

**Materials of the XII International Scientific Conference
«Information-Management Systems and Technologies»
23th – 25th September, 2024, Odesa**

The findings also suggest broader implications for the automotive industry, including the development of smarter, more responsive vehicle systems that can adapt to changing conditions and driver behaviors. Future research could explore the further refinement of AI algorithms to enhance their predictive capabilities and extend their applicability to other aspects of vehicle performance, such as battery health monitoring and autonomous driving systems.

References

- [1] M. Lungu and C. Mariana, "Real-time applications using artificial intelligence," in *International Journal of Advanced Research in Artificial Intelligence*, vol. 5, no. 2, pp. 45-56, 2021.
- [2] R. Collobert, J. Weston, and L. Bottou, "Deep Learning for Chatbots," in *Journal of Machine Learning Research*, vol. 21, pp. 3765-3789, 2020.
- [3] AbuGhoush, R., & Abu Arqoub, R. Integrating AI-Based Models into Software Development for Enhancing Real-Time Applications. *International Journal of Advanced Computer Science and Applications (IJACSA)* 14(3): 45-52. doi:10.14569/IJACSA.2023.014031, 2023.
- [4] Artstein, R., Gandhe, S., & Traum, D. Dialogue Management for Real-Time Applications. *Journal of Artificial Intelligence Research* 73: 237-269. doi:10.1613/jair.1.11483, 2022.

УДК 004.8

**МЕТОД ІНТЕЛЕКТУАЛЬНОГО ВИЯВЛЕННЯ ТА
КЛАСИФІКАЦІЇ КІБЕРЗАЛЯКУВАНЬ У ТЕКСТОВОМУ
КОНТЕНТІ**

О. Собко^[0000-0001-5371-5788]

*Хмельницький національний університет, Україна
EMAIL: olena.sobko.ua@gmail.com*

**THE METHOD FOR INTELLIGENT DETECTION AND
CLASSIFICATION OF CYBERBULLYING IN TEXT CONTENT**

O. Sobko

Khmelnytskyi National University, Ukraine

**Materials of the XII International Scientific Conference
«Information-Management Systems and Technologies»
23th – 25th September, 2024, Odesa**

***Анотація.** У роботі розглянуто проблему розповсюдження кіберзалякувань у мережі Інтернет. Зроблено огляд існуючих підходів до виявлення кіберзалякувань у текстових повідомленнях користувачів, а також запропоновано метод інтелектуального виявлення та класифікації кіберзалякувань у текстовому контенті з використанням підходів машинного навчання.*

***Ключові слова:** кіберзалякування, мультикласова класифікація, текстове повідомлення, тип кіберзалякування, машинне навчання.*

***Abstract.** This paper examines the problem of the spread of cyberbullying on the Internet. An overview of existing approaches to the detection of cyberbullying in text messages of users is made, and the method for intelligent detection and classification of cyberbullying in text content using machine learning approaches is proposed.*

***Keywords:** cyberbullying, multiclass classification, text message, type of cyberbullying, machine learning.*

Кіберзалякування є значною проблемою сучасного інформаційного суспільства, що набуває все більшого поширення у зв'язку з активним використанням соціальних мереж, месенджерів та інших онлайн-платформ. Інтернет-користувачі, особливо молодь, дедалі частіше стають жертвами агресивних дій, що включають образи, погрози, залякування та приниження, які здійснюються через мережу. Така форма агресії має значний вплив на психічне здоров'я людей, призводячи до підвищеної тривожності, депресії, зниження самооцінки та навіть суїцидальних думок [1].

Важливо зазначити, що кіберзалякування може мати довготривалі наслідки, які відчуваються навіть після завершення агресивних дій, оскільки негативний контент може зберігатися в інтернеті та продовжувати впливати на жертву [2].

Кіберзалякування проявляється у різних формах, кожна з яких має свої особливості та може по-різному впливати на жертву. Зокрема, він може бути спрямований на жертв через їхню етнічну або релігійну приналежність.

Вікове підгрунття кіберзалякування часто стосується підлітків або літніх людей, які є менш захищеними в цифровому середовищі. Гендерне підгрунття включає дискримінацію за статтю або гендерною ідентичністю. Виявлення кіберзалякувань є актуальною та важливою задачею, оскільки його поширення негативно впливає на суспільство та окремих індивідів [3, 4].

