

ЛІП'ЯНИНА-ГОНЧАРЕНКО ХРИСТИНА

Західноукраїнський національний університет
<https://orcid.org/0000-0002-2441-6292>

КОМАР МИРОСЛАВ

Західноукраїнський національний університет
<https://orcid.org/0000-0001-6541-0359>

ЮРКІВ ХРИСТИНА

Західноукраїнський національний університет
<https://orcid.org/0009-0007-4917-3251>

ЛУК'ЯНЧУК ВАСИЛЬ

Західноукраїнський національний університет
<https://orcid.org/0009-0009-8829-0316>

КОНЦЕПТУАЛЬНА МОДЕЛЬ ІНТЕЛЕКТУАЛЬНОЇ ОЦІНКИ НАСЛІДКІВ ТЕХНОГЕННИХ КАТАСТРОФ

У статті досліджується тема техногенної безпеки в умовах сучасної війни та важливість розробки концептуальної моделі інтелектуальної оцінки наслідків техногенних катастроф. Завданням дослідження було зібрання даних, їх аналіз, розробка моделі для прогнозування наслідків, визначення параметрів впливу та розробка рекомендацій.

Отримані результати підтверджують важливість розробки такої моделі для оперативного визначення рівня небезпеки в разі техногенних катастроф. Модель базується на зборі даних, аналізі, моделюванні та оцінці наслідків, що дозволяє враховувати різноманітні аспекти катастроф та їх вплив на навколишнє середовище та людей.

Результати оцінки моделі показали її ефективність у класифікації та прогнозуванні рівня небезпеки з точністю вище 77%. У майбутньому планується розширити модель, враховуючи додаткові фактори, такі як погодні умови та датчикові показники, і застосувати її у різних сферах діяльності, де важлива оперативна оцінка рівня небезпеки.

Ключові слова: техногенна безпека, техногенні катастрофи, інтелектуальна оцінка, моделювання наслідків, оперативна реакція, ризик, дані, аналіз, попередження.

KHYRSTYNA LIPIANINA-HONCHARENKO

West Ukrainian National University

MYROSLAV KOMAR

West Ukrainian National University

KHYRSTYNA YURKIV

West Ukrainian National University

VASYL LUKIANCHUK

West Ukrainian National University

CONCEPTUAL MODEL OF INTELLIGENT ASSESSMENT OF CONSEQUENCES OF TECHNOGENIC CATASTROPHES

In the context of modern warfare, the relevance of technogenic safety becomes particularly crucial due to the increasing threat of technogenic disasters. This article is dedicated to studying the most current challenges related to technogenic safety during wartime and determining the necessity for the development of novel approaches for prevention and management of the consequences of such catastrophes. The research object encompasses a complex of issues associated with mitigating the aftermath of technogenic disasters, while the main goal is the development of a model for intelligent assessment of these consequences.

The aim of the work lies in devising a conceptual model for the intelligent assessment of the aftermath of technogenic catastrophes. To achieve this goal, the following tasks were undertaken: data collection regarding technogenic disasters, analysis of this data, development of a model for predicting consequences, determination of influencing parameters, and formulation of recommendations.

The obtained results underscore the importance of creating such a model for swift and effective determination of the danger level in case of technogenic catastrophes. The conceptual model is built on the foundation of data collection, analysis, consequences modeling, and assessment, which allows for the consideration of diverse aspects of catastrophes and their impact on the environment and human beings.

Evaluation results of the model demonstrated its effectiveness in classification and forecasting of danger levels. The achieved accuracy surpasses 77%, thus confirming the success of the approach. In the future, the model is intended to be expanded by considering additional factors, such as weather conditions and sensor readings, and applied across various fields of activity where operational assessment of danger levels is crucial.

In conclusion, the article is dedicated to the development of a conceptual model for intelligent assessment of the consequences of technogenic catastrophes, which utilizes data, analysis, and modeling to swiftly and accurately determine danger levels, thereby enhancing safety in emergency situations.

Keywords: technogenic safety, technogenic disasters, intelligent assessment, consequences modeling, operational response, risk, data, analysis, prevention.

Постановка проблеми у загальному вигляді

та її зв'язок із важливими науковими чи практичними завданнями

Сучасні виклики щодо техногенної безпеки, особливо в період війни, вимагають новітніх способів попередження та забезпечення безперебійного функціонування об'єктів підвищеної небезпеки. Такі об'єкти

зазвичай складають фундамент сучасного техногенного суспільства, котре в разі виникнення надзвичайної ситуації може опинитись у вкрай важкому та критичному стані.

Проблема ліквідації наслідків надзвичайних ситуацій як природного так і техногенного характеру є важливою державною функцією, а її актуальність обумовлена як кількістю природних катастроф так і розповсюдженістю різного спектра підприємств важкої, хімічної та атомно-енергетичної промисловості. Навіть за умов нормальної роботи на промислових підприємствах можливі незначні викиди небезпечних речовин у навколишнє середовище, що в момент аварії можуть набути масштабів техногенної катастрофи. Це становить значну загрозу для населення, території та навколишнього середовища.

