

УДК 004.8:004.934:004.032.26

DOI: 10.31673/2412-9070.2025.061206

О. О. СНИЦАРЕНКО, аспірант,
ORCID: 0009-0009-1552-9760

Київський національний університет імені Тараса Шевченка

ВИЯВЛЕННЯ ПРОПАГАНДИ В ТЕКСТОВІЙ МОДАЛЬНОСТІ ШЛЯХОМ ЗАСТОСУВАННЯ ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ ТРАНСФОРМЕРІВ

Державна пропагандистська машина російської федерації, маючи багатомільярдні щорічні бюджети, стратегічно використовує соціальну мережу Телеграм як основний канал для створення та поширення пропагандистського контенту. Ця діяльність істотно впливає на сприйняття війни в Україні, міжнародному інформаційному просторі, а також всередині рф. Так як Телеграм досягнув один мільярд активних користувачів, соціальна мережа відіграє ключову роль у формуванні суспільних наративів та потоків інформації. Крім того, сучасний розвиток великих мовних моделей робить розповсюдження пропаганди більш автоматизованим, масовим і швидким. У даній статті досліджуються глибокі нейронні мережі трансформери, їх роль і ефективність у виявленні російської пропаганди. У статті досліджено пропагандистські Телеграм канали, створено і виконано обробку набору даних, що ефективно відображає цифровий слід російської пропаганди у період повномасштабного вторгнення. Розроблена класифікаційна модель штучного інтелекту, яка донаведена на цьому наборі даних, досягає ефективності у виявленні пропаганди рф на рівні 96%.

Ключові слова: глибокі нейронні мережі; трансформери; обчислювальна пропаганда; соціальні мережі; Телеграм.

Вступ

Постановка проблеми. Через популярність соціальних мереж, люди дедалі частіше орієнтуються на думки інших при ухваленні як незначних, так і важливих рішень - наприклад, за кого голосувати на виборах або які ліки використати для лікування. Соціальні мережі стали ефективним інструментом для поширення пропаганди, неправдивої інформації та маніпулювання громадською думкою. Тому, виявлення пропаганди та встановлення достовірності інформації у соціальних мережах набули особливої актуальності, що призвело до активного наукового інтересу у таких напрямках, як фейкові новини [1], клікбейт [3], фальшиві відгуки [4], чутки [5], тощо. Пропаганда базується на використанні символів, емоцій, стереотипів і усталених уявлень, з метою впливу на сприйняття, мислення та поведінку людей. Пропаганда використовує різні техніки впливу, деякі наукові праці виявляють 18 технік [7].

При виявленні пропаганди рф у соціальних мережах, окремої уваги заслуговує соціальна мережа Телеграм, кількість активних користувачів якої у світі досягла 1 млрд [22]. Через повномасштабну російсько-українську війну, вплив і розповсюдженість Телеграму зросли. В Україні, серед опитаних, 81% обирають Телеграм для отримання новин [23]. У рф, де пропаганда і наративи інформаційної війни проти України створюються, Телеграм є головним інструментом розповсюдження політичних новин - добова аудиторія Телеграм зросла з 25 млн. до 48 млн. [21], тобто майже третина населення рф.

Через те, що Телеграм не має строгої модерації і дозволяє створення анонімних інформаційних каналів, проросійські пропагандистські представники використовують Телеграм як основну соціальну мережу [20]. Більше того, після блокування соціальних мереж компанії META, Телеграм став одним із головних каналів розповсюдження інформації в рф.

Проблематика виявлення пропаганди у цифровому середовищі набуває особливої актуальності в умовах гібридної війни. Соціальна мережа Телеграм стала основною платформою по-

© Сніцаренко О. О., 2025

поширення російської пропаганди, спрямованої на формування викривленого уявлення про війну в Україні як серед внутрішньої, так і серед міжнародної аудиторії.

Незважаючи на значний прогрес у галузі обробки природної мови (NLP), більшість існуючих рішень орієнтовані на англomовну текстову модальність, використовують популярні набори даних і не враховують специфіки Телеграм каналів, які поєднують елементи новинного контенту, коментарів та неформального дискурсу.

