ВИСНОВОК
про наукову новизну, теоретичне та практичне значення результатів дисертації

на тему «Методи виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту»
(назва роботи)

здобувача наукового ступеня доктора філософії

Собко Олени Віталіївни
(прізвище, ім'я, по батькові)

з галузі знань 12 – Інформаційні технології
(шифр, назва галузі знань)

за спеціальністю 122 – Комп'ютерні науки
(шифр, назва спеціальності)

Публічна презентація проведена на міжкафедральному фаховому науково-практичному семінарі кафедри комп'ютерних наук та кафедри комп'ютерної інженерії та інформаційних систем
назва)

«14» квітня 2025 року, протокол № 2

1. Актуальність теми дослідження.

Кіберзалякування стали однією з найбільш поширених форм агресивної поведінки в інтернеті в останні роки. За даними досліджень, приблизно 20–40 % підлітків у всьому світі стають жертвами кіберзалякувань, що негативно впливає на їхнє психічне здоров'я та соціальну взаємодію. Ефективна протидія цьому явищу є складним завданням через анонімність і швидкість поширення контенту в цифрових середовищах. Одним із найбільш перспективних підходів до виявлення кіберзалякувань є автоматизований аналіз текстового контенту із застосуванням технологій обробки природної мови. Зокрема, трансформерні моделі, такі як BERT, дозволяють враховувати контекст повідомлень і семантичні залежності, що сприяє підвищенню точності класифікації образливого контенту. Такі системи вже демонструють високі показники у виявленні агресивної поведінки в текстах соціальних мереж та месенджерів.

Однак, незважаючи на високу ефективність, автоматизовані системи стикаються з низкою проблем. Зокрема, проблеми культурної адаптації алгоритмів та залежність від якісного набору навчальних даних можуть впливати на результати аналізу. Крім того, такі моделі часто сприймаються як «чорні ящики», оскільки їхні результати важко інтерпретувати. Відсутність прозорих механізмів пояснення негативно впливає на їхнє впровадження в системи модерації контенту або правозахисні ініціативи. Тому важливо не лише розробляти ефективні методи виявлення кіберзалякувань, але й забезпечувати їхню відповідність етичним стандартам та пояснюваність. Нові програмні рішення, що розробляються для протидії кіберзалякувань, повинні бути не лише точними у їх виявленні, але й етичними, справедливими та поясненними. Вони повинні враховувати різні соціально-демографічні підгрупи, щоб уникнути будь-яких упереджень або дискримінації. Алгоритми повинні діяти прозоро, надаючи чіткі пояснення своїх рішень, забезпечуючи дотримання прав користувачів та захист особистої

інформації, а також запобігаючи можливим зловживанням. Як показав аналіз існуючих досліджень, на сьогодні немає рішень, які б в повній мірі забезпечували етику, справедливість та поясненість у задачі виявлення кіберзалякування.

2. Зв'язок роботи з науковими програмами, планами, темами.

Наведені в дисертації дослідження проводились в рамках виконання науково-дослідної роботи по темі №0121U112025 «Розроблення інформаційної технології прийняття контрольованих людиною критично-безпекових рішень за ментально-формальними моделями машинного навчання», у якій Собко О.В. була виконавцем і розробляла нейромережеві архітектури моделей глибокого навчання та виконувала їх опис, які були використані у дисертаційному дослідженні.

3. Наукова новизна отриманих результатів.

У дисертації одержані такі нові наукові результати:

1. Вперше запропоновано метод оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості, що забезпечує недискримінацію за віковою, гендерною, релігійною приналежністю, що дозволило підвищити якість навчання класифікаторів для виявлення кіберзалякувань.

2. Розроблено новий метод виявлення кіберзалякувань у текстовому контенті, який відрізняється від існуючих двоетапним виявленням кіберзалякувань, що полягає у нейромережевій ідентифікації наявності кіберзалякувань та подальшій нейромережевій мультилейбловій класифікації окремих типів кіберзалякувань, що дало можливість підвищити точність та якість виявлення кіберзалякувань.

