

# ЗАСТОСУВАННЯ АУГМЕНТАЦІЇ ДАНИХ ДЛЯ ПІДВИЩЕННЯ ТОЧНОСТІ ВИЯВЛЕННЯ ПРОПАГАНДИ В ІНТЕРНЕТ-ДЖЕРЕЛАХ НЕЙРОМЕРЕЖЕВИМИ МОДЕЛЯМИ ГЛИБОКОГО НАВЧАННЯ

**Молчанова Марина Олексіївна**

аспірант

Хмельницький національний університет, Україна

**Вступ.** Загроза маніпуляції ЗМІ на громадську думку спонукає до наукових досліджень пропаганди та мовних впливів, а також до аналізу комунікаційних факторів в контексті інформаційної безпеки [1]. Зростання споживання онлайн-контенту посилює ризики пропаганди, що загрожує національній безпеці несвоєчасне вирішення якої може призвести до руйнівних наслідків. У соціологічному енциклопедичному словнику термін «пропагандистський допис» розглядається у декількох варіантах: поширення в масах ідеології та політики певних класів, партій, держав; засіб маніпуляції масовою свідомістю [2].

Метод класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання призначений для автоматизованої ідентифікації текстів, які містять пропагандистські елементи. Метод працює шляхом 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».

Першим кроком є попередня обробка тексту для класифікації, що включає в себе ряд кроків препроцесінгу [3], таких як перетворення тексту у нижній регістр, видалення стоп-слів та елементів пунктуації тощо. Попередньо оброблений текст перетворюється у числові послідовності, які будуть подані навченій нейромережевій моделі глибокого навчання з гібридною архітектурою для подальшого одержання відсоткових показників наявності пропаганди.

Другим кроком є аналіз тексту навченою нейромережевою моделлю, який

виконає числову оцінку рівня наявності пропаганди у тексті. Для отримання даної числової оцінки використовується нейронна мережа глибокого навчання, що має гібридну архітектуру, яка поєднує в собі традиційну рекурентну нейромережеву модель [4] з довгостроковою пам'яттю із трансформерами, що забезпечують більш глибоке розуміння послідовності та контексту в текстовому контенті [5].

Кроком 3 є віднесення проаналізованого тексту до 1-го з 3-х класів: «текст без пропаганди», «пропагандистський текст» та «підозрілий текст». Для цього емпіричним шляхом були встановлені межі для кожного з класів. Клас «текст без пропаганди» має межі від 0 до 0.45, «підозрілий текст» має межі від 0.45 до 0.55, та «пропагандистський текст» від 0.55 до 1. Межі можуть змінюватись та налаштовуватись в залежності від видів пропаганди та специфіки користувацьких даних.

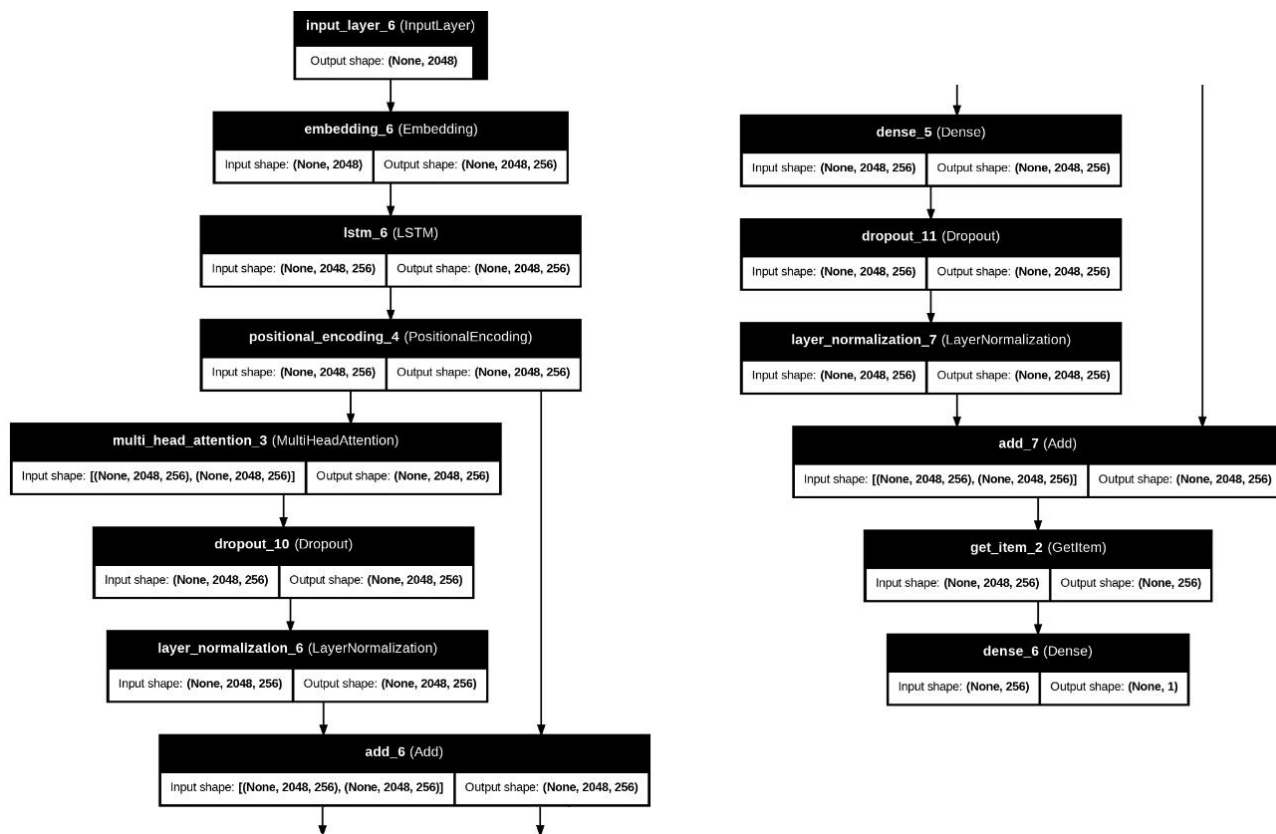
Відповідно, вихідними даними є відсоткова оцінка наявності пропаганди у тексті та присвоєння 1 з 3-х класів: «текст без пропаганди», «пропагандистський текст» та «підозрілий текст».