Дослідження є актуальним, оскільки техногенні катастрофи можуть мати серйозні наслідки для людей, навколишнього середовища та економіки. Швидке та ефективне визначення рівня небезпеки може допомогти уникнути або зменшити наслідки таких катастроф. Розроблення концептуальної моделі інтелектуальної оцінки наслідків техногенних катастроф, що використовує інформацію зібрану у свідків, може значно покращити процес визначення потенційної небезпеки, зменшити часові витрати та допомогти особам, які приймають рішення, бути більш гнучкими та адаптованими до конкретної ситуації. Крім того, можливість застосування моделі в інших сферах діяльності, де необхідно оперативним чином визначати рівень небезпеки, робить дослідження ще більш актуальним.

Тому важливим є дослідження техногенних катастроф, а розробка концептуальної моделі інтелектуальної оцінки наслідків техногенних катастроф є актуальною задачею.

Аналіз останніх досліджень чи публікацій

Існують різні методи визначення техногенних катастроф, включаючи візуальне спостереження, сенсори, аналіз даних зі сторонніх джерел тощо. Однак більшість з них обмежені в точності та швидкості реакції. Оцінка катастроф на основі соціальних мереж - новий напрям в дослідженнях ризиків. Вона дозволяє швидко виявляти небезпеки, аналізуючи повідомлення користувачів. Використання цих даних може покращити системи управління катастрофами, але є ризики недостатньої точності та розповсюдження недостовірної інформації.

Соціальні мережі стали важливим джерелом інформації для управління надзвичайними ситуаціями [1]. Дослідження [2] пропонує огляд методів керування та аналізу даних для покращення систем управління катастрофами. У [3] розглянуто зв'язок між аналізом даних з соцмереж та реагуванням на катастрофи. У [4] розроблено систему для мобілізації ресурсів на основі твітів. В [5] використано мультиагентну систему для аналізу даних з мікроблогів у реальному часі. У [6] порівняно алгоритми аналізу даних для класифікації твітів. В [7] розроблено модель для розподілу ресурсів після катастрофи. У [8] використано нейронний підхід для ідентифікації твітів під час катастрофи.

EAIMS [9] надає інструмент для аналізу кризових ситуацій зі соцмереж. CiviQue [10] спеціалізується на твітах з геотегами. [11] пропонує глибоку структуру виявлення змін. У [12] розроблено модель для визначення часу реагування служб. У [13] ґрунтується на парадигмі "людина як датчик". У [14] поєднує аналіз тексту і місцезнаходження для виявлення постраждалих. У [15] використовується гібридна структура для виявлення подій.

Незважаючи на наявність різних інструментів для аналізу техногенних катастроф та забезпечення автоматизованого реагування [14, 15], прийняття ефективних рішень все ще залишається проблемою, тому потрібно розробити концептуальну модель інтелектуальної оцінки наслідків техногенних катастроф для покращення подальшого оперативного реагування на них та дозволить запобігти можливим гуманітарно-економічним катастрофам.

Формулювання цілей статті

Мета роботи полягає в розробці концептуальної моделі інтелектуальної оцінки наслідків техногенних катастроф. Для досягнення цієї мети були виконані такі завдання: збір даних про техногенні катастрофи, аналіз цих даних, розробка моделі для прогнозування наслідків, визначення параметрів впливу та розробка рекомендацій.

Виклад основного матеріалу дослідження

Концептуальна модель оцінки наслідків техногенних катастроф передбачає використання комп'ютерних технологій та аналізу даних для забезпечення більш точної та оперативної оцінки наслідків катастроф.

Концептуальна модель оцінки наслідків техногенних катастроф представлена наступними блоками та схематично (рис 1):

Збір даних: збирання різноманітних даних про техногенну катастрофу, такі як місце, час, причина, тип техніки, що викликала катастрофу, та інші фактори, які можуть вплинути на наслідки катастрофи.

Аналіз даних: аналіз зібраних даних, щоб визначити масштаб техногенної катастрофи та наслідки, які вона може мати на довкілля, людей та майно. Для цього можуть використовуватись різні методи аналізу даних, включаючи статистичний аналіз, машинне навчання та інші методи.

1. Моделювання наслідків: розроблення моделі, яка допоможе прогнозувати наслідки техногенної катастрофи, використовуючи зібрані дані та аналіз. Ця модель може включати різні складові, такі як модель поширення забруднення в повітрі, воді та ґрунті, модель оцінки впливу на здоров'я людей та тварин, модель оцінки екологічних наслідків та інші.

2. Оцінка наслідків: використання моделі для оцінки наслідків техногенної катастрофи на довкілля, людей та майно. Оцінка може включати різні показники, такі як ризик для здоров'я людей, вартість збитків, екологічну шкоду та інші.