Усе це зумовлює необхідність створення збалансованого набору даних із Телеграм повідомлень, а також розроблення моделі штучного інтелекту на базі сучасних архітектур глибинних нейронних мереж-трансформерів, для ефективного виявляти пропаганди рф у Телеграм.

Аналіз останніх досліджень і публікацій. Основу сучасних підходів до виявлення пропаганди становлять глибокі нейронні мережі, основані на архітектурі трансформерів. Початком цього напрямку вважається праця [2], у якій уперше запропоновано механізм самоуваги (*self-attention*), що забезпечив ефективне моделювання довгострокових залежностей у тексті, без необхідності застосування рекурентних структур, такі як RNN. Цей підхід започаткував нову епоху у сфері обробки природної мови, дозволивши створювати моделі, здатні узагальнювати контекст на рівні усього документа, а не лише окремих речень, що є основою сучасних великих мовних моделей і так званих засновницьких моделей (*foundation models*).

Подальший розвиток NLP, включно з виявленням пропаганди, відбувся завдяки моделі *BERT*, яка запропонувала двонапрявне попереднє тренування на великих масивах текстів і стала стандартом для задач класифікації тексту, виявлення фейкових новин та пропаганди [15]. Після цього був оптимізований процес попереднього навчання *BERT*, що дозволило суттєво підвищити якість репрезентацій і стабільність тренування [12]. У архітектурі *XLNet* поєднали авторегресивне та авто-кодувальне навчання, досягаючи кращих результатів на семантичних задачах [16]. Надалі було запропоновано ефективний уніфікований підхід, згідно якого будь-яку задачу NLP можна формувати як задачу перетворення тексту на текст, що особливо важливо для задач, пов'язаних із перефразуванням, детекцією пропаганди та маніпуляцій у тексті [17].

Подальші модифікації, зокрема архітектура *ALBERT*, зменшили кількість параметрів шляхом факторизації матриць ембедингів без втрати точності, що зробило великі моделі придатними для тонких лінгвістичних завдань [10]. Для подальшої оптимізації був розроблений *DistilBERT* - стиснена версія *BERT*, яка зберігає більшість можливостей базової моделі, але споживає вдвічі менше обчислювальних ресурсів [8]. У моделі [11] покращили архітектуру завдяки розділенню контентної та позиційної уваги, що дозволило моделі глибше розуміти відносини між словами у складних реченнях.

У контексті багатомовних завдань, дослідження [14, 9] адаптували мультимовні трансформери для слов'янських мов, зокрема для російської, що стало основою для виявлення дезінформації у пострадянському інформаційному просторі. Більш того, у межах дослідження [13] систематизували найкращі практики російськомовної обробки текстів і створили ефективні базові моделі для класифікації пропаганди. *mBERT* модель показала гарний науковий результат при багатомовності [19].

Дослідження [18] доводить, що інтерпретованість результатів моделі є не менш важливою, ніж точність - запропоновано механізми пояснення, які дають змогу аналізувати, які саме пропагандистські техніки (емоційні апеляції, контраст, узагальнення) модель вважає релевантними.

Сукупно ці дослідження сформували актуальну теоретичну і практичну базу для створення сучасних систем виявлення пропаганди у текстовій модальності.

Формулювання мети статті. Метою даної статті є створення і розробка сучасної класифікаційної моделі штучного інтелекту на базі трансформерів для виявлення пропаганди рф у Телеграм повідомленнях, і порівняння ефективності моделей, а саме трансформерів, класичних моделей машинного навчання та ансамблевих моделей. Для тренування і донавчання моделей використовується набір даних Телеграм повідомлень, що відображає збалансований цифровий слід російської пропаганди, яка прямо або опосередковано підтримується державою агресором

під час повномасштабного вторгнення в Україну. Для досягнення цієї мети передбачено такі завдання дослідження:

1. Розробити та застосувати асинхронний алгоритм, який відповідно до заданих параметрів та конфігурацій, збирає і обробляє Телеграм повідомлення та відповідні метадані.
2. У відповідності до санкційних списків ЄС та системного аналізу пропаганди РФ і Телеграм каналів, зібрати та анотувати набір даних, що містять пропагандистські та нейтральні тексти.
3. Провести тренування і донавчання моделей штучного інтелекту на зразках для тренування та валідації.
4. Провести порівняльний експеримент на тестових зразках із використанням класичних методів машинного навчання, ансамблевих алгоритмів і трансформерів щодо ефективності виявлення пропаганди РФ, відповідно до метрик Precision, Accuracy, F1 та Recall.
5. Визначити найбільш ефективну архітектуру і напрями подальшого дослідження.