**Метою роботи** є дослідження ефективності застосування аугментації даних для підвищення точності виявлення пропаганди в інтернет-джерелах нейромережевими моделями глибокого навчання.

**Матеріали та методи.** Для отримання числової оцінки рівня наявності пропаганди у тексті використовується нейронна мережа глибокого навчання, що має гібридну архітектуру, яка поєднує в собі традиційну рекурентну нейромережеву модель з довгостроковою пам'яттю із трансформерами, що забезпечують більш глибоке розуміння послідовності та контексту в текстовому контенті. Використовувана для аналізу тексту та визначення відсоткової оцінки наявності пропаганди у тексті нейромережева архітектура схематично наведена на рис. 1.

Така архітектура нейромережі обумовлена тим, що LSTM добре підходить для обробки послідовних даних, але має обмеження щодо утримання довготривалих залежностей. Для текстів великої довжини, модель може мати

труднощі з утриманням контексту на протязі всього тексту. У той же час Multi-Head Attention шар є більш ефективним у збереженні довготривалих залежностей, оскільки враховують весь текст одразу. Проте, ця архітектура може вимагати значних обчислювальних ресурсів при роботі з довгими послідовностями.

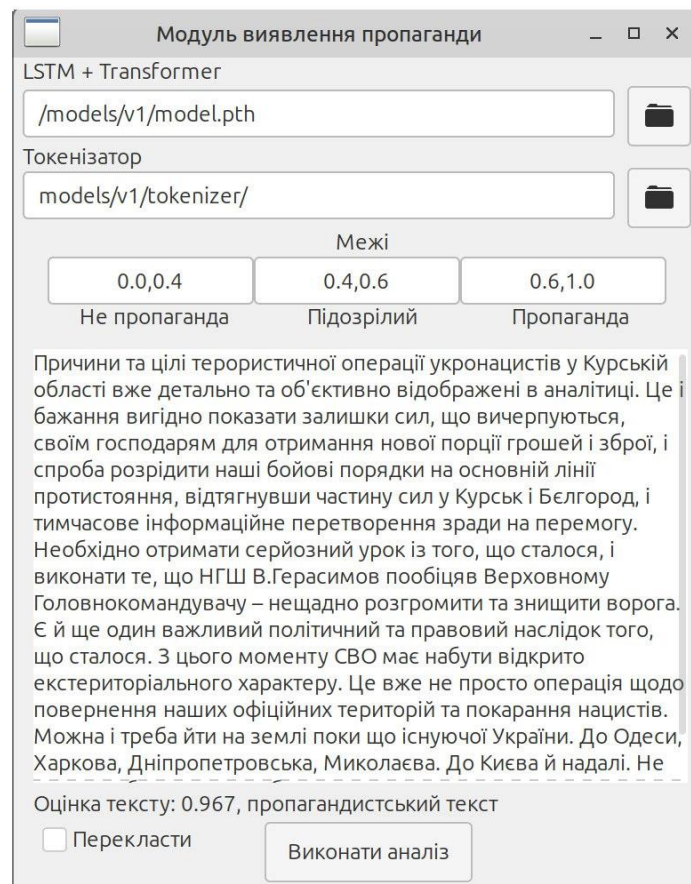


**Рис. 1. Гібридна архітектура нейронної мережі глибокого навчання для нейромережевого виявлення пропаганди в інтернет-джерелах**

Оскільки є проблема з недостатністю даних для навчання нейромережевих моделей, у рамках методу буде використано метод аугментації тексту для збільшення різноманітності навчальних даних. Для чого буде використано модель T5 (Text-to-Text Transfer Transformer), яка є трансформерною нейромережевою моделлю, що обробляє всі завдання як завдання перетворення тексту в текст. Це включає в себе переклад, узагальнення, питання-відповідь, перефразування.