3. Удосконалено метод інтерпретації результатів виявлення кіберзалякувань, який відрізняється від існуючих можливістю надавати візуальні пояснення для мультилейблової класифікації виявлених типів кіберзалякувань в альтернативних поданнях.

4. Теоретичне та практичне значення результатів дисертації.

Практичне значення отриманих результатів полягає у доведенні теоретичних результатів дисертаційної роботи та розробці інтелектуальної інформаційної системи виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту, що використовує розроблені методи оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості, виявлення і класифікації кіберзалякувань, а також інтерпретації результатів виявлення кіберзалякувань, та дозволяє підвищити точність та якість виявлення кіберзалякувань у текстовому контенті засобами штучного інтелекту й візуально пояснювати прийняті рішення.

Розроблена інтелектуальна інформаційна система для виявлення та класифікації кіберзалякувань у текстовому контенті. Інтелектуальна інформаційна система надає можливість оцінювати та коригувати репрезентативність датасетів для навчання моделей машинного навчання за етичними аспектами FATE-принципом справедливості; виявляти та

класифікувати типи кіберзалякувань у текстовому контенті. Також інтелектуальна інформаційна система дозволяє отримувати візуальні пояснення для мультилейблової класифікації виявлених типів кіберзалякувань, що сприяє підвищенню довіри до одержаних результатів класифікації типів кіберзалякувань.

5. Використання результатів роботи.

Теоретичні та практичні результати дослідження впровадженні у діяльності відділу протидії кіберзлочинам у Хмельницькій області Департаменту кіберполіції Національної поліції України; у ПП «Авіві» (довідка про впровадження); у ГО «ІТ-кластер міста Хмельницького» (довідка про впровадження); у ТОВ «Системи для бізнесу 2» (довідка про впровадження); у навчальному процесі Хмельницького національного університету (акт впровадження); при виконанні держбюджетної теми Хмельницького національного університету «Розроблення інформаційної технології прийняття контрольованих людиною критично-безпекових рішень за ментально-формальними моделями машинного навчання» (ДР № 0121U112025).

6. Особиста участь автора в одержанні наукових та практичних результатів, що викладені в дисертаційній роботі Собко О.В.
Дисертаційна робота виконана *на кафедрі комп'ютерних наук Хмельницького національного університету*,
(назва кафедри (відділу), назва установи)
наукові керівники д.т.н, професор, завідувач кафедри комп'ютерних наук Бармак О.В.; доктор філософії, старший дослідник Школи права Талліннського технічного університету (Естонія) Арчіл Чочіа
(науковий ступінь, вчене звання, посада, прізвище, ініціали)

Розглянувши звіт подібності щодо перевірки на плагіат, встановлено, що дисертаційна робота Собко О.В.
(прізвище, ініціали здобувача)
є результатом самостійних досліджень здобувача і не містить елементів плагіату та запозичень. Використані ідеї, результати і тексти інших авторів мають посилання на відповідне джерело.

Дисертація характеризується єдністю змісту та відповідає вимогам щодо її оформлення.

7. Перелік публікацій за темою дисертації із зазначенням особистого внеску здобувача.

За результатами досліджень опубліковано 11 наукових праць, серед яких 4 статті у фахових наукових журналах України, включених на дату опублікування до переліку наукових фахових видань України категорії Б. Крім цього, результати дисертації викладені у 5-х публікаціях, які засвідчують апробацію матеріалів дисертації (статті в матеріалах конференцій, що індексуються в наукометричній базі Scopus) та у 2-х свідоцтвах про реєстрацію авторського права на твір.

1. Собко О. В. Нейромережевий пошук і класифікація кіберзалякувань у текстових повідомленнях. *Науковий журнал «Information Technology: Computer*

Science, Software Engineering and Cybersecurity». 2024. № 4. С. 197–205. URL: https://doi.org/10.32782/IT/2024-4-23.

