

ПРОГНОЗУВАННЯ РИЗИКУ ЗАМІНУВАННЯ ТЕРИТОРІЙ З ВИКОРИСТАННЯМ АНСАМБЛЮ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ

Національний університет «Львівська політехніка»

Мінна небезпека залишається однією з найсерйозніших гуманітарних проблем сучасної України. Внаслідок повномасштабної війни значні території заміновані. За даними міжнародних організацій, потенційно небезпечними вважаються близько чверті площі країни. Традиційні методи розмінування потребують значних ресурсів та часу, тому виникає потреба у створенні інтелектуальних систем прогнозування ризику замінування, здатних автоматично визначати зони підвищеної небезпеки та покращити процес планування розмінування. Запропоновано метод прогнозування ризику замінування територій із застосуванням машинного навчання та геоінформаційних технологій. Сформовано багатозарову вибірку, що охоплює геопросторові, соціально-економічні та військові ознаки. Просторова сітка охоплює понад 0,55 млн комірок розміром 500×500 м, що забезпечує високий рівень деталізації покриття замінованих територій країни. Для аналізу даних використано чотири базові алгоритми: Logistic Regression, Random Forest, XGBoost та Invariant Risk Minimization. На їх основі реалізовано ансамблеву архітектуру типу Stacked Generalization, у якій метамодель поєднує результати базових моделей у єдиний прогноз. Для оцінювання якості моделей застосовано метрики AUC, Precision, Recall та F1-score. За результатами експериментів найвищу точність продемонстрував ансамбль моделей, перевищивши показники найкращої окремої моделі XGBoost. Це підтверджує ефективність інтеграції різних типів моделей для підвищення точності прогнозування. Встановлено, що використання ансамблевих алгоритмів дозволяє автоматизовано та з високою точністю прогнозувати ризик замінування територій, підвищуючи ефективність планування розмінування та безпеку населення.

Ключові слова: логістична регресія, машинне навчання, ансамбль моделей, прогнозування.

Вступ

Проблема мінування територій є однією з найгостріших гуманітарних і безпекових загроз сучасної України. Внаслідок повномасштабної збройної агресії значні площі сільськогосподарських угідь, лісових масивів та населених пунктів залишаються забрудненими боєприпасами, мінами та іншими вибухонебезпечними залишками війни. За оцінками Міністерства оборони України та міжнародних організацій, понад чверть території країни потенційно потребує обстеження або розмінування, через це Україна є однією з найзамінованіших держав світу.

Процеси гуманітарного розмінування традиційно базуються на польових обстеженнях та візуальному пошуку вибухонебезпечних предметів, що вимагає значних фінансових і людських ресурсів. Перевірка території вручну займає багато часу, тоді як кількість небезпечних ділянок зростає щодня. Тому критично важливо розробляти системи прогнозування ризику мінування, які дають змогу визначати пріоритети для розмінування, оптимізувати маршрути пошуку та підвищити безпеку цивільного населення.

Сучасні дослідження показують, що використання методів машинного навчання (ML) та геоінформаційних систем (GIS) дозволяє прогнозувати ймовірність мінування територій на основі поєднання різномірних джерел даних: супутникових знімків, карт бойових дій, даних про щільність населення, інфраструктуру та соціально-економічні показники. Такий підхід дає змогу автоматизувати процес оцінювання ризику мінування та визначати пріоритетні ділянки для подальшого розмінування. Водночас для України постає завдання адаптації подібних моделей з урахуванням локальних особливостей ландшафту, інтенсивності бойових дій та структури землекористування.

Наукова новизна роботи полягає у розробці підходу до прогнозування ризику мінування територій на основі ансамблевої моделі типу Stacked Generalization, що поєднує алгоритми Logistic Regression, Random Forest, XGBoost і Invariant Risk Minimization, забезпечуючи підвищення точності прогнозування ризику мінування територій.

Метою дослідження є створення автоматизованої системи прогнозування ризику замінування територій України на основі методів машинного навчання. Для досягнення мети передбачено виконання таких завдань:

- зібрати та підготувати комплекс даних з геопросторових, соціально-економічних і військових джерел;
- розробити та порівняти моделі машинного навчання для прогнозування ризику замінування;
- розробити ансамбль моделей машинного навчання з оптимізованими ваговими коефіцієнтами з урахуванням переваг базових моделей;
- розробити користувацький інтерфейс, що дасть змогу порівнювати результати роботи алгоритмів;
- проаналізувати вплив ключових ознак і закономірностей розподілу ризику.