3. Визначення параметрів: далі визначаються параметри, які впливають на наслідки катастрофи. Це можуть бути такі параметри, як тип виробництва, кількість людей та обладнання на місці події, метеорологічні умови, відстань до населених пунктів.

4. Розроблення рекомендацій: розроблення рекомендацій для зменшення наслідків техногенної катастрофи в майбутньому, таких як рекомендації щодо поліпшення безпеки на робочому місці, встановлення систем попередження про можливі аварії та швидкий реагування на виявлення проблем, аналіз попередніх аварій та проведення відповідних заходів для попередження подібних ситуацій в майбутньому та інше.

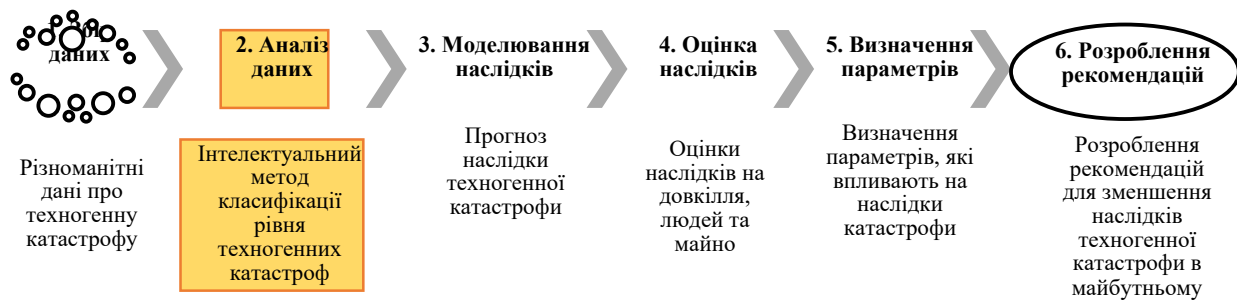


Рис. 1. Концептуальна модель інтелектуальної оцінки наслідків техногенних катастроф

Для проведення тестування розробленої концептуальної моделі інтелектуальної оцінки наслідків техногенних катастроф обрано мову Python. У якості вхідних даних використано записи про аварії на 12 різних підприємствах у 3 різних країнах, кожен рядок у яких містить інформацію про аварію [16] (табл. 1). Кожен рядок у записах містить інформацію про аварію, включаючи дату та час аварії, місце виникнення, тип промисловості, причини та наслідки аварії, відомості про кількість потерпілих, рівень катастрофи та потенційно можливий рівень катастрофи.

Далі дані про аварії були оброблені та підготовлені підходом Exploratory Data Analysis [17, 18] для використання в розробленому методі. Для цього були виконані наступні дії:

- Зчитування даних з файлу та перетворення у формат, придатний для подальшої обробки.
- Очищення та нормалізація даних, включаючи видалення непотрібної інформації та стандартизацію формату записів.
- Створення бази даних для зберігання оброблених даних та побудова відповідного інтерфейсу для доступу до цих даних.

Після попередньої обробки даних, які включали в себе зведення та очищення даних, а також інженерію ознак, було побудовано модель машинного навчання, яка використовувала ансамблеві методи класифікації. Для оцінки точності класифікації було використано крос-валідацію, що дозволило отримати точність моделі на методах. Після навчання та перевірки моделі на валідаційній вибірці, вона була застосована для оцінки наслідків техногенних катастроф на інших підприємствах та в інших регіонах.

Таблиця 1

Опис даних			
№	Параметр	Опис	Non-Null, Count, Dtype
1	Data	Часова позначка або інформація про час/дату	411, ненульовий, об'єкт
2	Countries	В якій країні сталася аварія	411, ненульовий, об'єкт
3	Local	Місто, в якому розташоване підприємство-виробник	411, ненульовий, об'єкт
4	Industry sector	До якого сектора належить об'єкт	411, ненульовий, об'єкт
5	Accident level	З I (серйозна) по VI (дуже серйозна), він показує, наскільки серйозною була аварія	411, ненульовий, об'єкт
6	Potential Accident Level	Залежно від рівня аварії, база даних також реєструє, наскільки серйозною могла бути аварія	411, ненульовий, об'єкт
7	Genre	Особа, яка повідомила про подію чоловічої статі жіночої статі	411, ненульовий, об'єкт
8	Employee or Third Party	Ким є очевидець працівником або третьою особою	411, ненульовий, об'єкт
9	Critical Risk	Деякий опис ризику, пов'язаного з аварією	411, ненульовий, об'єкт
10	Description	Детальний опис того, як сталася аварія	411, ненульовий, об'єкт

На першому етапі потрібно перевести дані в індексні значення. Далі виділяємо дані на тестові та навчальну вибірку (рис. 2).

```
Train set has 382 records out of 425 which is 90%
Dev set has 382 records out of 425 which is 90%
Test set has 43 records out of 425 which is 10%
```

