

<https://doi.org/10.31891/2219-9365-2024-80-19>

УДК 004.8

МОЛЧАНОВА Марина
Хмельницький національний університет
<https://orcid.org/0000-0001-9810-936X>
m.o.molchanova@gmail.com

НЕЙРОМЕРЕЖЕВЕ ВИЯВЛЕННЯ І КЛАСИФІКАЦІЯ ПРИЙОМІВ ТА ОБ'ЄКТІВ ПРОПАГАНДИ У ТЕКСТОВОМУ КОНТЕНТІ

Запропоновано підхід до нейромережевого виявлення і класифікації прийомів та об'єктів пропаганди у текстовому контенті, що складається з трьох послідовних етапів та забезпечує ефективне виявлення наявності пропаганди, класифікацію використаних технік пропаганди, та встановлення об'єктів виявленого пропагандистського впливу. Етап класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання дозволяє виявляти як явні, так і приховані пропагандистські меседжі, що забезпечує більш глибоке розуміння послідовності та контексту в текстовому контенті та дозволяє досягнути точності 97.83%. Етап виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень дозволяє перетворювати вхідні дані у вигляді тексту для аналізу та навчених окремим технікам 17 моделей машинного навчання у вихідні дані, які містять числові оцінки наявності кожної з технік пропаганди із візуальною аналітикою присутності детектованих маркерів пропаганди, що забезпечує виявлення різних пропагандистських технік з мінімальною точністю 82,03%. Етап виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень характеризується розширенням множини об'єктів пропаганди за рахунок додавання варіантів їх словесних подань і використанням контекстних вікон для виявлення взаємозв'язків між використаними прийомами та об'єктами пропаганди, що дало можливість виявлення об'єктів пропаганди та візуально їх інтерпретувати. Для візуальної інтерпретації одержаних результатів, формується візуальна аналітика щодо знайдених технік та об'єктів пропаганди, що дозволяє візуально спостерігати об'єкти впливу в рамках використовуваних технік пропаганди. Внаслідок розробки підходу, вирішено ряд проблем в напрямку автоматизації виявлення пропаганди, таких як відсутність комплексного аналізу взаємозв'язків прийомів і об'єктів пропаганди в текстах, та відсутність узагальнень для об'єктів пропаганди і їх альтернативних згадок у текстах.

Ключові слова: візуальна інтерпретація, прийоми пропаганди, об'єкти пропаганди, маркери пропаганди.

MOLCHANOVA Maryna
Khmelnitskyi National University

NEURAL NETWORK DETECTION AND CLASSIFICATION OF PROPAGANDA TECHNIQUES AND OBJECTS IN THE TEXT CONTENT

Approach to neural network detection and classification of techniques and objects of propaganda in text content is proposed, which consists of three consecutive stages and provides effective detection of propaganda presence, classification of used propaganda techniques, and establishment of objects of detected propaganda influence. The step of classifying texts by propaganda content by neural network models of deep learning allows the detection of both explicit and hidden propaganda messages, based on the combination of traditional recurrent neural networks with long-term memory with transformative neural networks, which provides a deeper understanding of sequence and context in textual content and allows to achieve accuracy of 97.83%. The stage of detecting propaganda techniques by markers with a visual interpretation of the decisions made allows you to transform input data in the form of text for analysis and 17 machine learning models trained by individual technicians into output data that contain numerical estimates of the presence of each of the propaganda techniques and marked text with visual analytics of the presence of detected markers propaganda, which provides the detection of various propaganda techniques with a minimum accuracy of 82.03%, which is better than known analogues in terms of detection propaganda regardless of the methods used. The stage of identifying objects of propaganda by neural network models of deep learning with visual interpretation of the decisions made is characterized by the expansion of the set of objects of propaganda due to the addition of variants of their verbal representations and the use of context windows to identify relationships between the techniques used and objects of propaganda, which made it possible improve the results of detection of propaganda objects and visually interpret them. Significant problems in the direction of identifying propaganda have been solved, namely: the lack of a comprehensive analysis of the interrelationships of techniques and objects of propaganda in texts; lack of generalizations for objects of propaganda and their alternative mentions in texts. For the visual interpretation of the obtained results, a visual analytics is formed regarding the found techniques and objects of propaganda, which allows you to visually observe the objects of influence within the framework of the used propaganda techniques.

Keywords: visual interpretation, propaganda techniques, propaganda objects, propaganda markers.

ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

На сучасному етапі інтернет-технології створюють сприятливі умови для все більш швидкого, масового та ефективного поширення пропаганди. Нові методи генерації текстів забезпечують стрімке зростання кількості контенту, тому виникає потреба у постійному моніторингу нових способів створення пропагандистського вмісту й удосконалення методів ідентифікації пропагандистських маніпуляцій, що є важливою задачею протидії дезінформації та інформаційної безпеки [1]. При цьому, виявлення політичної пропаганди у медіапросторі доводиться виконувати вручну, часто без аналізу взаємозв'язку між

використаними прийомами та обраними об'єктами пропаганди, що суттєво знижує оперативність та якість результатів. Тому наразі є актуальною автоматизація виявлення прийомів і об'єктів пропаганди нейромережевими засобами, їх класифікація та візуальна аналітика одержаних результатів.

У статті пропонується підхід до нейромережевого виявлення і класифікації прийомів та об'єктів пропаганди у текстовому контенті, що складається з трьох послідовних етапів та забезпечує ефективне виявлення наявності пропаганди, класифікацію використаних прийомів пропаганди, та встановлення об'єктів виявленого пропагандистського впливу.

АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Питання своєчасного виявлення пропаганди [2] є актуальними питаннями обробки природної мови [3], що не поступаються сучасним викликам семантичного аналізу, таким як виявлення аб'юзивного змісту [4, 5] та аналізу емоційної тональності [6, 7] тощо. Ідентифікація пропагандистського вмісту включає не тільки виявлення самого вмісту, але й визначення його цільової аудиторії, об'єктів пропаганди. Це є ключовим аспектом розуміння механізмів маніпуляції, тому привертає значну увагу сучасних дослідників.

Однією з труднощів виявлення пропаганди є відсутність надійно маркованих джерел даних. Для ідентифікації пропагандистських матеріалів у текстових новинах у дослідженні [8] була запропонована бібліотека пропагандистських текстів ProText. Мітки достовірності присвоюються текстам у ProText після як ручної, так і автоматичної перевірки за допомогою методів фактчекінгу. Дослідники використали методи обробки природної мови для розробки системи, яка застосовує глибоке навчання для автоматизованого виявлення пропаганди в новинних текстах.