Основний матеріал

Набір даних. Для створення набору даних Телеграм повідомлень розроблено асинхронний алгоритм, який використовує Телеграм API [24]. Період збору і обробки даних охоплює перший рік від початку повномасштабного вторгнення росії в Україну - з 24 лютого 2022 до 24 лютого 2023 року (2024 і 2025 роки входять до мультимодального дослідження автора). Для класифікації текстів як пропагандистських або незалежних автор спирається на 1) санкційні списки ЄС, в які входять автори каналів, 2) ключові слова та 3) свій системний аналіз Телеграм каналів, який враховує кількість підписників, індекс цитування, загальне охоплення, охоплення по різних країнам, індекс залучення, категорії тощо.

Рис. 1 і Рис. 2 містять частину ключових слів, при наявності яких Телеграм повідомлення класифікується алгоритмом як пов'язане із відповідною категорією - пропаганда або незалежний ресурс.

```
SEED_PHRASES_PROP = [{"денацификация"}, {"демилитаризация"}, {"русский"}, {"мир"}, {"киевский"}, {"режим"}, {"украинский"}, {"фашизм"}, {"нацист"}, {"НАТО"}, {"спецоперация"}, {"биологическая"}, {"зсу"}, {"Азов"}, {"на"}, {"Украина"}]
```

Рис. 1. Ключові слова пропаганди

```
SEED_PHRASES_NON_PROP = [{"российский"}, {"вторжение"}, {"аннексия"}, {"правительство"}, {"украина"}, {"война"}, {"аннексия"}, {"в"}, {"украина"}, {"легитимный"}, {"власть"}, {"легитимный"}, {"власть"}, {"в"}, {"украина"}, {"в"}, {"украине"}, {"легитимный"}, {"президент"}, {"легитимный"}, {"президент"}, {"в"}, {"украина"}, {"трибунал"}, {"военный"}, {"преступление"}, {"гуманитарный"}, {"катастрофа"}, {"пропагандист"}, {"в"}, {"украина"}, {"незаконный"}, {"референдум"}, {"нелегитимный"}, {"референдум"}]
```

Рис. 2. Ключові слова незалежних ресурсів

Ключові слова для відбору повідомлень пропагандистських каналів були обрані автором, враховуючи новини з російсько-української війни, тобто дані слова повинні бути актуальними для пропаганди відповідного часового періоду та підтверджені трьома великими мовними моделями (LLM) - ChatGPT [25], Google Gemini [26], та Claude [27]. Ключові слова незалежних ресурсів системно виключають пропаганду і не мають явного політичного забарвлення. Приклад таблиці, де пропагандистські повідомлення агрегуються алгоритмом для подальшої обробки наведені на рис. 3.

Статистика даних. У таблиці 1 представлено основну статистику набору даних, який використовується для тренування, валідації і донавчання моделей, а також для порівняння їхньої ефективності на тестових зразках. Таблиця 1 показує кількість зразків у кожному класі для

кожної вибірки й загалом середню кількість слів на зразок, а також максимальну, мінімальну та унікальну кількість слів.