**Materials of the XII International Scientific Conference
«Information-Management Systems and Technologies»
23th – 25th September, 2024, Odesa**

Метою роботи є розробка методу інтелектуального виявлення та класифікації кіберзалякувань у текстовому контенті. Кроки розробленого методу подано на рис. 1.

Вхідними даними методу інтелектуального виявлення та класифікації кіберзалякувань у текстовому контенті є векторизатор TFID, який перетворює текстовий зразок на числовий формат, що є придатним для подальшої обробки моделлю машинного навчання. Модель машинного навчання BERT навчена на наборі даних Cyberbullying Classification [5], що має такі класи кіберзалякувань: Age, Ethnicity, Gender, Religion, Other type of cyberbullying, Not cyberbullying.

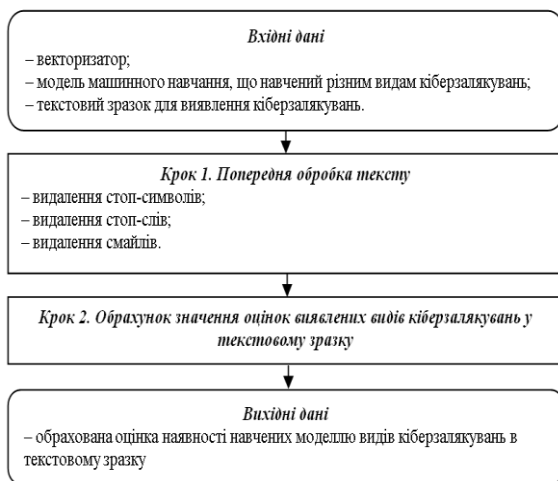


Рисунок 1. Кроки методу інтелектуального виявлення та класифікації кіберзалякувань у текстовому контенті

Після попередньої обробки текстового зразка, а саме очищення від стоп-слів та символів, видалення смайлів відбувається обрахунок значення оцінок виявлених видів кіберзалякувань у текстовому зразку, таким чином, відбувається мультикласова класифікація кіберзалякувань у текстовому зразку. Було проведено аналіз ефективності розробленого методу інтелектуального виявлення та

**Materials of the XII International Scientific Conference
«Information-Management Systems and Technologies»
23th – 25th September, 2024, Odesa**

класифікації кіберзалякувань у текстовому контенті, які показали наступні результати макрометрик: Accurasy 0.94, Precision 0.93, Recall 0.93, F1 0.93. Результати мультикласової класифікації кіберзалякувань мікрометрик наведено в таблиці 1.

Таблиця 1. Результати мікрометрик мультикласової класифікації кіберзалякувань

	Precision	Recall	F1
Religion	0.95	0.97	0.96
Age	0.99	0.98	0.98
Ethnicity	0.99	0.99	0.99
Gender	0.91	0.91	0.91
Not cyberbullying	0.83	0.82	0.83

Отже, згідно проведеного дослідження ефективності, розроблений метод інтелектуального виявлення та класифікації кіберзалякувань у текстовому контенті з використання моделі машинного навчання на основі BERT, показує точність понад 82% для мультикласового виявлення типів кіберзалякувань.

Література

- [1] Darcy & Roy Press. URL: <https://drpress.org/ojs/index.php/EHSS/article/download/17833/17343> (дата звернення: 29.08.2024).
- [2] Method for cyberbullying neuronetwork detection using cloud services and object-oriented model / М. МОЛЧАНОВА et al. Herald of Khmelnytskyi National University. Technical sciences. 2024. Vol. 333, no. 2. P. 200–206. URL: <https://doi.org/10.31891/2307-5732-2024-333-2> (date of access: 29.08.2024).
- [3] ProTect: a hybrid deep learning model for proactive detection of cyberbullying on social media / Т. Nitya Harshitha et al. Frontiers in Artificial Intelligence. 2024. Vol. 7. URL: <https://doi.org/10.3389/frai.2024.1269366> (date of access: 29.08.2024).
- [4] Aliyeva Ç. O., Yağanoğlu M. Deep learning approach to detect cyberbullying on twitter. Multimedia Tools and Applications. 2024. URL: <https://doi.org/10.1007/s11042-024-19869-3> (date of access: 29.08.2024).
- [5] Cyberbullying Classification. Kaggle: Your Machine Learning and Data Science Community. URL: <https://www.kaggle.com/datasets/andrewmvd/cyberbullying-classification> (date of access: 29.08.2024).