На сьогоднішній день у виявленні пропаганди з текстових даних розрізняють такі підходи: на базі виявлення іменованих об'єктів та на базі класифікації текстів.

При розпізнаванні пропаганди як завдання розпізнавання іменованих сутностей (NER) виникає проблема, що пропагандистські фрагменти тексту зазвичай набагато довші, ніж звичайні фрагменти, такі як імена осіб, і можуть складатися з десятків слів [9]. У дослідженні [10] аналізується вплив довжини текстових фрагментів на ефективність розпізнавання пропаганди, показуючи, що з ростом довжини фрагментів складність завдання зростає. Було проведено оцінку декількох поширених підходів до розв'язання цієї задачі, щоб визначити, наскільки добре вони відновлюють розподіл довжини фрагментів. Крім того, запропоновано метод, що включає адаптивний рівень згортки, який покращує обмін інформацією між віддаленими словами. Цей підхід дозволяє зберегти точність обробки довгих фрагментів без втрати загальної продуктивності.

Аналіз літератури у сфері виявлення пропагандистських прийомів та об'єктів пропаганди показав наявність двох основних проблем: відсутність комплексного аналізу взаємозв'язків між прийомами та об'єктами пропаганди в текстах, а також брак узагальнень щодо об'єктів пропаганди та їх альтернативних згадувань. Ідентифікація пропаганди шляхом виявлення лише іменованих сутностей не надає повного розуміння спрямованості пропаганди через використані прийоми [11]; водночас виявлення пропагандистських прийомів на рівні документа не відображає цільових об'єктів пропаганди [12]. Додатково, при вирішенні задачі виявлення пропаганди за допомогою NER виникає проблема, що текстові фрагменти, які містять пропаганду, значно довші за типові для NER сутності і можуть містити значну кількість слів.

ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є розробка підходу до виявлення і класифікація прийомів та об'єктів пропаганди у текстовому контенті з використанням нейромережових засобів, а також забезпечення візуальної аналітики одержаних результатів.

ПІДХІД ДО НЕЙРОМЕРЕЖЕВОГО ВИЯВЛЕННЯ І КЛАСИФІКАЦІЇ ПРИЙОМІВ ТА ОБ'ЄКТІВ ПРОПАГАНДИ

Запропонований підхід складається з трьох послідовних етапів та забезпечує автоматизоване виявлення наявності пропаганди, класифікацію використаних технік пропаганди, встановлення об'єктів виявленого пропагандистського впливу та візуальну аналітику одержаних результатів (рисунок 1).

Вхідними даними підходу є текст для аналізу, навчена модель глибокого навчання для бінарної класифікації (пропаганда / не пропаганда), множина нейромережових моделей глибокого навчання для ідентифікації прийомів пропаганди (окрема нейромережа для кожного з 17 прийомів пропаганди).

На першому етапі відбувається аналіз рівня пропаганди та віднесення до 1 з категорій: «текст без пропаганди», «підозрілий текст», «пропагандистський текст». Віднесення до категорій здійснюється відповідно до отриманої нейромережевої оцінки рівня пропаганди.

Другий та третій етапи відбуваються лише для текстів, що віднесені в категорію «пропагандистський текст». На другому етапі вхідне текстове представлення подається по черзі на 17 навчених нейромережових моделей для аналізу наявності 17 технік. Відповідно, вихідними даними другого

етапу буде нейромережева оцінка наявності технік пропаганди за маркерами [13] та поданням із візуальною інтерпретацією результатів.

Третій етап працює використовуючи проміжні дані попередніх етапів. Здійснює перетворення вхідних даних у множину тематичних об'єктів пропаганди та відповідне візуальне подання.