message_id	channel	date	text	media	has_image	has_video	weak_label_channel	seed_phrase
15814	147763	@rian_ru	2022-02-24 02:57:23+00:00	Путин: Россия будет стремиться к демилитаризац...	{'image': False, 'video': False}	False	False	propaganda [демилитаризация]
15815	147789	@rian_ru	2022-02-24 03:49:48+00:00	Постпред Украины при ООН, комментируя выступле...	{'image': False, 'video': False}	False	False	propaganda [спецоперация]
15816	147801	@rian_ru	2022-02-24 04:03:31+00:00	Путин объявляет о начале специальной военной о...	{'image': False, 'video': True}	False	True	propaganda [денацификация]
15817	147977	@rian_ru	2022-02-24 07:38:13+00:00	Пушилин: штаб украинской спецоперации в Донб...	{'image': False, 'video': False}	False	False	propaganda [спецоперация]
15818	148015	@rian_ru	2022-02-24 08:21:37+00:00	Лукашенко заявил, что** белорусские войска н...	{'image': False, 'video': False}	False	False	propaganda [спецоперация]
...	...	...	...	...	...	...	...	...
23928	141402	@SolovievLive	2022-11-23 07:57:06+00:00	! **Проект «Ставка». Обзор ситуации на фронтах...	{'image': False, 'video': True}	False	True	propaganda [спецоперация]
23929	141407	@SolovievLive	2022-11-23 08:07:32+00:00	☀ **Пенсионерка с Урала пожертвовала деньги ра...	{'image': True, 'video': False}	True	False	propaganda [спецоперация]
23930	141410	@SolovievLive	2022-11-23 08:39:04+00:00	**Пономарев* собрался идти на Кремль с несущес...	{'image': False, 'video': True}	False	True	propaganda [спецоперация]
23931	141569	@SolovievLive	2022-11-23 16:56:34+00:00	Сегодня испугавшиеся наших ударов по объектам ...	{'image': False, 'video': False}	False	False	propaganda [киевский, режим]
23932	141586	@SolovievLive	2022-11-23 18:34:33+00:00	** !! **** !! **** **Путин информировал участни...	{'image': False, 'video': False}	False	False	propaganda [спецоперация]

Рис. 3. Агрегація пропагандистських повідомлень

Статистика набору даних

Зразки у тренуванні	Зразки у валідації	Зразки у тестуванні	Загальна кількість
9213	967	2324	12504
мін кількість слів	мак кількість слів	середній показник слів	кількість унікальних слів
8	715	98	92311

У таблиці 2, у позитивному класі (пропаганда) міститься 7502 текстів, тоді як у зразках з незалежних ресурсів - 5002.

Баланс класів

Етап тренування	Етап тренування	Етап валідації	Етап валідації	Етап тестування	Етап тестування
пропаганда	незалежний ресурс	пропаганда	незалежний ресурс	пропаганда	незалежний ресурс
5527	3685	580	386	1394	929

Аналіз довжини текстів (у символах) показав, що пропагандистські Телеграм повідомлення, як правило, довші, тоді як незалежні - коротші, з рівномірнішим розподілом довжини.

Векторизація даних. Для класичних моделей машинного навчання та ансамблевих моделей використовується метод векторизації TF-IDF, щоб представити текстову модальності у вигляді статичних векторів - числового формату, який подається на вхід традиційним класифікаторам SVM, LG, Decision Tree, Random Forest та XGBoost. TF-IDF - це метод, який підсилює вагу слова у тексті пропорційно частоті його появи у документі, але зменшує її залежно від того, скільки документів у корпусі містять це слово. Кожен текст розглядається як окремий документ, а весь корпус - як основа для обчислення вагів.

Моделі трансформери не потребують такої векторизації, так як вони вже містять ембедінги-вектори, які під час донавчання динамічно підлаштовуються разом із класифікаційною головою (додатковим модулем класифікації), щоб максимізувати точність кінцевого результату класифікації.

Моделі. Для визначення ефективності глибоких нейронних мереж трансформерів у виявленні пропаганди рф у текстовій модальності Телеграму, використано наступні моделі і алгоритми: традиційні алгоритми машинного навчання, а саме: SVM, LR, Decision Tree; ансамблеві моделі: Random Forrest, XGBoost; тонке донавчання BERT-моделей - застосування попередньо навчених моделей безпосередньо до набору даних, зокрема: модель BERT [15]; модель DeepPavlov [13], попередньо навчена на новинному корпусі російською мовою та інших тематичних підкорпусах; RuBERT [14]; багатомовну модель mBERT [19].

Результати дослідження. У таблиці 3 наведено результати дослідження ефективності виявлення пропаганди для класичних моделей, ансамблевих моделей і трансформерів, за метриками Accuracy, Precision, Recall та F1. Серед класичних моделей найкращий результат показала SVM (TF-IDF), серед ансамблевих моделей - XGBoost (TF-IDF).