Собко О. В. виконано огляд предметної області виявлення та класифікації кіберзалякувань, наведено опис розроблених методів оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості, виявлення і класифікації кіберзалякувань, а також інтерпретації результатів виявлення кіберзалякувань; описано взаємодію розроблених методів для вирішення задачі виявлення та класифікації кіберзалякувань; проведено аналіз отриманих результатів, підготовлено чернетку рукопису.

2. Собко О. В., Бармак О. В. Метод аналізу та формування репрезентативних вибірок текстових даних із використанням моделей машинного навчання. *Науковий журнал «Computer Science and Applied Mathematics».* 2024. № 2. С. 83–92. URL: https://doi.org/10.26661/2786-6254-2024-2-09

Собко О. В. виконано огляд предметної області, описано розроблений метод оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості, виконано апробацію методу на розробленому програмному забезпеченні, описано отримані результати, підготовлено чернетки рукопису.

3. Собко О. В. Метод класифікації кіберзалякувань в україномовному текстовому контенті засобами штучного інтелекту. *Наука і техніка сьогодні.* 2024. № 13 (41). С. 1252–1263. URL: https://doi.org/10.52058/2786-6025-2024-13(41)-1252-1263

Собко О. В. виконано огляд існуючих рішень щодо класифікації кіберзалякувань в україномовному текстовому контенті, наведено розроблений метод виявлення і класифікації кіберзалякувань в україномовному текстовому контенті, розроблено програмне забезпечення та виконано апробацію методу, описано отримані результати, виконано підготовку чернетки рукопису.

4. Собко О. В. Метод інтерпретації результатів виявлення кіберзалякувань у текстовому контенті засобами штучного інтелекту. *Науковий журнал «Вісник Хмельницького національного університету», серія: Технічні науки».* 2024. № 6, Т. 1 (343). С. 302–309. URL: https://doi.org/10.31891/2307-5732-2024-343-6-45

Собко О. В. виконано відомих підходів та засобів для інтерпретації результатів нейромереж для задачі виявлення кіберзалякувань, висвітлено розроблений метод інтерпретації результатів виявлення кіберзалякувань, виконано апробацію методу на розробленому програмному забезпеченні, описано отримані результати, виконано підготовку чернетки рукопису.

5. Method for Analysis and Formation of Representative Text Datasets / O. Sobko, O. Mazurets, M. Molchanova, I. Krak, O. Barmak. *CEUR Workshop Proceedings*, 2025, vol. 3899, pp. 84–98. URL: https://ceur-ws.org/Vol-3899/paper9.pdf *(індексована в наукометричній базі Scopus).*

Собко О. В. висвітлено огляд відомих методів та рішень щодо оцінювання та коригування репрезентативності датасетів, описано і апробовано розроблений метод оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості.

6. Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network / I. Krak, O. Zalutska, M. Molchanova, O. Mazurets, R. Bahrii, O. Sobko, O. Barmak. *CEUR Workshop Proceedings*, 2024, vol. 3688, pp. 16–28. URL: https://ceur-ws.org/Vol-3688/paper2.pdf *(індексована в наукометричній базі Scopus).*

Собко О. В. висвітлено огляд відомих методів та рішень щодо автоматизованого виявлення образливих висловлювань у текстових даних; розроблено методологію виявлення образливих висловлювань для української мови.

7. Method for Neural Network Cyberbullying Detection in Text Content With Visual Analytic / I. Krak, O. Sobko, M. Molchanova, I. Tymofiiev, O. Mazurets, O. Barmak. *CEUR Workshop Proceedings*, 2025, vol. 3917, pp. 298–309. URL: https://ceur-ws.org/Vol-3917/paper57.pdf *(індексована в наукометричній базі Scopus).*

Собко О. В. описано і апробовано розроблений метод виявлення та класифікації кіберзалякувань у текстовому контенті, з подальшою візуальною інтерпретацією результатів; описано розроблене експериментальне програмного забезпечення для апробації рішень щодо виявлення, класифікації та візуальної інтерпретації кіберзалякувань у текстовому контенті.