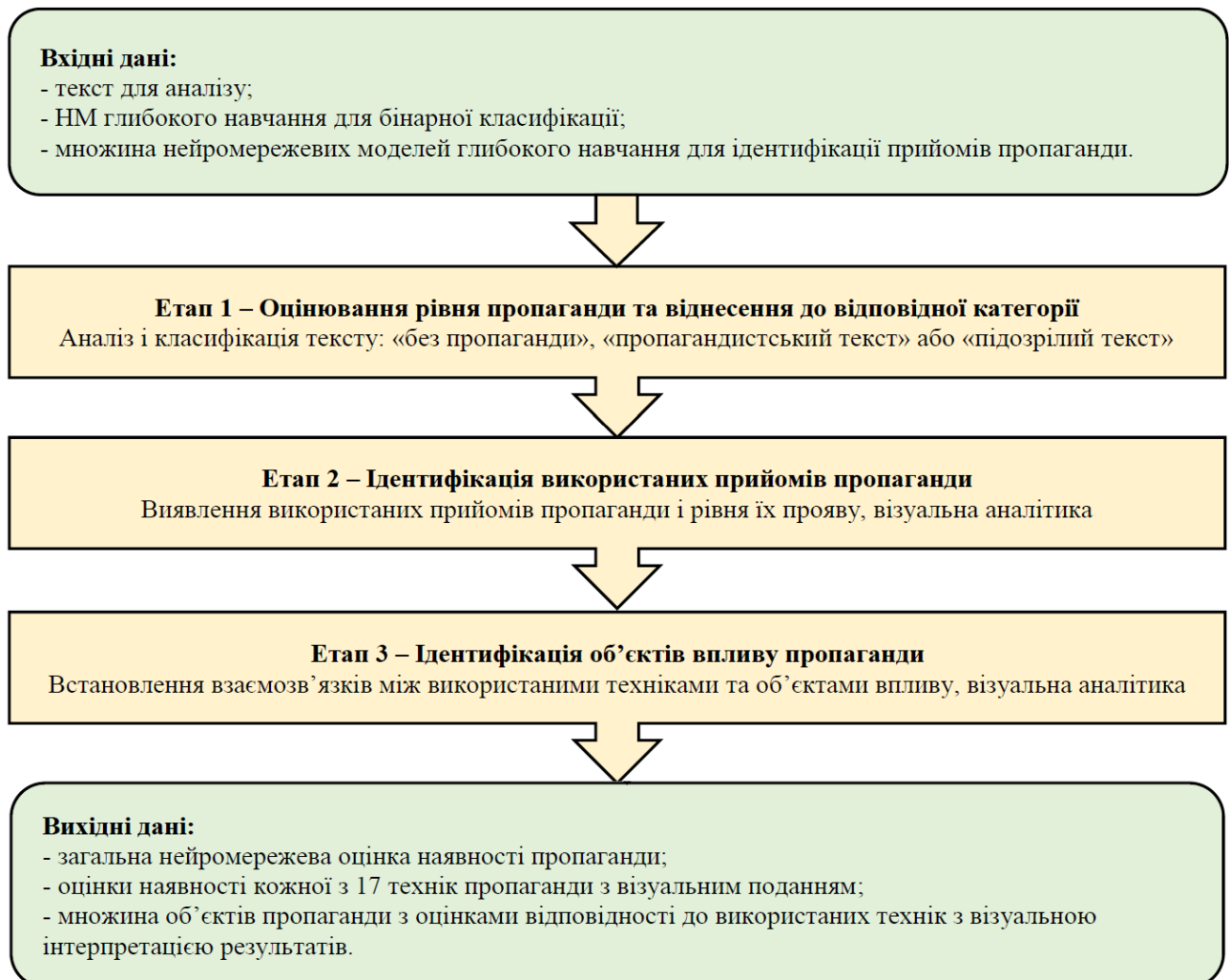


Рис. 1. Підхід до нейромережевого виявлення і класифікації прийомів та об'єктів пропаганди

Таким чином, запропонований підхід до виявлення і класифікації прийомів та об'єктів пропаганди у текстовому контенті з використанням нейромережевих засобів містить етапи класифікації текстів за вмістом пропаганди, виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень та виявлення об'єктів пропаганди з візуальною інтерпретацією прийнятих рішень.

КЛАСИФІКАЦІЯ ТЕКСТІВ ЗА ВМІСТОМ ПРОПАГАНДИ

Етап класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання призначений для автоматизованої ідентифікації текстів, які містять пропагандистські елементи. Особливістю етапу є те, що він дозволяє виявляти як явні, так і приховані пропагандистські меседжі, ґрунтуючись на об'єднанні традиційних рекурентних нейромереж з довгостроковою пам'яттю із трансформерами, а також на використанні механізму аугментації навчальних текстових даних, що дозволяє розширити кількість навчальних зразків. Схему етапу наведено на рисунку 2.

Етап здійснює перетворення вхідних даних у вигляді навченої нейромережевої моделі глибокого навчання та тексту для класифікації у вихідні дані у вигляді відсоткової оцінки наявності пропаганди у тексті та присвоєння 1 з 3-х класів: «текст без пропаганди», «пропагандистський текст» та «підозрілий текст». Межі категорій визначаються емпірично в залежності від потреби точності класифікації, однак за замовчуванням рекомендуються такі межі: "текст без пропаганди" охоплює значення від 0 до 0.45, "підозрілий текст" має значення від 0.45 до 0.55, а "пропагандистський текст" від 0.55 до 1.

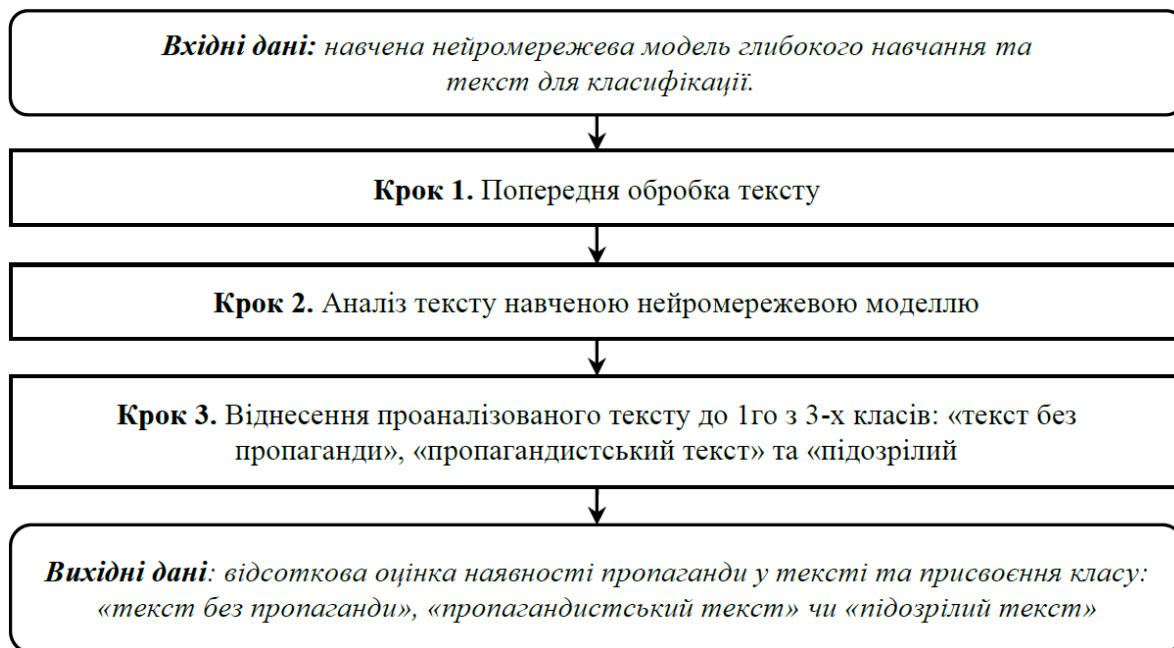


Рис. 2. Кроки для класифікації текстів за вмістом пропаганди неймережевими моделями

Аналіз тексту виконується нейронною мережею глибокого навчання з гібридною архітектурою, яка об'єднує традиційну рекурентну нейронну мережу з довготривалою пам'яттю та трансформера. Це дозволяє досягти більш глибокого розуміння послідовності та контексту в тексті.

ВИЯВЛЕННЯ ПРИЙОМІВ ПРОПАГАНДИ ЗА МАРКЕРАМИ ІЗ ВІЗУАЛЬНОЮ ІНТЕРПРЕТАЦІЄЮ ПРИЙНЯТИХ РІШЕНЬ

Етап виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень дозволяє перетворювати вхідні дані у вигляді тексту для аналізу та навчених окремим технікам моделей машинного навчання у вихідні дані, які містять числові оцінки наявності кожної з технік пропаганди та розмічений текст із візуальною аналітикою присутності детектованих маркерів пропаганди. Схему етапу наведено на рисунку 3.