Результати застосування класичних моделей, ансамблевих моделей та трансформерів до тестових зразків

Модель	Accuracy	Precision	Recall	F1
<i>Класичні моделі</i>				
Logistic Regression (TF-IDF)	82.5	80.3	79.1	79.9
SVM (TF-IDF)	84.1	81.0	80.1	80.5
Decision Tree (TF-IDF)	80.9	76.8	76.2	76.5
<i>Ансамблеві моделі</i>				
Random Forest (TF-IDF)	83.2	82.1	81.3	81.7
XGBoost (TF-IDF)	84.4	83.2	82.6	82.9
<i>Трансформери</i>				
BERT (донавчена)	85.7	87.4	88.1	89.7
DeepPavlov (донавчена)	96.3	95.1	94.0	93.3
mBERT (донавчена)	87.2	85.5	84.6	83.7
RuBERT (донавчена)	93.1	92.5	92.2	91.5

Найкращими серед трансформерів є 1) DeepPavlov - модель, яка добре себе зарекомендувала, так як її первинне тренування відбувалося російською мовою та завдяки іншим архітектурним рішенням та 2) багатомовна модель RuBERT. Обидві моделі навчалися протягом 15 епох, з розміром батчу 16 та швидкістю навчання $2e-5$. Модель DeepPavlov досягла точності 96.3%.

Порівняння результатів усіх моделей за метриками Accuracy, Precision, Recall та F1 наведені на рис. 4.

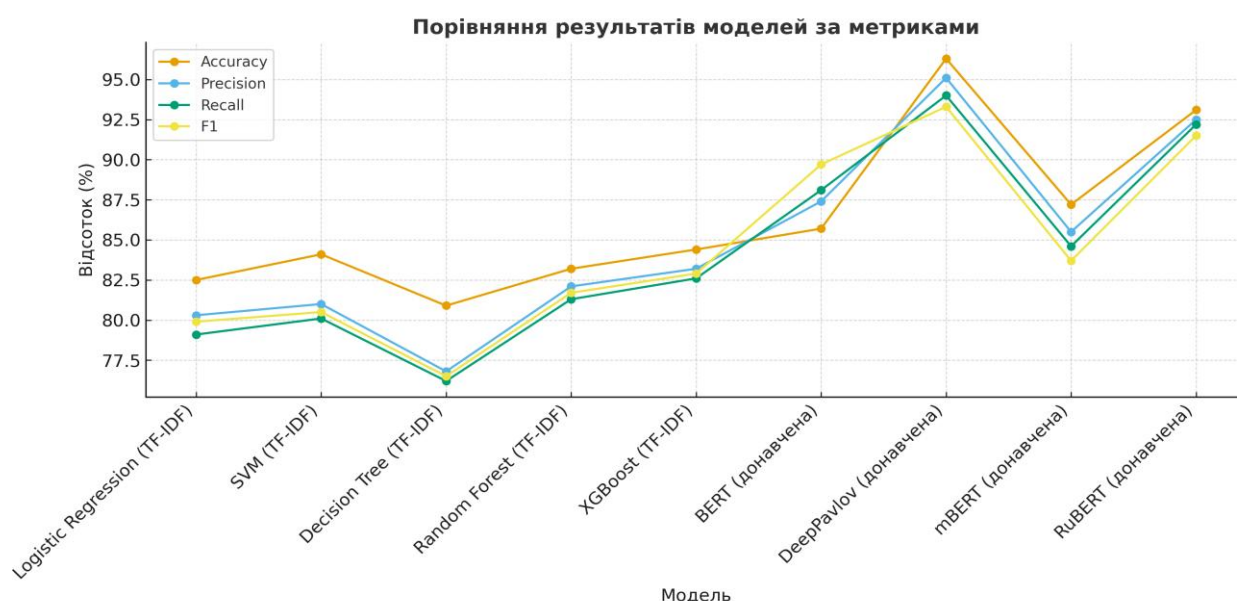


Рис. 4. Порівняння результатів моделей

Висновки

Соціальна мережа Телеграм є головним інструментом розповсюдження пропаганди рф під час російсько-української війни.

Результати дослідження продемонстрували, що моделі трансформерів DeepPavlov і ruBERT є ефективнішими на 4-8%, ніж інші моделі трансформерів у виявленні пропаганди рф за основними метриками та обидві моделі демонструють вищу ефективність, ніж базові класичні й ансамблеві моделі, на рівні 10-15%.