Рис. 2. Результати поділу даних

Проведемо класифікацію Potential Accident Level на основі ансамблевих методів: AdaBoost, Gradient Boost, XGBoost та CatBoost. Проведемо до кожного типу моделювання звіт про класифікацію [19], що показує представлення основних показників класифікації для кожного класу. Це дає глибше уявлення про поведінку класифікатора над глобальною точністю, яка може маскувати функціональні недоліки в одному класі багатокласової проблеми. Візуальні звіти про класифікацію використовуються для порівняння моделей класифікації, щоб вибрати моделі, які є «темнішими», мають сильніші показники класифікації або більш збалансовані.

Аналізуючи оцінку моделювання AdaBoost, бачимо, що клас Industry Sector (клас №0) має найвищу точність класифікації (75%) та передбачення (82%), підтвержену оцінкою f1-score на рівні 0,78 (рис. 3а, 3б). Водночас, середня точність класифікації Potential Accident Level складає 63% для 382 значень (рис. 3а).

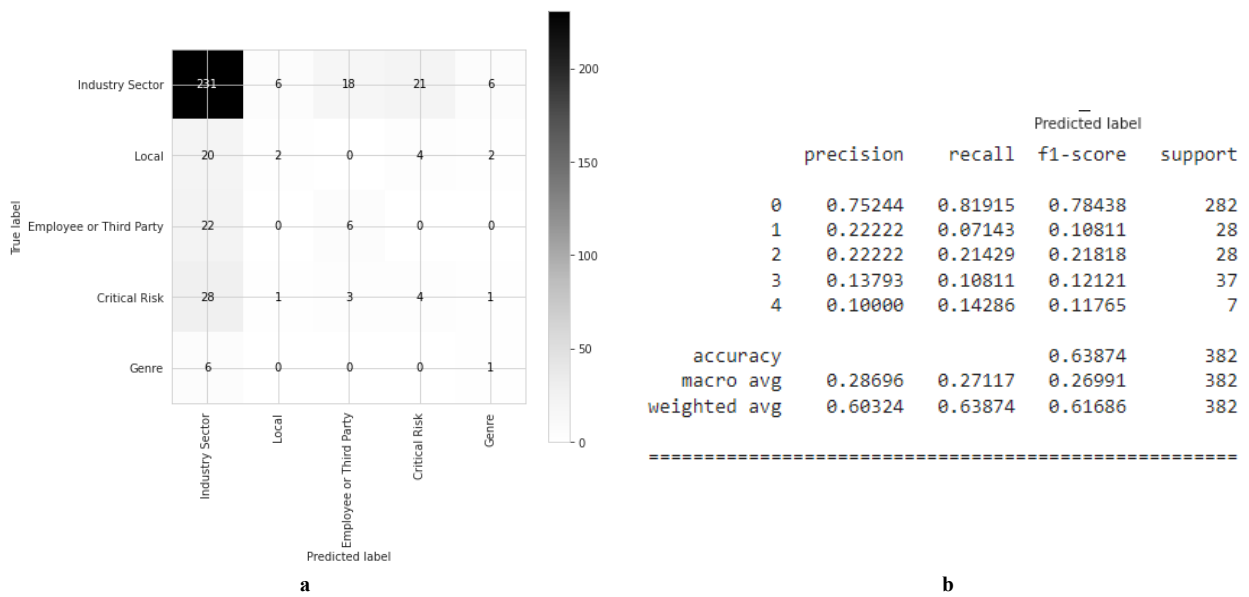


Рис. 3. Оцінка моделювання AdaBoost

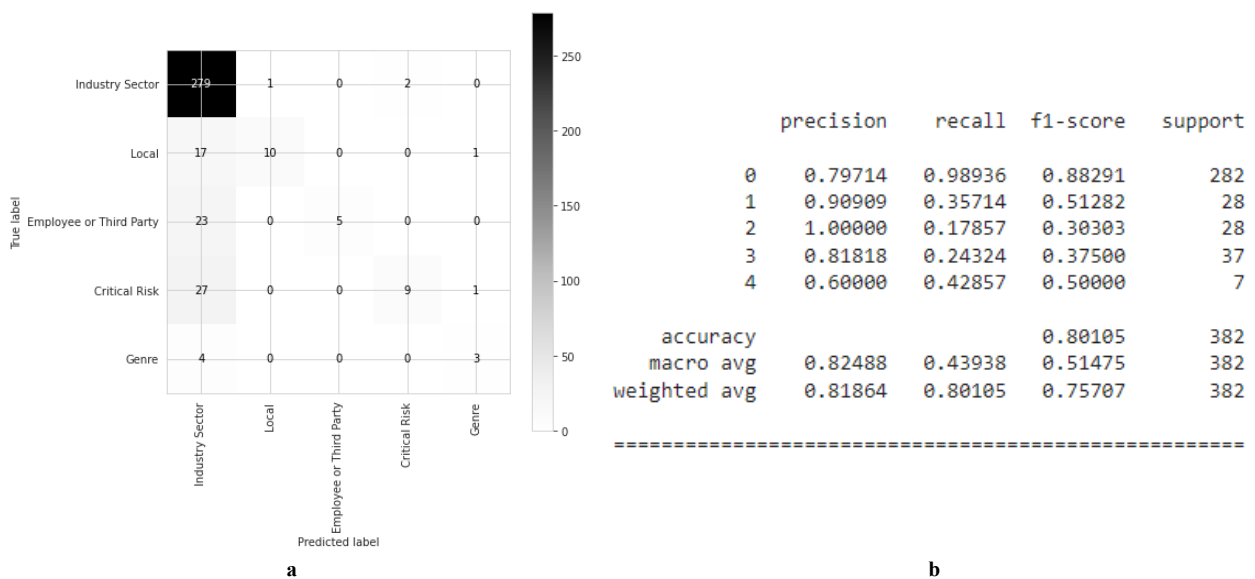


Рис. 4. Оцінка моделювання Gradient Boost

Оцінюючи моделювання Gradient Boost (рис. 4), варто зазначити, що клас Employee or Third Party (клас №2) має 100% точність класифікації, але лише 17% передбачення правильно, підтверджено це оцінкою