Практичне значення результатів полягає у можливості використання моделі як інструмента оцінювання ризику замінувань для центру протимінної діяльності, гуманітарних організацій та розробки системи для інформування населення. Отримані результати можуть бути інтегровані у геоінформаційну систему оцінювання мінної небезпеки, що підвищить ефективність планування розмінування та зменшить ризики для цивільного населення.

Для побудови моделей машинного навчання застосовано комплексний підхід, який поєднує геопросторові, соціально-економічні та історичні дані. Метою є створення моделі, здатної оцінювати ризик мінної небезпеки на рівні просторових комірок (grid cells) з використанням даних з відкритих джерел.

Для формування навчальної вибірки в межах дослідження використано дані з відкритих геопросторових, історичних та OSINT-джерел. Геопросторові дані включають інформацію з Open Street Map, що містить відомості про дорожню мережу, залізницю, річки, населені пункти та будівлі. Дані бойових дій та мінування отримано з системи Mine Action Information Management System for Ukraine (IMSMA-UA), яку адмініструє Український центр протимінної діяльності. До того ж, враховано офіційні звіти Державної служби України з надзвичайних ситуацій, де наведено осередки підтверджених випадків мінування та розмінування.

У дослідженні розглянуто три групи моделей: базові лінійні моделі, представлені логістичною регресією та Support Vector Machine з радіальним ядром [1]; дерева рішень та ансамблеві методи, до яких належать Random Forest, Gradient Boosting (XGBoost, LightGBM) [2]; а також нейронні мережі та сучасні інтерпретовані підходи, зокрема Multi-Layer Perceptron і модель на основі концепції Invariant Risk Minimization [3], що забезпечує стійкість до зсувів у даних.

Для порівняння результатів моделі навчено на вибірці з 70 % даних, а тестування проводилося на 30 % методом stratified train-test split. Для перевірки стабільності моделей використовувалася блокова крос-валідація з урахуванням географічної кореляції між суміжними комірками.

Оцінювання якості розроблених моделей здійснювалося за сукупністю показників, що відображають різні аспекти їх ефективності. Основними метриками виступали AUC-ROC — площа під кривою, яка характеризує здатність моделі розрізняти позитивні та негативні класи, та AUC-PR — площа під кривою Precision та Recall, що є інформативнішою в умовах суттєвого дисбалансу класів [4]. Для комплексної оцінки збалансованості між точністю (Precision) і повнотою (Recall) використовувалась метрика F1-score як їх гармонічне середнє значення.

Аналіз останніх досліджень і публікацій

Проблематика прогнозування ризику мінування територій після бойових дій привертає дедалі більшу увагу у зв'язку зі зростанням масштабів мінної небезпеки. Застосування методів машинного навчання та геоінформаційних технологій дає можливість створювати автоматизовані системи оцінювання ризику на основі історичних даних, супутникових зображень та соціально-географічних показників.

Суттєвий внесок у формування підходів до моделювання ризику мінування зроблено у дослідженні [5], де запропоновано модель прогнозування небезпечних ділянок на прикладі Камбоджі з використанням алгоритмів логістичної регресії та методу опорних векторів SVM. Автори показали, що навіть класичні алгоритми машинного навчання, навчені на даних попередніх операцій

розмінування, можуть результативно ідентифікувати території з підвищеною ймовірністю замінування. Важливим внеском роботи є обґрунтування підходів до опрацювання незбалансованих вибірок, що є типовим викликом для задач гуманітарного розмінування. Модель продемонструвала високі показники точності (0,89) та AUC (0,93), завдяки чому це дослідження є важливим кроком у розвитку методів автоматизованої оцінки ризику мінування територій.

Подальший розвиток цього напрямку отримано у роботі [6], де описано систему Desk-AId (Desk Assessment with Geospatial AI) для оцінювання ризику замінування на основі геопросторових даних. Ця система застосовує методи штучного інтелекту для дистанційного аналізу територій, що дає змогу виконувати попередню оцінку ризику без необхідності виїзду фахівців у потенційно небезпечні зони. Desk-AId інтегрує відкриті супутникові дані (Sentinel-2, Copernicus DEM, OSM) та алгоритми глибинного навчання для побудови карт мінування територій. Важливим результатом роботи є формування глобальної структури ознак, релевантних для задач оцінювання ризику, яка може бути адаптована до різних країн і регіонів.