Створений набір даних, за допомогою розробленого асинхронного алгоритму та відповідних параметрів й конфігурацій обробки даних з Телеграм каналів, робить запропоновані алгоритми машинного навчання ефективними у виявленні пропаганди рф на рівні 80-96%.

Подальше дослідження доцільно спрямувати на додавання мультимодальних аспектів (текст та зображення або метадані), використання крослінгвістичних моделей і розроблення систем онлайн-моніторингу пропагандистських кампаній у соціальних мережах. Враховуючи результати і висновки цієї статті, подальше дослідження буде стосуватися збільшення і покращення набору даних пропаганди країни агресора; додавання інших модальностей; використання інших глибоких нейронних мереж; оптимізації алгоритмів; процесінгу у реальному часі та використання великих мовних моделей для обробки даних.

Список літератури

1. Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). *The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans*. *Journal of Data and Information Quality (JDIQ)*, 11(3).
2. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. *Attention Is All You Need // Advances in Neural Information Processing Systems (NeurIPS 2017)*. — Red. Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. — Long Beach, CA, USA: Curran Associates, Inc., 2017. — P. 5998–6008. — ISBN 978-1-5108-6096-4.
3. Chen, Y., Conroy, N. J., & Rubin, V. L. (2015). *Misleading online content: Recognizing clickbait as false news*. In: MDD.
4. Rubin, V. L., Conroy, N. J., & Chen, Y. (2015). *Towards news verification: Deception detection methods for news discourse*.
5. Wu, K., Yang, S., & Zhu, K. Q. (2015). *False rumors detection on Sina Weibo by propagation structures*. In: *Proceedings of the IEEE 31st International Conference on Data Engineering (ICDE'15)*. IEEE, pp. 651–662.
6. Geissler, D., Bar, D., Pröllochs, N., & Feuerriegel, S. (2022). *Russian propaganda on social media during the 2022 invasion of Ukraine*. arXiv preprint arXiv:2211.04154.
7. Da San Martino, G., Yu, S., Barrón-Cedeño, A., Petrov, R., & Nakov, P. (2019). *Fine-Grained Analysis of Propaganda in News Articles*. Association for Computational Linguistics. DOI: 10.18653/v1/D19-1565.
8. Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). *DistilBERT: A distilled version of BERT: smaller, faster, cheaper and lighter*. arXiv preprint arXiv:1910.01108.
9. Zaytsev, A., Kuratov, Y., Ustalov, D., & Kutuzov, A. (2018). *DeepPavlov at RusEval 2018: Russian Question Answering and Machine Reading Evaluation*. In: *First Workshop on Evaluation of Human Language Technologies for Slavic Languages*, pp. 1–6.
10. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). *ALBERT: A Lite BERT for Self-Supervised Learning of Language Representations*. In: *8th International Conference on Learning Representations (ICLR)*. Addis Ababa, Ethiopia: OpenReview.net.
11. He, P., Liu, X., Gao, J., & Chen, W. (2021). *DeBERTa: Decoding-Enhanced BERT with Disentangled Attention*. arXiv preprint arXiv:2006.03654.
12. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach*. arXiv preprint arXiv:1907.11692.