Для застосування аугментації даних для підвищення точності виявлення пропаганди в інтернет-джерелах нейромережевими моделями глибокого

навчання було створено програмну реалізацію, яка складається із ноутбуків реалізованих у хмарному сервісі «Google Colab» (для навчання нейромережевої моделі гібридної архітектури та для розширення отриманого набору даних методом аугментації тексту), а також застосунок з графічним інтерфейсом користувача на мові Python використовуючи середовище розробки PyCharm. Приклад роботи розробленого застосунку наведено на рис. 2.



**Рис. 2. Програмна реалізація методу класифікації текстів за вмістом пропаганди нейромережевими моделями глибокого навчання**

**Результати та обговорення** Експеримент було проведено з застосуванням розширенням отриманого набору даних методом аугментації, та без розширення набору даних. У ході дослідження вдалось досягнути точності у 97.8%. Результати експерименту щодо навчання нейромережі гібридної архітектури з використанням аугментації навчальної вибірки та без аугментації наведені в таблиці 1 (без використання аугментації) та в таблиці 2 (з використанням аугментації).

Таблиця 1

Результати навчання нейромережі без використання аугментації  
навчальної вибірки

Кількість епох	Accuracy, %	Precision, %	Recall, %
5	92.31	93.5	94.8
6	93.42	94.2	95.3
7	93.14	94	95.2
8	95.15	95.5	96.5
9	96.18	96.0	97
10	96.19	96.1	97.1
<b>15</b>	<b>96.94</b>	<b>96.5</b>	<b>97.8</b>
20	96.46	96.3	97.6

Таблиця 2

Результати навчання нейромережі з використанням аугментації  
навчальної вибірки

Кількість епох	Accuracy, %	Precision, %	Recall, %
5	85.49	86.51	87.12
6	86.3	87.2	87.8
7	92.89	93.06	93.53
8	93.15	93.47	93.71
9	95.26	95.52	96.0
10	95.12	95.41	95.87
15	97.11	97.23	97.58
<b>20</b>	<b>97.83</b>	<b>98.17</b>	<b>98.16</b>

При використанні лише 2-х категорій – «пропаганда» та «не пропаганда», найбільша кількість помилок зосереджена в проміжку 0.4 – 0.6 (рис. 3). Тому уведення категорії «підозрілий текст» із можливістю плаваючих меж адаптивним предметній області є доцільним.

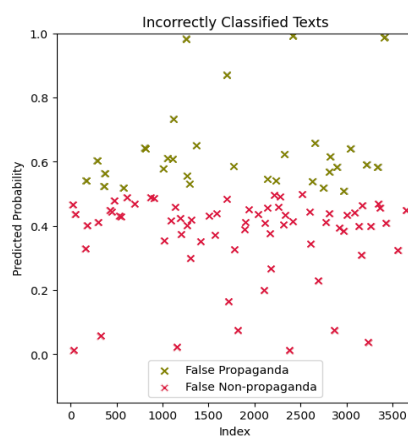


Рис. 3. Розподіл некоректно ідентифікованих до категорій «пропаганда» та «не пропаганда» зразків текстів

Як видно з таблиць 1 та 2, вдалося досягти Accuracy 96,94% при гібридній архітектурі без використання аугментації навчальної вибірки та 97.83% з використанням аугментації навчальної вибірки.

**Висновки.** Таким чином, із застосуванням аугментації кращі показники досягаються при більшій кількості епох. Це пояснюється розширенням навчальної вибірки, що призводить до потреби більшої кількості епох. В той же час, при застосуванні аугментації вдалося досягнути точності 97.83 %, при тому що без аугментації цей показник максимально досяг рівня 96.94%.

Отримані результати говорять про спроможність ефективно класифікувати тексти за вмістом пропаганди нейромережевими моделями глибокого навчання. Застосування додаткової категорії «підозрілий текст» дозволило підняти показники Precision та Recall, що у свою чергу дає можливість автоматизованої модерації текстів на предмет пропаганди з помилками не більше 1.83% для хибного виявлення пропаганди.

## ПОСИЛАННЯ

1. Молчанова М. О., Мазурець О. В., Собко О. В., Віт Р. В., Назаров В. В. Алгоритм виявлення аб'юзивного вмісту в україномовному аудіоконтенті для імплементації в об'єктно-орієтовану інформаційну систему. Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2024. №1 (331). С. 101-106.

2. Slobodzian V., Kovalchuk O., Molchanova M., Sobko O., Mazurets O., Barmak O., Krak I. Text Data Vectorization Model of Ukrainian-Language Internet Communication Content. CEUR Workshop Proceedings, 2022, vol. 3171, pp. 561-571.

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., Мазурець О. В., Собко О. В., Кліменко В. І., Андрощук В. І. Метод нейромережевого виявлення кібербулінгу з використанням хмарних сервісів та об'єктно-орієнтованої моделі. Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2024. №2 (333). С. 200-206.

5. Молчанова М. О., Залуцька О. О., Бармак О. В. Метод інтелектуального аналізу тональності текстів. Матеріали XII Всеукраїнської науково-практичної конференції «Глушковські читання». Київ, 2023. с. 113-116.