Етап здійснює перетворення вхідних даних у вигляді множини навчених неймереж до оцінки сили прояву кожного маркера з доповненою множини маркерів, множини навчених неймереж для оцінки сили прояву кожного прийому пропаганди та тексту для аналізу у вихідні дані у вигляді оцінки наявності кожного прийому пропаганди за текстом та доповненою розміткою з множини маркерів та візуального подання прийнятих неймережевими моделями рішень щодо наявності прийомів пропаганди.

В рамках запропонованого підходу буде виконано виявлення 17 основних прийомів пропаганди: «Appeal to fear-prejudice», «Name Calling», «Causal Oversimplification», «Doubt», «Exaggeration», «Flag-Waving», «Labeling», «Loaded Language», «Minimisation», «Repetition», «Thought terminating Cliches», «Appeal to Authority», «Black and White Fallacy», «Slogans», «Reductio ad hitlerum», «Red Herring», «Whataboutism» [14] та 4 види маркерів: «емоційність тексту», «булінг», «страх» та «мова ворожнечі».

Маркери показують, як різні пропагандистські прийоми пов'язані з емоційністю тексту, булінгом, страхом і мовою ворожнечі. Наприклад, прийом «Appeal to Fear-Prejudice» характеризується високим рівнем маркеру «емоційність тексту», «булінгу» і «страху», але нейтральний щодо маркера «мова ворожнечі». Інші прийоми, такі як «Causal Oversimplification» і «Minimization» зазвичай мають нейтральний вплив на всі чотири маркера.

Таким чином, етап виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень для тестового тексту дозволяє одержувати числові оцінки наявності кожної з технік пропаганди та розмічений текст із візуальною аналітикою присутності детектованих маркерів пропаганди.

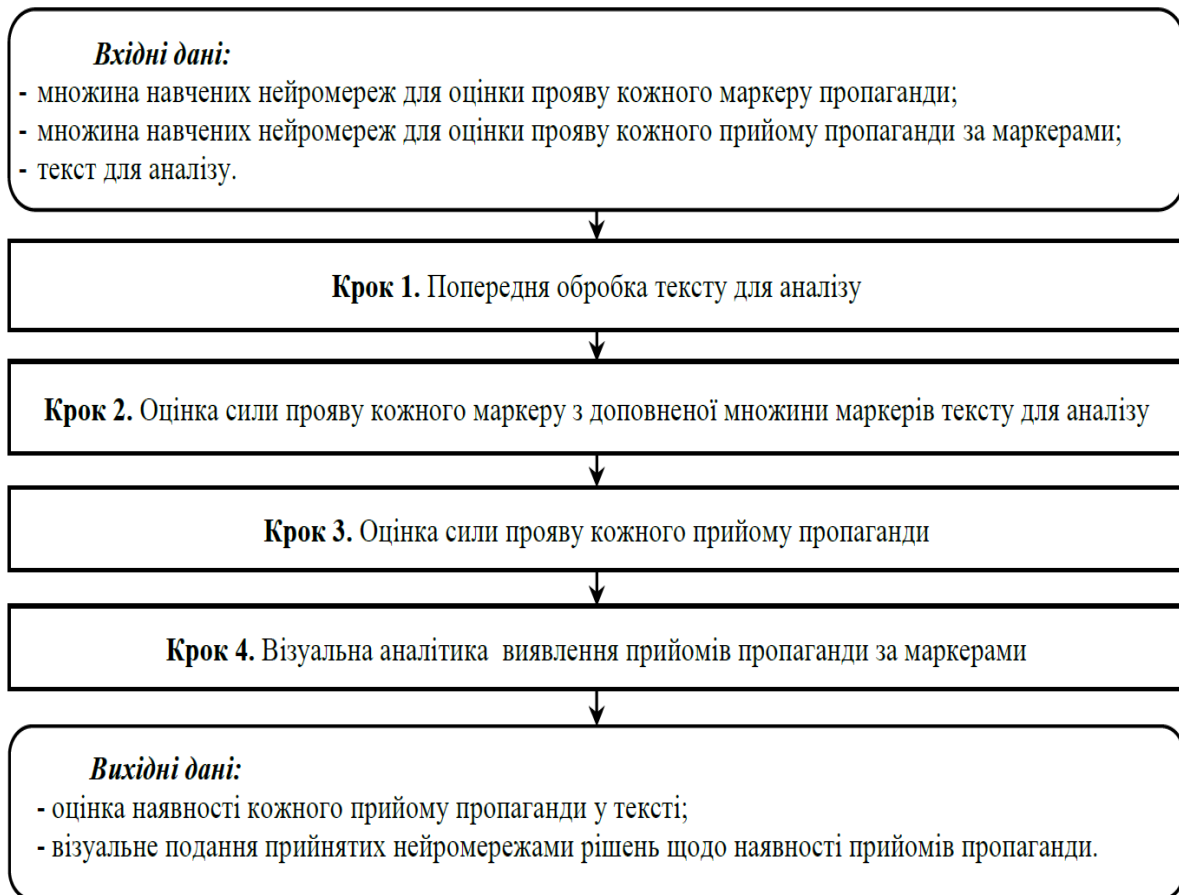


Рис. 3. Кроки для виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень

ВИЯВЛЕННЯ ОБ'ЄКТІВ ПРОПАГАНДИ З ВІЗУАЛЬНОЮ ІНТЕРПРЕТАЦІЄЮ ПРИЙНЯТИХ РІШЕНЬ

Етап виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень характеризується розширенням множини об'єктів пропаганди за рахунок додавання варіантів їх словесних подань і використанням контекстних вікон для виявлення взаємозв'язків між використаними прийомами та об'єктами пропаганди, що дало змогу покращити результати виявлення об'єктів пропаганди та візуально їх інтерпретувати. Схему етапу наведено на рисунку 4.

Етап виявлення об'єктів пропаганди з візуальною інтерпретацією прийнятих рішень виконує перетворення вхідних даних у вигляді тестового тексту для виявлення об'єктів пропаганди та множини використаних технік у тестовому тексті і множини навчених нейромережеских моделей для кожного прийому пропаганди у вихідні дані у вигляді візуального представлення семантичної моделі пропаганди для тексту, яке включає: множини об'єктів пропаганди; множини слів що представляють об'єкти з їх семантичною оцінкою; множини важливості зв'язків між техніками та об'єктами пропаганди у тексті з оцінками семантичної важливості.

