

**Катков Ю.І.,**

доктор технічних наук, доцент, доцент  
кафедри Комп'ютерних наук,  
Державний університет інформаційно-  
комунікаційних технологій

**Бай Я.В.,**

аспірант кафедри Комп'ютерних наук,  
Державний університет інформаційно-  
комунікаційних технологій

(м.Київ, Україна)

## **АВТОМАТИЧНА КЛАСИФІКАЦІЯ ПОВІДОМЛЕНЬ ПРО ПОВІТРЯНІ ЗАГРОЗИ У TELEGRAM-КАНАЛАХ**

*Вступ.* В умовах сучасних воєнних конфліктів швидкість та точність обробки інформації про повітряні загрози стає критично важливою для забезпечення безпеки цивільного населення. Дослідження присвячене розробці системи автоматичної класифікації повідомлень про повітряні загрози у Telegram-каналах з використанням методів машинного навчання та обробки природної мови. Запропонована система досягла точності класифікації 88.2% та може обробляти до 1000 повідомлень на хвилину в режимі реального часу.

Telegram-канали стали основним джерелом інформації про повітряні загрози в Україні під час російсько-української війни. Щоденно публікується тисячі повідомлень різного характеру: від офіційних сповіщень про повітряну тривогу до неперевірених чуток та дезінформації. Автоматична класифікація таких повідомлень є критично важливою задачею для забезпечення швидкого реагування на реальні загрози, фільтрації неперевірених даних та дезінформації, підтримки ситуаційної обізнаності органів влади та зменшення паніки серед населення

Основними викликами є багатозначність української мови, використання специфічної військової термінології, наявність емоційно забарвленого контенту та швидка еволюція мовних патернів у воєнний час.

*Методологія.* Для дослідження було зібрано датасет з 50000 повідомлень з 25 українських Telegram-каналів за період 2022-2024 року. Повідомлення були розмічені експертами на 6 категорій:

1. Офіційні сповіщення (15000 повідомлень) – повідомлення від органів влади;
2. Перевірена інформація (12000) – підтверджена інформація від надійних джерел;

3. Непереверені дані (8000) – чутки та не підтверджена інформація;
4. Дезінформація (5000) – навмисно хибна інформація;
5. Емоційні реакції (7000) – коментарі та емоційні відгуки;
6. Нерелевантний контент (3000) – повідомлення не пов'язані з загрозами.

Запропонована система базується на поєднанні трьох компонентів:

1. Модель-трансформер – змінена модель на основі існуючої моделі `uk-bert-base` для розуміння семантики;
2. CNN модель – для виявлення локальних текстових патернів;
3. Модуль часових ознак – для аналізу залежностей в часі.

Фінальна класифікація здійснюється з такими ваговими коефіцієнтами: BERT – 0.5, CNN – 0.3, часові ознаки – 0.2.

Система реалізована з використанням таких складників мови програмування Python 3.9 та фреймворків PyTorch, Transformers, FastAPI. Для обробки великих обсягів даних використано архітектуру мікросервісів з горизонтальним масштабуванням.

*Експериментальні результати.* Наведемо отримані результати тестування на тестовому наборі даних у таблиці нижче.

Категорія	Точність	Повнота	Оцінка F1
Офіційні сповіщення	0.873	0.867	0.870
Перевірена інформація	0.824	0.818	0.821
Непереверені дані	0.891	0.805	0.898
Дезінформація	0.867	0.852	0.859
Емоційні реакції	0.856	0.874	0.765
Нерелевантний контент	0.882	0.886	0.884

Отже, загальна точність становить 88.2%. Порівняємо отримані результати з іншими методами:

Метод	Точність	Оцінка F1	Час висновку
Naive Bayes	0.741	0.723	8.6 мс
SVM + TF-IDF	0.823	0.809	12.5 мс
LSTM	0.865	0.806	30.4 мс
BERT-base	0.878	0.815	80.7 мс
Розроблена модель	0.882	0.833	120.1 мс

Найчастіші помилки спостерігаються при класифікації сатиричних повідомлень (часто класифікуються як дезінформація), повідомлень з подвійним значенням (неоднозначність інтерпретації) та нових термінів та неологізмів воєнного часу.

*Висновки.* Розроблена система автоматичної класифікації повідомлень про повітряні загрози демонструє високу ефективність та практичну значущість. В рамках цього дослідження була досягнута точність 88.2% на реальних даних з Telegram-каналів, забезпечена обробка великих обсягів даних в режимі реального часу, створено спеціалізовану систему для української мови.

Напрямами подальших досліджень можуть стати:

- інтеграція мультимодальних даних (зображення, аудіо);
- розробка механізмів пояснюваності рішень;
- адаптація для інших типів надзвичайних ситуацій;
- дослідження впливу дезінформації на поведінку населення.

Практична значущість дослідження полягає в можливості значного покращення систем цивільного захисту та забезпечення інформаційної безпеки в умовах воєнного стану.

***Список використаних джерел:***

1. Aggarwal C.C., Wang H. Text mining in social networks. Social Network Data Analytics. Boston, MA, 2011. P. 353–378.
2. Attention is all you need / A. Vaswani et al. Advances in Neural Information Processing Systems. 2017.
3. Bert: Pre-training of deep bidirectional transformers for language understanding / J. Devlin et al. 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). 2019. P. 4171–4186.