f1-score на рівні 0,30 для 28 значень. Відзначимо також високу точність у класифікації Industry Sector (клас №0), з 80% розпізнаних значень та 99% правильних передбачень, що відображено оцінкою f1-score 0,88 для 282 значень. Середній показник точності прогнозових значень Potential Accident Level склав 80% для вибірки розміром 382 значення.

Метод XGBoost (рис. 5) найточніше класифікує клас Employee or Third Party (клас №2) з 100% точністю, але з правильним передбаченням всього 4%. Оцінка f1-score для цього класу становить 0,07 для 282 значень. Також відзначимо високу точність у класифікації Industry Sector (клас №0), де розпізнано 75% значень з 99% правильних передбачень. Оцінка f1-score для цього класу дорівнює 0,85 для 282 значень. Середня точність прогнозових значень класифікації Potential Accident Level становить 75% для вибірки розміром 382 значення.

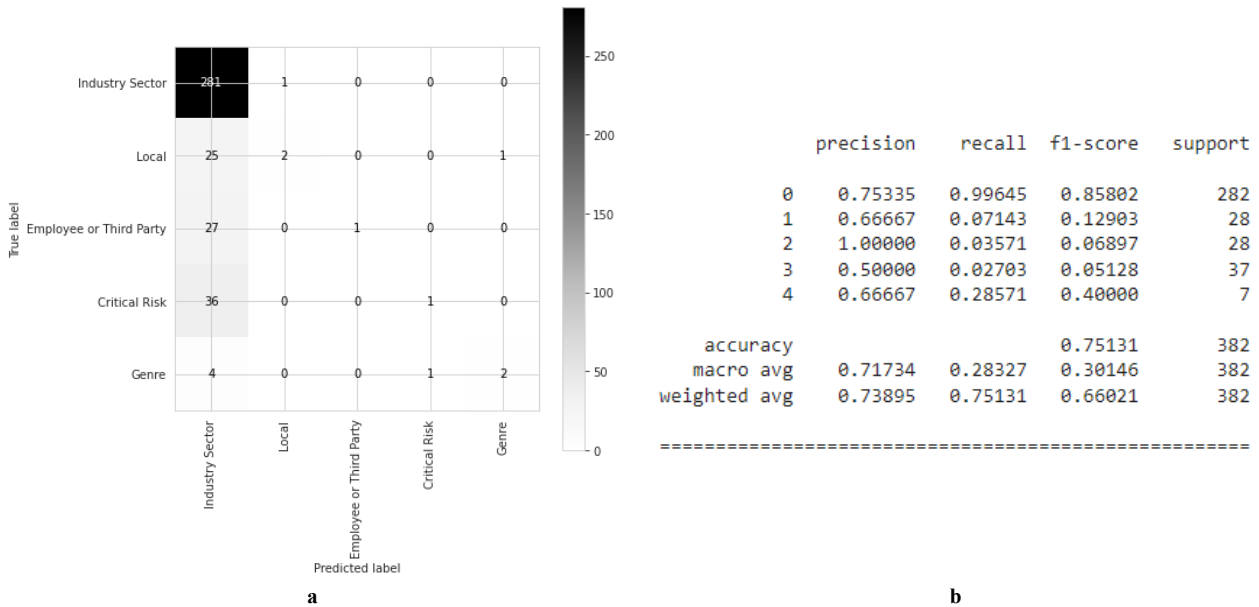


Рис. 5. Оцінка моделювання XGBoost

Метод LigthGBM (рис.6) найточніше класифікує класи Local (клас №1), Employee or Third Party (клас №2) та Genre (клас №4) з 100% точністю, але з відповідними правильними передбаченнями всього 18%, 3% та 14%. Оцінки f1-score підтверджують це. Варто також відзначити високу точність у класифікації Industry Sector (клас №0), де розпізнано 76% значень з 99% правильних передбачень. Оцінка f1-score для цього класу дорівнює 0,86 для 282 значень. Середня точність прогнозових значень класифікації Potential Accident Level становить 76% для вибірки розміром 382 значення.