8. Text Data Vectorization Model of Ukrainian-Language Internet Communication Content / V. Slobodzian, O. Kovalchuk, M. Molchanova, O. Sobko, O. Mazurets, O. Barmak, I. Krak. *CEUR Workshop Proceedings*, 2022, vol. 3171, pp. 561–571. URL: https://ceur-ws.org/Vol-3171/paper45.pdf *(індексована в наукометричних базах Scopus, Web of Science).*

Собко О. В. виконано огляд існуючих підходів до вирішення задачі, аналіз та визначення доцільних методів аналізу текстового контенту для моделі української мови в завданнях для аналізу обробки природної мови з Інтернет-джерел.

9. Visual Analytics-Based Method for Sentiment Analysis of COVID-19 Ukrainian Tweets / O. Kovalchuk, V. Slobodzian, O. Sobko, M. Molchanova, O. Mazurets, O. Barmak, I. Krak, N. Savina. *Lecture Notes on Data Engineering and Communications Technologies*, 2023, vol. 149, pp. 591–607. URL: https://doi.org/10.1007/978-3-031-16203-9_33 *(індексована в наукометричних базах Scopus, Web of Science).*

Собко О. В. виконано аналіз існуючих класифікаторів для класифікації текстів за сентиментом та обґрунтування доцільності використання конкретних моделей машинного навчання для задач обробки природної мови.

10. А. с. № 132920 Україна. Комп'ютерна програма «Інтелектуальна інформаційна система для оцінювання та коригування репрезентативності текстових датасетів» / О.В. Собко. 2025.

Собко О. В. розроблено модулі та компоненти комп'ютерної програми для апробації розробленого методу оцінювання та коригування репрезентативності датасету за FATE-принципом справедливості.

11. А. с. № 132921 Україна. Комп'ютерна програма «Інтелектуальна інформаційна система для виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту» / О.В. Собко. 2025.

Собко О. В. розроблено модулі та компоненти комп'ютерної програми для апробації розробленого методу виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту.

ВВАЖАТИ, що дисертаційна робота Собко О.В.
(прізвище, ініціали здобувача)
«Методи виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту»,
(назва)
яка подана на здобуття ступеня доктора філософії, за своїм науковим рівнем та практичною цінністю, змістом та оформленням повністю відповідає вимогам пп. 6, 7, 8, 9 «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради Закладу вищої освіти, наукової установи про присудження ступеня доктора філософії», затвердженому постановою Кабінету Міністрів України від 12 січня 2022 р. № 44, та відповідає напрямку наукового дослідження освітньо-наукової програми Хмельницького національного університету зі спеціальності 122 Комп'ютерні науки.
(шифр, назва)

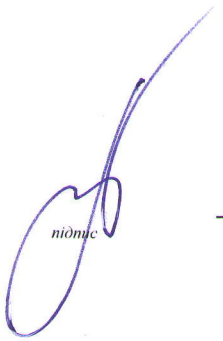
РЕКОМЕНДУВАТИ:

Дисертаційну роботу «Методи виявлення та класифікації кіберзалякувань у текстовому контенті засобами штучного інтелекту»,
назва роботи
подану Собко Оленою Віталіївною
прізвище, ім'я, по батькові
на здобуття ступеня доктора філософії, до захисту.

Головуючий публічної презентації:

доктор технічних наук,
(науковий ступінь)

професор, декан факультету інформаційних технологій ХНУ
(вчене звання, посада)


підпис

Тетяна ГОВОРУЩЕНКО
Ім'я ПРІЗВИЩЕ