Таким чином, основними кроками етапу є формування множини об'єктів пропаганди шляхом розпізнавання іменованих сутностей, попередня обробка тексту й розширення множини об'єктів пропаганди за рахунок додавання варіантів їх словесних подань, побудова контекстних вікон для словесних подань кожного об'єкту пропаганди з урахуванням порогового значення мінімального розміру вікна, оцінювання рівня використаних прийомів пропаганди в межах контекстних вікон за нейромережевими моделями, й побудова множини важливих зв'язків між техніками та об'єктами пропаганди з урахуванням порогового значення мінімального рівня прояву пропаганди.

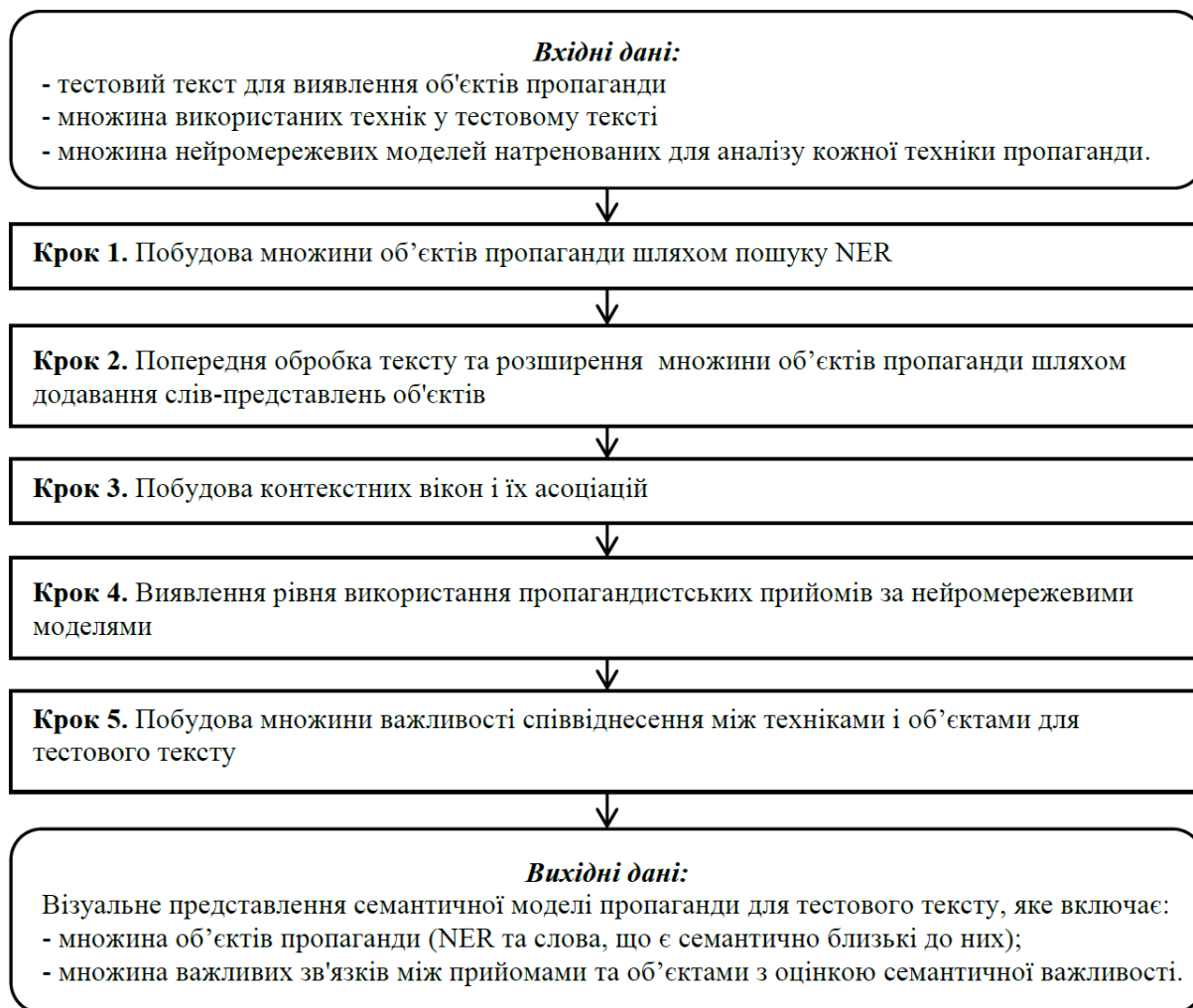


Рис. 4. Кроки для виявлення об'єктів пропаганди з візуальною інтерпретацією прийнятих рішень

ЕКСПЕРИМЕНТ, РЕЗУЛЬТАТИ ТА ДИСКУСІЯ

У якості даних дослідження для навчання і оцінки продуктивності нейромереж для ідентифікації пропагандистського вмісту та виявлення прийомів було використано набори даних від команди «Analysis Project» [15] та [16]. Дані представляють корпус новинних статей, який було анотовано вручну на рівні фрагментів із застосуванням вісімнадцяти пропагандистських прийомів. Загальний обсяг набору даних складає 788 статей.

Для визначення наявності пропаганди, використаних прийомів та об'єктів пропаганди було використано спеціалізоване створене програмне забезпечення. Отримані результати у вигляді візуальної аналітики зіставлялися з висновками авторитетних джерел і експертів, що займаються виявленням та протидією пропаганді. Для аналізу обрано дописи з соціальних мереж, які були оброблені «Центром стратегічних комунікацій» [17]. Ці дописи вже мають розмітку та експертні висновки, що дозволяє порівнювати їх із результатами роботи запропонованого методу. Приклад роботи створеного програмного забезпечення наведено на рис.5.

Підхід до нейромережового виявлення і класифікації прийомів та об'єктів пропаганди дозволяє шляхом об'єднання традиційних рекурентних нейромереж з довгостроковою пам'яттю із трансформерами забезпечити більш глибоке розуміння послідовності та контексту в текстовому контенті, що дає ефект у вигляді отриманого показника Ассигасу (точності) 0.98 для ідентифікації пропаганди, що переважає відомі аналоги.