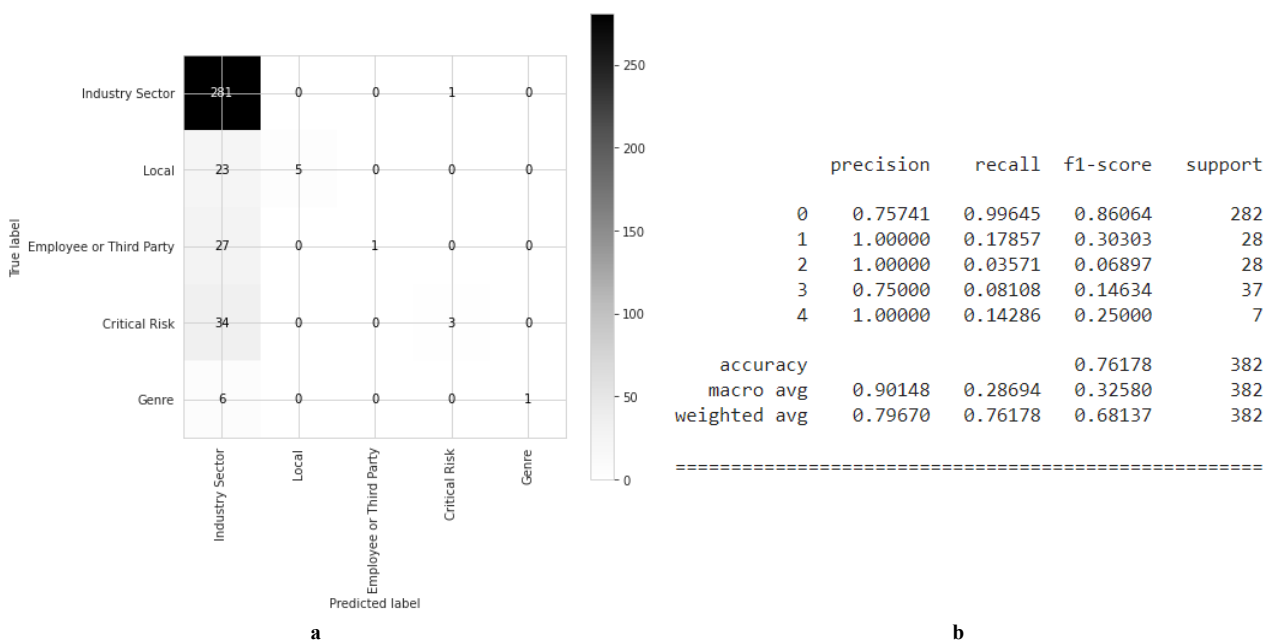


Рис. 6. Оцінка моделювання LigthGBM

Метод CatBoost (рис. 7) показав найкращі результати моделювання. Найбільш точно він класифікує

клас Employee or Third Party (клас №2) з 100% точністю, з правильними передбаченнями на рівні 25%. Оцінка f1-score для цього класу становить 0,40 для 28 значень. Також важливо зазначити високу точність класифікації Industry Sector (клас №0), де розпізнано 80% значень з 99% правильними передбаченнями. Оцінка f1-score для цього класу дорівнює 0,89 для 282 значень. В середньому точність прогнозних значень класифікації Potential Accident Level становить 81% для вибірки розміром 382 значення, що є дуже добрим результатом.

Після проведення моделювання найкращим методом для класифікації та визначення Potential Accident Level виявився CatBoost. Моделювання на тестовій вибірці за цим методом (рис.8) показало, що при класифікації Potential Accident Level відносно Industry Sector (клас №0) розпізнано 80% значень, з 97% правильними передбаченнями. Оцінка f1-score для цього класу склала 0,88 для 34 значень, що є високою точністю. В середньому точність прогнозних значень класифікації Potential Accident Level становить 77% для вибірки розміром 43 значення, що також є добрим результатом.