Інноваційний підхід до підвищення стійкості моделей представлено в системі RELand (Risk Estimation of Landmines) [7], яка поєднує принципи Invariant Risk Minimization та глибинного навчання для узагальнення даних у нових географічних умовах з різними розподілами ознак. На відміну від традиційних методів, RELand враховує структуру «легких» і «складних» середовищ, що зменшує ризик перенавчання та підвищує стійкість до просторових зсувів. Експериментальні результати на даних Колумбії підтверджують перевагу цього підходу через покращення ROC і PR-кривих. У цьому ж напрямі розвиваються інтерпретовані нейронні моделі зокрема, застосування Sparse Feature Masking для побудови карт важливості ознак підвищує прозорість моделі та відповідає сучасним тенденціям створення пояснюваних (Explainable AI) систем для критично важливих задач безпеки.

У низці робіт увага зосереджується на застосуванні методів машинного навчання для виявлення мін та оцінки небезпеки територій. Зокрема, у роботі [8] проаналізовано можливості алгоритмів для прогнозування мінної небезпеки на підставі геопросторових даних та технічних характеристик мінних полів. У дослідженні [9] запропоновано підхід на основі глибоких згорткових мереж для виявлення мін за оптичними зображеннями, що демонструє ефективність використання комп'ютерного зору в задачах гуманітарного розмінування. У роботі [10] аналізуються сигнали георадара (GPR) із застосуванням згорткових нейронних мереж, що дало змогу підвищити точність і стабільність автоматичного розпізнавання мін. Ці дослідження загалом підтверджують, що інтеграція методів машинного навчання та глибинних моделей підвищує можливості виявлення мін у різних середовищах.

Методи Out-of-Distribution (OOD) привертають особливу увагу завдяки здатності забезпечувати стабільність моделей під час їх застосування в нових географічних регіонах з іншими розподілами ознак. Зокрема, у роботі [11] сформовано теоретичні засади IRM, які стали основою для подальших практичних реалізацій, зокрема системи RELand. Застосування OOD підходів є перспективним напрямом для створення універсальних моделей прогнозування мінування територій на основі геопросторових даних.

У роботі [12] описано інтелектуальну систему підтримки прийняття рішень для задач гуманітарного розмінування, орієнтовану на комплексну оцінку ризику територій. Запропонований підхід поєднує геоінформаційний аналіз і багатофакторне оцінювання з метою формування рекомендацій щодо пріоритетності обстеження ділянок. На відміну від методів, що ґрунтуються виключно на алгоритмах машинного навчання, система TIRAMISU робить акцент на інтеграції різноманітних джерел інформації, що є важливим для оперативного планування робіт з розмінування. Водночас відсутність автоматизованого прогнозування на основі великих обсягів історичних даних обмежує масштабованість і адаптивність цього підходу.

Окремі дослідження, зокрема [13], розглядають можливості використання радарних і гіперспектральних даних для безпосереднього виявлення мін, проте такі підходи залишаються дорогавартісними й технічно складними для масштабного використання. На противагу цьому, поєднання супутникових знімків, відкритих картографічних ресурсів та методів машинного навчання є значно практичнішим для оперативного оцінювання ризиків у реальних умовах, де важливими є швидкість, масштабованість і невисока вартість збору даних.

Таким чином, аналіз сучасних наукових досліджень свідчить, що наявні підходи до прогнозування ризику мінування територій переважно зосереджуються або на застосуванні окремих алгоритмів машинного навчання в межах локальних регіонів, або на геоінформаційних системах підтримки прийняття рішень без повноцінного автоматизованого прогнозування на основі великих масивів даних. Новітні рішення, зокрема Desk-AId та RELand, демонструють ефективність

використання геопросторових ознак і методів узагальнення в умовах просторових зсувів, проте не розглядають інтеграцію моделей у межах ансамблевої архітектури. Це зумовлює актуальність розроблення комплексного підходу, що поєднує методи машинного навчання з метою підвищення ефективності прогнозування ризику мінування територій. Подальші дослідження доцільно спрямувати на створення регіонально адаптованих моделей, які враховують локальні особливості бойових дій, ландшафт і характер мінування, а також на їх інтеграцію в системи інформування населення про небезпечні та потенційно заміновані території.

Результати дослідження

У ході дослідження проведено попередню оцінку ризику замінування територій України, що слугувало основою для подальшої побудови моделі машинного навчання. Станом на початок 2025 року територія України, ймовірно уражена мінами чи іншими вибухонебезпечними залишками, оцінюється приблизно у 138 тис. км² землі та додатково 14 тис. км² водної поверхні. Сукупно цей показник складає близько чверті від загальної площі країни. Для формування вибірки заміновану територію України розбито на сітку з комітками розміром 500 на 500 метрів. За такого розміру клітинки загальна кількість осередків у вибірці становила близько 0,55 млн елементів, кожен з яких характеризує локальну ділянку земної поверхні з відповідними геопросторовими атрибутами.

