

УДК 004.9:070

## **ВИЯВЛЕННЯ ПРИХОВАНОЇ ПРОПАГАНДИ В МЕДІАТЕКСТАХ**

Самелюк А. А.

Науковий керівник – д.т.н., проф. Терзіян В. Я.

Харківський національний університет радіоелектроніки, каф. ШІ

м. Харків, Україна

тел.: +38(050) 246-12-64, e-mail: andrii.sameliuk@nure.ua

The rapid development of technology and the ubiquity of the media have led to the spread of media texts containing covert propaganda. Such propaganda aims to influence public opinion, often implicitly, and has serious implications for democratic societies. Methods for detecting and analyzing hidden propaganda in media texts using natural language processing (NLP) methods are researched. Various unsupervised and supervised learning strategies are emphasized, such as topic modeling, clustering, and deep learning methodologies. The significance of devising methods capable of identifying propaganda across diverse languages and cultural settings is addressed. Also, the limitations of these methods are considered.

Виявлення прихованої пропаганди в медіатекстах представляє значні труднощі через її заплутані характеристики. Методи обробки природної мови (NLP) пропонують багатообіцяючий підхід до ідентифікації та аналізу такої пропаганди. Алгоритми машинного навчання, зокрема навчання з учителем та без, можуть бути застосовані для завдання виявлення пропаганди.

Одним із популярних методів навчання без учителя для виявлення пропаганди є тематичне моделювання, яке може ідентифікувати приховані тематичні структури в медіатекстах. Прихований розподіл Діріхле (LDA) — це широко використовувана техніка моделювання тем, яка може виявити приховану пропаганду шляхом виявлення тем, які постійно пов'язані з певними медіа-джерелами чи політичними ідеологіями. Аналізуючи розподіл тем у різних ЗМІ, можливо виявити потенційні упередження та приховані наміри [1].

Іншим підходом до навчання без учителя є кластеризація, яка об'єднує подібні медіатексти на основі їх змісту та мовних особливостей. Алгоритми кластеризації, такі як k-середніх та ієрархічна кластеризація, можуть виявити закономірності поширення пропаганди, досліджуючи розподіл контенту між різними джерелами інформації, виявляючи кластери схожих медіатекстів.

Методи глибокого навчання, такі як CNN, RNN і Transformer, також можна використовувати для виявлення прихованої пропаганди. Ці моделі здатні виявляти складні мовні патерни та семантичні зв'язки в медіатекстах. Попередньо треновані мовні моделі, такі як BERT і GPT, можна налаштувати для розпізнавання конкретних пропагандистських

прийомів, таких як завантаження логічними помилками або емоційні заклики [2].

На додаток до цих підходів можливо використовувати аналіз настроїв, щоб виявити пропаганду, вивчаючи емоційний тон медіатекстів. Пропагандисти часто використовують емоційно насичену мову, щоб викликати сильну реакцію та вплинути на громадську думку. Аналізуючи настрої медіаконтенту, можливо виявити випадки маніпуляції та потенційну тактику пропаганди.

Виявлення міжмовної та міжкультурної пропаганди є ще одним важливим напрямом для подальших досліджень. Оскільки пропаганда не обмежується однією мовою чи культурою, важливо розробити методи, здатні виявляти пропаганду в різних мовах і культурних контекстах. Трансферне навчання та інші міжмовні методи можуть бути використані для адаптації існуючих моделей до нових мов і доменів, підвищуючи ефективність виявлення пропаганди в глобальному масштабі [3].

Незважаючи на прогрес у методах NLP для виявлення прихованої пропаганди, залишається кілька проблем і обмежень. До них належать потреба у великих високоякісних розмічених наборах даних, складність розрізнення справжньої думки від навмисної маніпуляції, а також ризик хибних позитивних чи хибно негативних результатів. Крім того, оскільки пропагандисти адаптують свою тактику та використовують більш складні методи, потрібно постійно вдосконалювати методи та розробляти нові підходи.

Виявлення прихованої пропаганди в медіатекстах є складним завданням, але методи NLP мають великий потенціал для вирішення даної проблеми. Використовуючи алгоритми машинного навчання та моделі глибокого навчання, можливо виявляти та аналізувати приховані повідомлення в текстах. Проте вкрай важливо визнати обмеження цих методів і постійно прагнути до вдосконалення, щоб ефективно протидіяти поширенню пропаганди.

Список використаних джерел:

1. Martino, G. D. S., Cresci, S., Barrón-Cedeño, A., Yu, S., Di Pietro, R., & Nakov, P. (2020). A survey on computational propaganda detection. <https://arxiv.org/pdf/2007.08024.pdf>.
2. Chavan, T., & Kane, A. (2022). Large Language Models for Multi-label Propaganda Detection. <https://arxiv.org/pdf/2210.08209.pdf>.
3. Chaudhari, D. D., & Pawar, A. V. (2022). A Systematic Comparison of Machine Learning and NLP Techniques to Unveil Propaganda in Social Media. *Journal of Information Technology Research (JITR)*, 15(1), 1-14. [https://www.researchgate.net/publication/342789189\\_A\\_Survey\\_on\\_Computational\\_Propaganda\\_Detection](https://www.researchgate.net/publication/342789189_A_Survey_on_Computational_Propaganda_Detection).