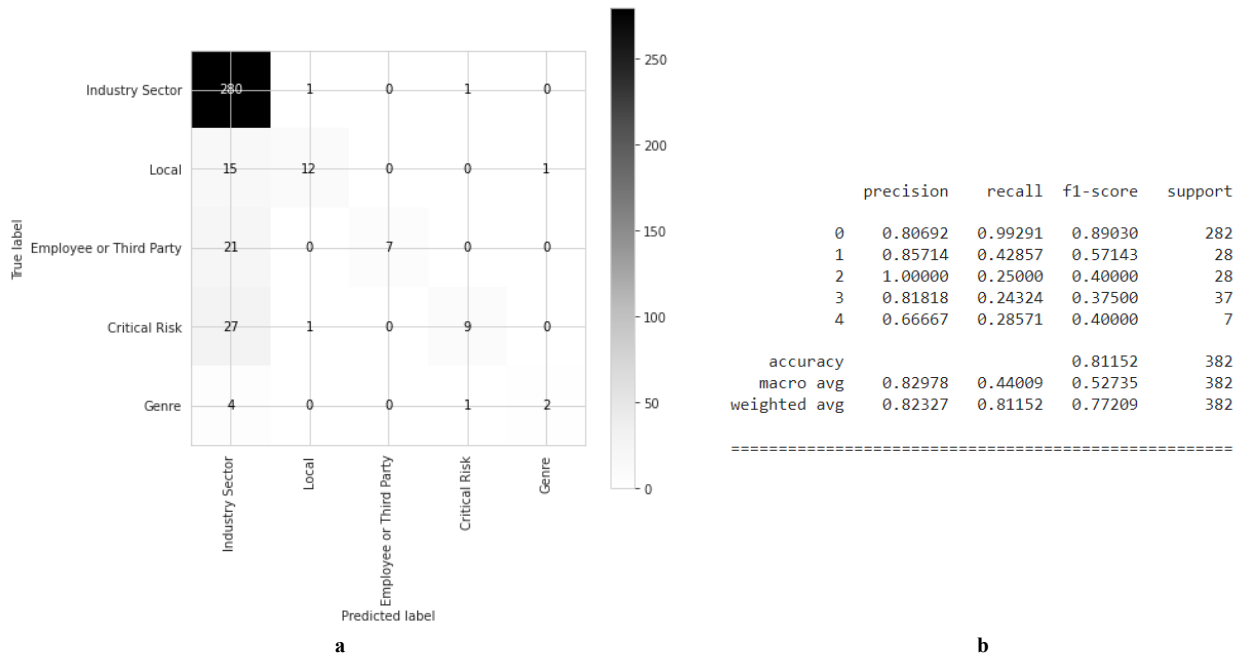


Рис. 7. Оцінка моделювання CatBoost

	precision	recall	f1-score	support
0	0.805	0.971	0.880	34
1	0.000	0.000	0.000	2
2	0.000	0.000	0.000	3
3	0.000	0.000	0.000	3
4	0.000	0.000	0.000	1
accuracy			0.767	43
macro avg	0.161	0.194	0.176	43
weighted avg	0.636	0.767	0.696	43

CatBoost Test Accuracy - 76.74418604651163

Рис. 8. Оцінка моделювання CatBoost на тестовій вибірці

На рисунку 9 видно, що система передбачила рівень потенційної небезпеки майже вірно (окрім 4 рядка) для перших 5 рядків тестової вибірки.

y_test	test_preds[:5]
256 0	array([[0],
56 0	[0],
72 0	[0],
389 4	[3],
131 0	[0]])

а. тестова вибірка

б. результат передбачення

Рис. 9. Результат передбачення рівня потенційної небезпеки

тестування показали перспективні результати в точному прогнозуванні рівня потенційних аварій. Метод CatBoost виявився найефективнішим для цього завдання, досягнувши високої точності в класифікації рівня аварій.

Результати тестування на невідомому наборі даних показали, що модель змогла правильно класифікувати більшість тестових точок даних. Матриця плутанини показала, що модель здатна точно класифікувати рівень небезпеки V і I в аваріях з точністю 0,97 і 0,94 відповідно. Загальна середня точність результатів склала 0,68, що вказує на можливості для подальшого вдосконалення.

Порівняно з існуючими методами, які спираються на дані соціальних мереж, ця інтелектуальна модель є надійним і точним способом класифікації рівня техногенних катастроф. Можливість реалізації цього методу в мобільному додатку дозволить швидко реагувати на аварії та запобігати можливим катастрофам.

Майбутні дослідження можуть спрямовуватися на підвищення точності моделі та розробку комплексного підходу, який враховуватиме інші фактори, впливаючи на рівень небезпеки. Вивчення потенціалу додаткових джерел даних, таких як погодні умови або датчики в реальному часі, також може поліпшити точність моделі.

Розширення досліджень на інші галузі промисловості дозволить перевірити ефективність методу для класифікації рівня наслідків техногенних катастроф в інших контекстах. Також можна використовувати підхід, розроблений у цьому дослідженні, для виявлення ризиків для безпеки та здоров'я працівників.