Для навчання моделей сформовано комплексний набір просторових даних, що охоплює офіційні джерела державних і громадських організацій, а також відкриті аналітичні платформи. Основну частину інформації отримано з інформаційної системи управління протимінною діяльністю України, карт ДСНС та аналітичного ресурсу DeepState [14]. Дані з цих джерел інтегровано у єдину базу, адаптовану для машинного аналізу та подальшого прогнозування ризику мінування.

Система IMSMA-UA є офіційним реєстром, що ведеться центром протимінної діяльності відповідно до міжнародних стандартів IMAS. З неї отримано дані підтверджених, підозрюваних і очищених ділянок мінування, а також їхні статуси. На основі цих відомостей сформовано цільову змінну, де клас 1 відповідає осередкам із підтвердженим або ймовірним мінуванням, а клас 0 — ділянкам без ознак небезпеки. Дані з карти мінної небезпеки Державної служби України з надзвичайних ситуацій використано для уточнення меж зон потенційної небезпеки [15]. Кожен об'єкт з карти містить категорію: «Небезпечна територія (підозра на мінування)», «Територія у процесі розмінування»; «Очищена територія». Аналітичний ресурс DeepState містить регулярно оновлювані дані про лінію бойових дій, які використано для визначення відстані до актуальної лінії фронту, одного з ключових факторів ризику. Для кожного осередку обчислено мінімальну відстань до лінії фронту (`dist_to_front`, км), середню інтенсивність змін фронту за останні 30 днів (`front_activity_index`) та частоту зміни статусу території (окупована чи звільнена), що відображає динаміку бойових дій.

Для формування вибірки використовувались групи ознак, що відображають різні аспекти мінної небезпеки територій. Геопросторові ознаки включали тип землекористування, наявність населених пунктів і об'єктів критичної інфраструктури, щільність забудови, а також відстань до основних транспортних шляхів і водних об'єктів. Військово-історичні ознаки охоплювали відстань до лінії фронту, індекс активності бойових дій, частоту зміни контролю території, статус ділянки (окупована / деокупована) та історію зафіксованих бойових зіткнень. Загалом у фінальній моделі використовувалось понад 20 агрегованих ознак, нормалізованих і приведених до єдиного просторового масштабу комірок.

Враховуючи специфіку задачі прогнозування мінної небезпеки, навчальна вибірка характеризується суттєвим дисбалансом класів, оскільки частка осередків з підтвердженим або ймовірним мінуванням не перевищує 19 %, тоді як більшість комірок належить до безпечного класу. Для коректної оцінки якості моделей використовувались метрики, чутливі до дисбалансу класів, зокрема AUC-PR та F1-score. До того ж, під час навчання моделей застосовувалось *cost-sensitive* навчання шляхом зважування класів (*class weights*), що дозволяє зменшити вплив домінування негативного класу та підвищити здатність моделей виявляти небезпечні ділянки.

Для побудови ансамблевої моделі застосовано підхід *Stacked Generalization* з дворівневою структурою навчання. З метою запобігання витоку інформації та коректної оцінки узагальнювальної здатності навчання метамоделі здійснювалося на основі *out-of-fold* прогнозів базових алгоритмів, отриманих у процесі крос-валідації. Базові моделі формували прогнози для відповідних валідаційних підмножин, які надалі використовувалися як вхідні ознаки другого рівня ансамблю.

Такий підхід забезпечує інтеграцію взаємодоповнювальних властивостей різних алгоритмів без