Множина іменованих сутностей з семантично-близькими об'єктами за аналізом контекстних залежностей:

україна, LOC, країна (0.66), практика (0.25), політика (0.22), техніка (0.20)
 сполучений штатів, LOC, списаний (0.29), західний (0.24)
 європейський союз, LOC, фронт (0.23), говорити (0.18)
 захід, LOC, західний (0.50), заява (0.28), зникати (0.23), захищати (0.20)

Множина використаних об'єктів пропаганди у тексті:

Західна допомога **Україні** – це лише фарс, створений для того, щоб виглядати, ніби вони підтримують нашу **країну**. Насправді, всі ці обіцянки та кредити не приносять реальної користі. Вони і створюють ілюзію підтримки, тоді як наші люди продовжують страждати. Військова **техніка**, яку нам постачають, – це застарілі моделі, що не здатні протистояти сучасним викликам. Західні **політики** роблять вигляд, що допомагають, але їхня допомога – це пустий звук. Подивімося на обіцянки **Сполучених Штатів**. Вони заявляють про мільярди доларів допомоги, але насправді ці кошти майже не доходять до **фронту**. Куди **зникають** ці гроші? Невідомо. І хоч бачимо гучні **заяви** про поставки сучасної зброї, в реальності ми отримуємо лише старі моделі, які давно **списані** на заході. Те саме стосується і допомоги від **Європейського Союзу**. Їхні обіцянки підтримки звучать голосно, але на **практиці** вони обмежуються символічними жестами. Західні **країни** постійно **говорять** про важливість підтримки **України**, але де реальні дії? Де сучасна **техніка**, яка допомогла б нам **захищати** наші землі? Натомість ми отримуємо старе озброєння, яке не відповідає сучасним стандартам війни. Це показує справжнє ставлення **Заходу** до нашої боротьби. Вони просто намагаються заспокоїти власну совість, не вкладаючи реальних зусиль у нашу перемогу.

Сили прояву використаних технік та приналежність їм тематичних об'єктів:

Використані техніки:

1. Doubt (Сумнів). Виражена на 0.689
2. Loaded Language (Заряджена мова). Виражена на 0.387
3. Minimisation (Мінімізація). Виражена на 0.511

Оцінка приналежності об'єктів пропаганди технікам:

{**україна** (LOC) Доповнена тематична множина: [країна, практика, політика, техніка]} **Оцінки приналежності:** [Doubt 0.78; Loaded Language 0.24; Minimisation 0.67]
 {**сполучений штатів** (LOC) Доповнена тематична множина: [списаний, західний]} **Оцінки приналежності:** [Doubt 0.64; Loaded Language 0.31; Minimisation 0.7]
 {**європейський союз** (LOC) Доповнена тематична множина: [фронт, говорити]} **Оцінки приналежності:** [Doubt 0.17; Loaded Language 0.12; Minimisation 0.39]
 {**захід** (LOC) Доповнена тематична множина: [західний, заява, зникати, захищати]} **Оцінки приналежності:** [Doubt 0.53; Loaded Language 0.4; Minimisation 0.09]

Рис. 5. Приклад виявлення прийомів та об'єктів у пропагандистському контенті

За проведеним експериментом по навченим неймережевим моделям глибокого навчання для виокремлення використаних прийомів пропаганди, вдалось досягнути показників від 0.82 до 0.97 за метрикою Accuarcy. Отримані значення точності для прийомів пропаганди наведені на рис.6.

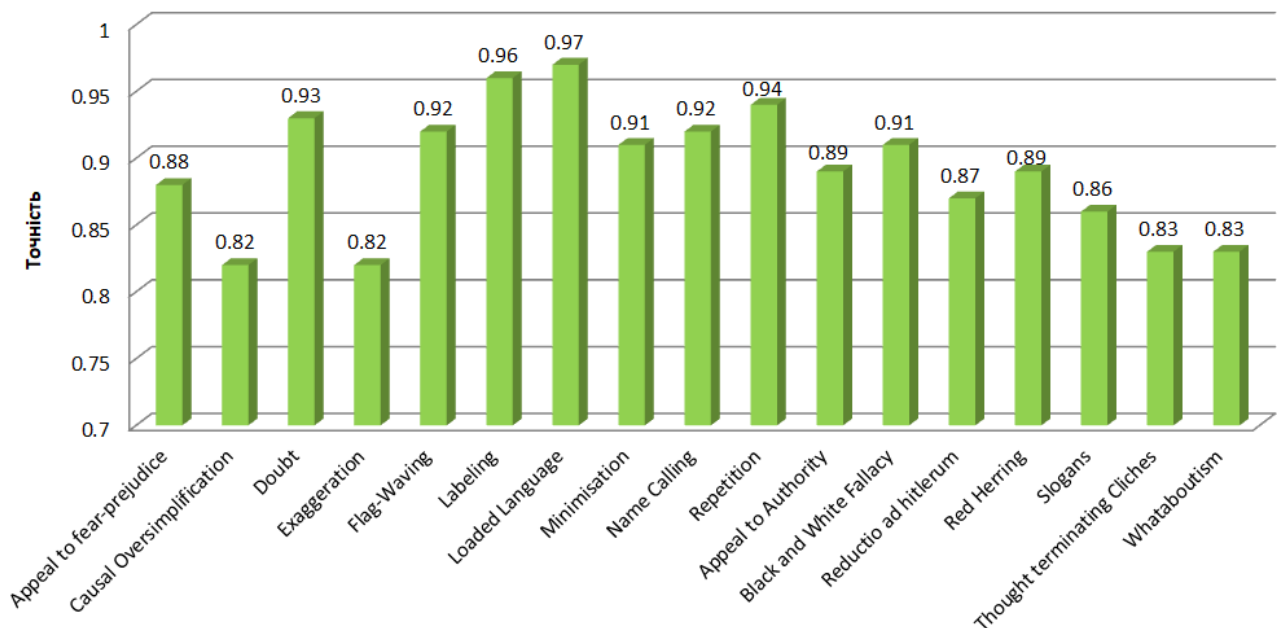


Рис. 6. Оцінка виявлення прийомів пропаганди за метрикою Accuarcy

Результати показують, що запропонований метод здатний ефективно виявляти техніки пропаганди та надає можливість візуально визначити, які саме дані вплинули на рішення моделей щодо виявлених прийомів пропаганди в текстах користувачів.

Таким чином, в результаті дослідження ефективності розробленого підходу до виявлення і класифікації прийомів та об'єктів пропаганди у текстовому контенті з використанням неймережевих засобів було виявлено, що він дозволяє одержувати результати, які цілком корелюють із результатами, одержаними експертами. При цьому внаслідок розгляду пропаганди як цілісної моделі та використання візуальної аналітики одержаних результатів, вдалося досягти комплексного аналізу взаємозв'язків прийомів та об'єктів пропаганди, а також забезпечити узагальнення для об'єктів пропаганди та їх альтернативних згадувань у текстах.

ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Розроблено підхід до нейромережевого виявлення і класифікації прийомів та об'єктів пропаганди у текстовому контенті, що складається з трьох послідовних етапів та забезпечує ефективне виявлення наявності пропаганди, класифікацію використаних технік пропаганди, та встановлення об'єктів виявленого пропагандистського впливу.

Етап класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання призначений для автоматизованої ідентифікації текстів, які містять пропагандистські елементи. Особливістю етапу є те, що він дозволяє виявляти як явні, так і приховані пропагандистські меседжі, ґрунтуючись на об'єднанні традиційних рекурентних нейромереж з довгостроковою пам'яттю із трансформерами, а також на використанні механізму аугментації навчальних текстових даних, що дозволяє розширити кількість навчальних зразків. Рішення ґрунтується на об'єднанні традиційних рекурентних нейромереж з довгостроковою пам'яттю із нейромережами трансформерами, що забезпечує більш глибоке розуміння послідовності та контексту в текстовому контенті та дозволяє досягнути точності 97.83 %.

Етап виявлення прийомів пропаганди за маркерами із візуальною інтерпретацією прийнятих рішень дозволяє перетворювати вхідні дані у вигляді тексту для аналізу та навчених окремим технікам 17 моделей машинного навчання у вихідні дані, які містять числові оцінки наявності кожної з технік пропаганди та розмічений текст із візуальною аналітикою присутності детектованих маркерів пропаганди. Отримані результати забезпечили виявлення різних пропагандистських технік з мінімальною точністю 82,03%, що краще за відомі аналоги щодо виявлення пропаганди незалежно від використовуваних методик.

Етап виявлення об'єктів пропаганди нейромережевими моделями глибокого навчання з візуальною інтерпретацією прийнятих рішень характеризується розширенням множини об'єктів пропаганди за рахунок додавання варіантів їх словесних подань і використанням контекстних вікон для виявлення взаємозв'язків між використаними прийомами та об'єктами пропаганди, що дало змогу покращити результати виявлення об'єктів пропаганди та візуально їх інтерпретувати. В результаті дослідження ефективності підходу було виявлено, що він дозволяє одержувати результати, які цілком корелюють із результатами, одержаними експертами.

Таким чином, в рамках основної мети підвищення точності та якості виявлення прийомів та об'єктів пропаганди за семантичними маркерами у текстовому контенті засобами штучного інтелекту з подальшим поясненням прийнятих рішень, розроблено підхід до вирішення проблеми ідентифікації об'єктів пропаганди, який дозволяє знайти в пропагандистських текстах, на кого і на що спрямовані пропагандистські прийоми. Вирішено вагомі проблеми в напрямку виявлення пропаганди, а саме: відсутність комплексного аналізу взаємозв'язків прийомів та об'єктів пропаганди в текстах; відсутність узагальнень для об'єктів пропаганди та їх альтернативних згадок у текстах. Для візуальної інтерпретації одержаних результатів, формується візуальна аналітика щодо знайдених технік та об'єктів пропаганди, що дозволяє візуально спостерігати об'єкти впливу в рамках використовуваних технік пропаганди.

Література

1. Молчанова М.О. Дослідження ефективності методу класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання. Інформаційні технології і автоматизація. Матеріали XVII міжнародної науково-практичної конференції. 31 жовтня – 1 листопада 2024 р. Одеса, ОНТУ. 2024. С.665-668.
2. Martino G.D.S. SemEval-2020 Task 11: Detection of Propaganda Techniques in News Articles / G.D.S. Martino, A. Barron-Cedeno, H. Wachsmuth, R. Petrov, P. Nakov // Proceedings of the Fourteenth Workshop on Semantic Evaluation. – 2020. – С. 1377–1414.
3. Zalutska O., Molchanova M., Sobko O., Mazurets O., Pasichnyk O., Barmak O., Krak I. Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network. CEUR Workshop Proceedings, 2023, vol. 3387, pp. 344-356
4. Молчанова М.О. Алгоритм виявлення аб'юзивного вмісту в україномовному аудіоконтенті для імплементації в об'єктно-орієнтовану інформаційну систему / М.О. Молчанова, О.В. Мазурець, О.В. Собко та ін. // Вісник Хмельницького національного університету. Технічні науки. – 2024. – № 1 (331). – С. 101–106.
5. Krak I. Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network / I. Krak, O. Zalutska, M. Molchanova, O. Mazurets, R. Bahrii, O. Sobko, O. Barmak // CEUR Workshop Proceedings. – 2024. – Vol. 3688. – С. 16–28.
6. Молчанова М.О. Метод інтелектуального аналізу тональності текстів / М.О. Молчанова, О.О. Залуцька, О.В. Бармак // Матеріали XII Всеукраїнської науково-практичної конференції «Глушковські читання». – Київ, 2023. – С. 113–116.
7. Nazarov V. Information System for Detecting Abusive Speech in Audio Content by Means of Natural Language / V. Nazarov, M. Molchanova // Proceedings of V International Scientific and Practical Conference «Modern strategies of global scientific solutions». – Stockholm, Sweden : International Scientific Unity, 2023. – С. 132–135.