13. Chernykh, P., Mikhailovskaya, A., Miftahutdinov, Z., & Ustalov, D. (2021). DeepPavlov Goes to Russia: Baselines and Best Practices for Russian Language Processing. In: *International Conference on Computational Linguistics and Intelligent Text Processing*, pp. 176–187. Springer.
14. Arkhipov, M., Trofimova, M., Kuratov, Y., & Sorokin, A. (2019). Tuning multilingual transformers for language-specific named entity recognition. In: *Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing*, pp. 89–93.
15. Jacob Devlin et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 4171–4186.
16. Yang, Z., Dai, Z., Yang, Y., Carbonell, J. G., Salakhutdinov, R., & Le, Q. V. *XLNet: Generalized Autoregressive Pretraining for Language Understanding* // *Advances in Neural Information Processing Systems (NeurIPS 2019)*. — Red. Wallach, H., Larochelle, H., Beygelzimer, A., d’Alché-Buc, F., Fox, E., Garnett, R. — Vancouver, Canada: Curran Associates, Inc., 2019. — P. 5753–5763. — arXiv:1906.08237.
17. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer (T5) // *Journal of Machine Learning Research (JMLR)*. — 2020. — Vol. 21, No. 140. — P. 1–67. — arXiv:1910.10683.
18. Yu, S., Da San Martino, G., Barrón-Cedeño, A., & Nakov, P. (2021). Interpretable Propaganda Detection in News Articles. In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pp. 1597–1605.
19. Pires, T., Schlinger, E., & Garrette, D. How Multilingual is Multilingual BERT? // *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*. — Florence, Italy: Association for Computational Linguistics, 2019. — P. 4996–5001. — DOI: 10.18653/v1/P19-1493.
20. [Онлайн ресурс] What pro-Russian Telegram channels are sharing ahead of the German parliamentary election. Texty.org.ua. URL: <https://texty.org.ua/articles/114515/>
21. [Онлайн ресурс] В топ-10 Telegram-каналов РФ отсутствует не только оппозиционный, но даже адекватный контент – The Bell. URL: <https://uatv.ua/v-top-10-telegram-kanalov-rf-otsutstvuet-ne-tolko-oppozitsionnyj-no-dazhe-adekvatnyj-kontent-detali-ot-spetskora-the-bell>
22. [Онлайн ресурс] Кількість активних користувачів Telegram перевищила мільярд. Forbes.ua, 19.03.2025. URL: <https://forbes.ua/news/kilkist-aktivnikh-koristuvachiv-telegram-perevishchila-milyard-19032025-28148>
23. [Онлайн ресурс] Україна на шостому місці у світі за кількістю Telegram-каналів із понад мільйоном підписників – Telemetro. Detector Media, 19.12.2023. URL: <https://ms.detector.media/sotsmerezhi/post/33786/>
24. Telegram [Електронний ресурс] // Офіційний вебсайт платформи Telegram. — Режим доступу: <https://core.telegram.org/> (дата звернення: 08.10.2025).
25. OpenAI ChatGPT [Електронний ресурс] // Офіційний вебсервіс компанії OpenAI для взаємодії з мовною моделлю GPT. — Режим доступу: <https://chat.openai.com/> (дата звернення: 08.10.2025).
26. Google Gemini [Електронний ресурс] // Офіційний вебсайт багатомодальної моделі штучного інтелекту Google Gemini. — Режим доступу: <https://gemini.google.com/> (дата звернення: 08.10.2025).
27. Claude AI [Електронний ресурс] // Офіційний вебсайт компанії Anthropic для доступу до мовної моделі Claude AI. — Режим доступу: <https://claude.ai/> (дата звернення: 08.10.2025).

*O. Snitsarenko***PROPAGANDA DETECTION IN TEXTUAL MODALITY ON TELEGRAM BY MEANS OF TRANSFORMERS**

The state propaganda machine of the Russian Federation, supported by multi-billion annual budgets, strategically employs the Telegram platform as the primary channel for creating and disseminating propagandistic content. This activity significantly affects the perception of the war in Ukraine, the international information environment, and within Russia itself. Since Telegram already has one billion active users, the platform plays a crucial role in shaping public narratives and information flows. Moreover, the modern development of large language models makes the dissemination of propaganda more automated and rapid. This paper explores deep transformer-based neural networks, their role, and effectiveness in detecting Russian propaganda, using an algorithmically created dataset of Telegram channels that reflects the digital footprint of Russian propaganda during the full-scale invasion period. An effective transformer-based model is proposed, post-trained, and evaluated on this dataset.

The results of the study demonstrated that certain transformer-based models outperform other multilingual models, including other transformers. These models also show higher effectiveness compared to classical and ensemble baselines. Moreover, the developed approach to processing and constructing efficient Telegram datasets—through an asynchronous algorithm with well-developed systematic approach, the randomized sampling and adherence to data construction principles and configurations—has proven to make machine learning algorithms highly effective, achieving accuracy levels between 80% and 96%.

Future research should focus on expanding multimodal aspects (text combined with images and metadata), exploring cross-linguistic transformer models, and developing online monitoring systems for tracking propaganda campaigns in social networks.

Keywords: deep neural networks; transformers; computational propaganda; social media; Telegram.
