice conf. NA SBU. Kyiv*, 2021. P. 96-98. [in Ukrainian].
4. Y. V. Tarasenko, Main problems of computerized detection of differences

- between information influence and propaganda. Modern movement of science: theses XII International Scientific and Practical Conference, 2021. P. 380-380. [in Ukrainian].
5. Propaganda. [cited 1 Apr 2024]. <https://vue.gov.ua/Пропаганда> [in Ukrainian].
  6. K. Sprenkamp, D. J. Gordon, L. Zavolokina, Large Language Models for Propaganda Detection, in: Computation and Language, 2023. doi: 10.48550/arXiv.2310.06422.
  7. X. Liu, K. Ma, K. Ji, Zh. Chen, B. Yang, Graph-based multi-information integration network with external news environment perception for Propaganda detection, in: International Journal of Web Information Systems (2024).
  8. V. Solopova, OI. Popescu, C. Benzmüller, Automated Multilingual Detection of Pro-Kremlin Propaganda in Newspapers and Telegram Posts, in: Datenbank Spektrum (2023), 5–14. doi: 10.1007/s13222-023-00437-2.
  9. Russia Today’s Disinformation Campaign. [cited 2 Apr 2024]. <https://blogs.state.gov/stories/2014/04/29/russia-today-s-disinformation-campaign>
  10. Canada sanctions Russian propagandists, singers, actors, musicians, and Wagner Group media. [cited 2 Apr 2024]. <https://english.nv.ua/life/canada-sanctions-russian-propagandists-singers-actors-musicians-and-wagner-group-media-50302091.html>
  11. O. Zalutska, M. Molchanova, O. Sobko, O. Mazurets, O. Pasichnyk, O. Barmak, I. Krak, Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network, in: CEUR Workshop Proceedings, 2023, pp. 561–571.

Одержано: 12.02.2024

Внутрішня рецензія отримана: 19.02.2024

Зовнішня рецензія отримана: 08.03.2024

### **Про авторів:**

<sup>1,2</sup>*Крак Юрій Васильович*,  
доктор фізико-математичних наук,  
професор.  
<http://orcid.org/0000-0002-8043-0785>.

<sup>3</sup>*Дідур Володимир Олександрович*,  
студент.  
<http://orcid.org/0009-0004-3149-3781>.

<sup>3</sup>*Молчанова Марина Олексіївна*,  
викладач.  
<http://orcid.org/0000-0001-9810-936X>.

<sup>3</sup>*Мазурець Олександр Вікторович*,  
кандидат технічних наук, доцент.  
<http://orcid.org/0000-0002-8900-0650>.

<sup>3</sup>*Собко Олена Віталіївна*, викладач.  
<http://orcid.org/0000-0001-5371-5788>.

<sup>3</sup>*Залуцька Ольга Олександрівна*,  
студент.  
<http://orcid.org/0000-0003-1242-3548>.

<sup>3</sup>*Бармак Олександр Володимирович*,  
доктор технічних наук, професор.  
<http://orcid.org/0000-0003-0739-9678>.

### **Місце роботи авторів:**

<sup>1</sup>Київський національний  
університет імені Тараса Шевченка  
Тел. (+38) (044) 239-33-29  
E-mail: [info@knu.ua](mailto:info@knu.ua)  
Сайт: [knu.ua](http://knu.ua)

<sup>2</sup>Інститут кібернетики імені  
В.М. Глушкова НАН України  
Тел. (+38) (044) 526-20-08  
E-mail: [incyb@incyb.kiev.ua](mailto:incyb@incyb.kiev.ua)  
Сайт: [incyb.kiev.ua](http://incyb.kiev.ua)

<sup>3</sup>Хмельницький національний  
університет,  
Тел. (+38) (0382) 67-02-76  
E-mail: [centr@khmn.edu.ua](mailto:centr@khmn.edu.ua)  
Сайт: [khmn.edu.ua](http://khmn.edu.ua)