8. Ahmad P.N. Robust Benchmark for Propagandist Text Detection and Mining High-Quality Data / P.N. Ahmad, Y. Liu, G. Ali, M.A. Wani, M. ElAffendi // *Mathematics*. – 2023. – Vol. 11. – С. 2668.
9. Молчанова М.О. Застосування аугментації даних для підвищення точності виявлення пропаганди в інтернет-джерелах нейромережевими моделями глибокого навчання / М.О. Молчанова // *Матеріали VIII Міжнародної науково-практичної конференції «Перспективи сучасної науки: теорія і практика»*. – Львів, 2024. – С. 199–205.
10. Przybyla P. Long Named Entity Recognition for Propaganda Detection and Beyond / P. Przybyla, K. Kaczynski // *Proceedings of the International Conference of the Spanish Society for Natural Language Processing*. – 2023.
11. Krak I. Method for Political Propaganda Detection in Internet Content Using Recurrent Neural Network Models Ensemble / I. Krak, V. Didur, M. Molchanova, O. Mazurets, O. Zalutska, E. Manziuk, O. Barmak // *CEUR Workshop Proceedings*. – 2024. – Vol. 3806. – С. 312–324.
12. Молчанова М.О. Метод виявлення та класифікації прийомів пропаганди у текстовому контенті засобами штучного інтелекту / М.О. Молчанова // *Матеріали XII Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології IUST-ODESSA-2024»*. – Одеса, 2024. – С. 251–254.
13. Krak I. Method for Neural Network Detecting Propaganda Techniques by Markers With Visual Analytic / I. Krak, M. Molchanova, O. Mazurets, O. Sobko, O. Zalutska, O. Barmak // *CEUR Workshop Proceedings*. – 2024. – Vol. 3790. – С. 158–170.
14. Yoosuf S. Fine-Grained Propaganda Detection with Fine-Tuned BERT / S. Yoosuf, Y. Yang // *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*. – Hong Kong, China, 2019. – С. 87–91.
15. Propaganda Analysis Project [Електронний ресурс]. – 2023. – Режим доступу: <https://propaganda.math.unipd.it/index.html>.
16. Zenodo. Propaganda. Propy Corpus 1.0 [Електронний ресурс]. – 2019. – Режим доступу: <https://zenodo.org/records/3271522#.XS6qRUUzau4>.
17. Центр стратегічних комунікацій [Електронний ресурс]. – Режим доступу : <https://spravdi.gov.ua/>.

References

1. Molchanova, M.O. Research on the Effectiveness of Text Classification Methods for Propaganda Content Using Deep Learning Neural Network Models. *Information Technologies and Automation. Proceedings of the XVII International Scientific and Practical Conference, October 31 – November 1, 2024, Odessa, ONTU. 2024*. pp. 665-668.
2. Martino G.D.S. SemEval-2020 Task 11: Detection of Propaganda Techniques in News Articles / G.D.S. Martino, A. Barron-Cedeno, H. Wachsmuth, R. Petrov, P. Nakov // *Proceedings of the Fourteenth Workshop on Semantic Evaluation*. – 2020. – S. 1377–1414.
3. Zalutska O., Molchanova M., Sobko O., Mazurets O., Pasichnyk O., Barmak O., Krak I. Method for Sentiment Analysis of Ukrainian-Language Reviews in E-Commerce Using RoBERTa Neural Network. *CEUR Workshop Proceedings, 2023, vol. 3387*, pp. 344-356.
4. Molchanova M.O. Alhorytm vyivlennia abiuzyvnoho vmistu v ukrainomovnomu audiokontenti dlia implementatsii v obiektno-orientovanu informatsiinu systemu / M.O. Molchanova, O.V. Mazurets, O.V. Sobko ta in. // *Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky*. – 2024. – № 1 (331). – S. 101–106.
5. Krak I. Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network / I. Krak, O. Zalutska, M. Molchanova, O. Mazurets, R. Bahrii, O. Sobko, O. Barmak // *CEUR Workshop Proceedings*. – 2024. – Vol. 3688. – S. 16–28.
6. Molchanova M.O. Metod intelektualnoho analizu tonalnosti tekstiv / M.O. Molchanova, O.O. Zalutska, O.V. Barmak // *Materialy XII vseukrainskoi naukovo-praktychnoi konferentsii «Hlushkovski chytannia»*. – Kyiv, 2023. – S. 113–116.
7. Nazarov V. Information System for Detecting Abusive Speech in Audio Content by Means of Natural Language / V. Nazarov, M. Molchanova // *Proceedings of V International Scientific and Practical Conference «Modern strategies of global scientific solutions»*. – Stockholm, Sweden : International Scientific Unity, 2023. – S. 132–135.
8. Ahmad P.N. Robust Benchmark for Propagandist Text Detection and Mining High-Quality Data / P.N. Ahmad, Y. Liu, G. Ali, M.A. Wani, M. ElAffendi // *Mathematics*. – 2023. – Vol. 11. – S. 2668.
9. Molchanova M.O. Zastosuvannia auhmentatsii danykh dlia pidvyshchennia tochnosti vyivlennia propahandy v internet-dzherelakh neiomerezhovyimi modeliamy hlybokoho navchannia / M.O. Molchanova // *Materialy VIII Mizhnarodnoi naukovo-praktychnoi konferentsii «Perspektyvy suchasnoi nauky: teoriia i praktyka»*. – Lviv, 2024. – S. 199–205.
10. Przybyla P. Long Named Entity Recognition for Propaganda Detection and Beyond / P. Przybyla, K. Kaczynski // *Proceedings of the International Conference of the Spanish Society for Natural Language Processing*. – 2023.
11. Krak I. Method for Political Propaganda Detection in Internet Content Using Recurrent Neural Network Models Ensemble / I. Krak, V. Didur, M. Molchanova, O. Mazurets, O. Zalutska, E. Manziuk, O. Barmak // *CEUR Workshop Proceedings*. – 2024. – Vol. 3806. – S. 312–324.
12. Molchanova M.O. Metod vyivlennia ta klasyfikatsii pryiomiv propahandy u tekstovomu kontenti zasobamy shuchnoho intelektu / M.O. Molchanova // *Materialy XII Mizhnarodnoi naukovo-praktychnoi konferentsii «Informatsiini upravliaiuchi systemy ta tekhnologii IUST-ODESSA-2024»*. – Odesa, 2024. – S. 251–254.
13. Krak I. Method for Neural Network Detecting Propaganda Techniques by Markers With Visual Analytic / I. Krak, M. Molchanova, O. Mazurets, O. Sobko, O. Zalutska, O. Barmak // *CEUR Workshop Proceedings*. – 2024. – Vol. 3790. – S. 158–170.
14. Yoosuf S. Fine-Grained Propaganda Detection with Fine-Tuned BERT / S. Yoosuf, Y. Yang // *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*. – Hong Kong, China, 2019. – S. 87–91
15. Propaganda Analysis Project [Elektronnyi resurs]. – 2023. – Rezhym dostupu: <https://propaganda.math.unipd.it/index.html>.
16. Zenodo. Propaganda. Propy Corpus 1.0 [Elektronnyi resurs]. – 2019. – Rezhym dostupu: <https://zenodo.org/records/3271522#.XS6qRUUzau4>.
17. Tsentr stratehichnykh komunikatsii [Elektronnyi resurs]. – Rezhym dostupu: <https://spravdi.gov.ua/>.