Загалом запропонований інтелектуальний метод має значний потенціал для впровадження в управління техногенними катастрофами, допомагаючи передбачати рівень небезпеки з високою точністю.

Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

Для зменшення часових витрат на визначення потенційної небезпеки в регіоні авторами розроблено концептуальну модель інтелектуальної оцінки наслідків техногенних катастроф. Для реалізації інтелектуального методу оцінки промислової безпеки регіону у якості вхідних даних використано записи про аварії на 12 різних підприємствах у 3 різних країнах, кожен рядок у яких містить інформацію про аварію.

Оцінка моделювання для кількісних даних показали наступні результати. Класифікація Potential Accident Level відносно Industry Sector (клас №0), де було розпізнано 80% значень, з яких 97% правильно модель передбачила. Оцінка f1-score рівна 0,88 для 34 значень, що є достатньо високою точністю. В середньому точність отриманих прогнозних значень класифікації Potential Accident Level рівна 77% для вибірки в розмірі 43 значення, що є хорошим результатом.

У подальших дослідженнях, буде запропоновано підходи для підвищення точності моделі та розроблено більш комплексний підхід, який може враховувати інші фактори, які можуть вплинути на рівень небезпеки техногенних катастроф: погодні умови, показання датчиків та ін. Також, буде розглянуто можливість застосування методу в інших сферах діяльності, де необхідно оперативно визначати рівень небезпеки (екологія, транспорт).

References

1. Luna S., Pennock M. J. Social media applications and emergency management: A literature review and research agenda. *International Journal of Disaster Risk Reduction*. 2018. Vol. 28. P. 565—577.
2. Li T. Data-Driven Techniques in Disaster Information Management. *ACM Computing Surveys*. 2017. Vol. 50, no. 1. P. 1—45.
3. Nazer T. H. Intelligent Disaster Response via Social Media Analysis A Survey. *ACM SIGKDD Explorations Newsletter*. 2017. Vol. 19, no. 1. P. 46—59.
4. Bhoi A., Pujari S. P., Balabantaray R. C. A deep learning-based social media text analysis framework for disaster resource management. *Social Network Analysis and Mining*. 2020. Vol. 10, no. 1.
5. Şahin C., Rokne J., Alhaji R. Emergency Detection and Evacuation Planning Using Social Media. *Lecture Notes in Social Networks*. Cham, 2018. P. 149—164.
6. Huang Q., Xiao Y. Geographic Situational Awareness: Mining Tweets for Disaster Preparedness, Emergency Response, Impact, and Recovery. *ISPRS International Journal of Geo-Information*. 2015. Vol. 4, no. 3. P. 1549—1568.
7. Basu M., Bit S. D., Ghosh S. Utilizing microblogs for optimized real-time resource allocation in post-disaster scenarios. *Social Network Analysis and Mining*. 2021. Vol. 12, no. 1.
8. Madichetty S., M S. A Neural-Based Approach for Detecting the Situational Information From Twitter During Disaster. *IEEE Transactions on Computational Social Systems*. 2021. Vol. 8, no. 4. P. 870—880.
9. McCreadie R., Macdonald C., Ounis I. Eaims: emergency analysis identification and management system, in *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, Series SIGIR '16 (ACM, New York, NY, 2016)*, pp. 1101—1104.
10. Kanojia D., Kumar V. (2016) Ramamritham, Civique: using social media to detect urban emergencies. *CoRR abs/1610.04377*
11. Zheng Z. Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters. *Remote Sensing of Environment*. 2021.

Vol. 265. P. 112636.

12. Bandyopadhyay M., Singh V. Development of agent based model for predicting emergency response time. *Perspectives in Science*. 2016. Vol. 8. P. 138—141.

13. Avvenuti M. A framework for detecting unfolding emergencies using humans as sensors. *SpringerPlus*. 2016. Vol. 5, no. 1.

14. Karam E., Hussein W., Gharib T. F. Integrating location and textual information for detecting affected people in a crisis. *Social Network Analysis and Mining*. 2021. Vol. 11, no. 1.

15. Adel H. Improving Crisis Events Detection Using DistilBERT with Hunger Games Search Algorithm. *Mathematics*. 2022. Vol. 10, no. 3. P. 447.

16. Industrial Safety and Health Analytics Database. Kaggle: Your Machine Learning and Data Science Community. URL: <https://www.kaggle.com/datasets/ihmstefanini/industrial-safety-and-health-analytics-database>

17. Komorowski M. Exploratory Data Analysis. Secondary Analysis of Electronic Health Records. Cham, 2016. P. 185—203.

18. Li Vigni M., Durante C., Cocchi M. Exploratory Data Analysis. Data Handling in Science and Technology. 2013. P. 55—126.

19. Classification Report — Yellowbrick v1.5 documentation. Yellowbrick: Machine Learning Visualization — Yellowbrick v1.5 documentation. URL: https://www.scikit-yb.org/en/latest/api/classifier/classification_report.html