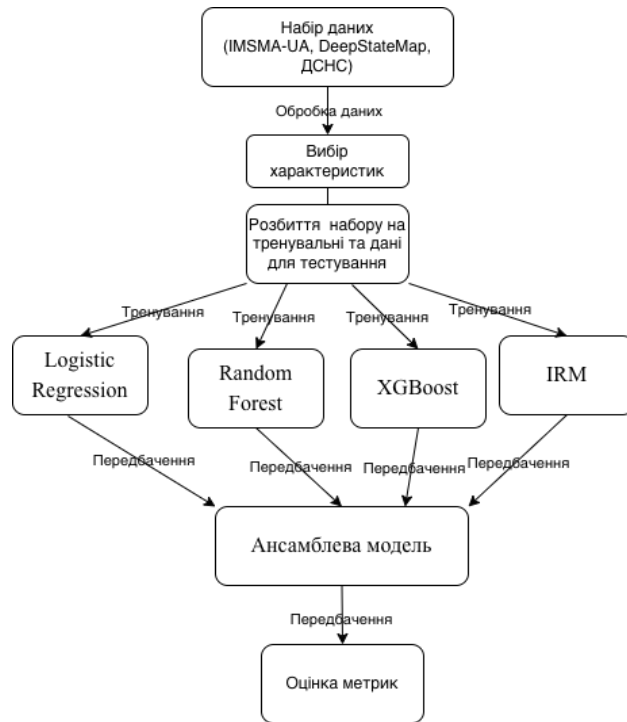


Рис. 1. Схема запропонованої ансамблевої моделі

залучення інформації з тестової вибірки та сприяє підвищенню стабільності результатів.

Архітектуру процесу побудови ансамблевої моделі показано на рис. 1. Вона відображає основні етапи, від збору та обробки даних до формування підсумкового прогнозу з використанням алгоритмів машинного навчання.

На рис. 1 показано загальну послідовність етапів прогнозування ризику мінної небезпеки. Спочатку здійснюється інтеграція просторових даних з офіційних та відкритих джерел, після чого відбувається їх попередня обробка, вибір релевантних ознак та розподіл на тренувальний і тестовий набори. Далі кожна з базових моделей навчається незалежно, формуючи власні проміжні передбачення. Ці результати узгоджуються метамоделлю в межах ансамблевого підходу, який забезпечує інтеграцію виходів базових алгоритмів у єдиний фінальний прогноз. Завершальний етап передбачає оцінювання ефективності моделі за допомогою метрик класифікації, результати якого подано в таблиці.

Порівняння результатів моделей машинного навчання на тестовій вибірці

Модель	AUC-ROC	AUC-PR	Precision	Recall	F1-Score
Logistic Regression	0,871	0,412	0,712	0,862	0,779
Random Forest	0,927	0,615	0,795	0,861	0,827
XGBoost	0,948	0,684	0,883	0,907	0,895
IRM	0,941	0,671	0,753	0,811	0,781

Для практичної реалізації запропонованої моделі прогнозування ризику міннування розроблено мобільний застосунок. На рис. 2 продемонстровано приклад візуалізації результатів прогнозування у мобільному застосунку, де червоні круги позначають осередки з підвищеною ймовірністю міннування. Застосунок слугує клієнтським інтерфейсом для моделі, що працює на сервері. Архітектура такого підходу передбачає відправлення запиту до серверної частини для отримання координат прогнозованих осередків та відповідної ймовірності міннування для відображення результату у вигляді кольорового маркера на карті. Такий підхід дозволяє поєднати наукову модель прогнозування з реальним прикладним інструментом для інформування населення про небезпечні території.

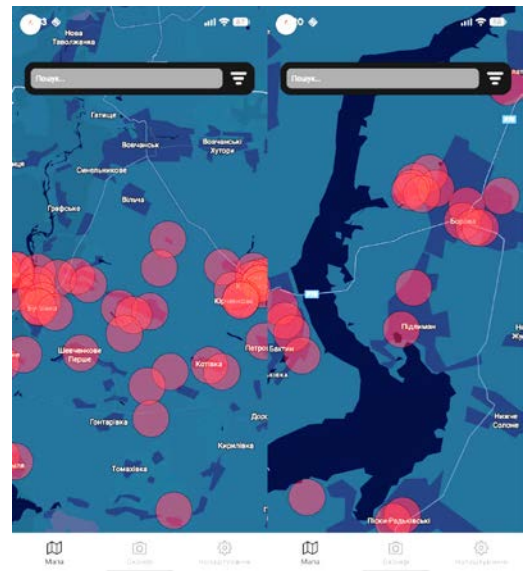


Рис. 2. Інтерфейс мобільного застосунку для візуалізації прогнозованих зон мінної небезпеки

Графік на рис. 3 демонструє порівняння ROC-кривих для всіх моделей. Найкращі результати показала модель XGBoost з AUC-ROC 0.948, що свідчить про високу здатність відокремлювати ризикові осередки (класи 1) від безпечних (класи 0). Модель IRM має схожий результат, проте характеризується кращою узагальненістю на нових регіонах, що підтверджується стабільністю показників у крос-валідації $\sigma < 0,01$. Logistic Regression продемонструвала показник Recall (0,862), що означає мінімізацію пропущених небезпечних ділянок, проте за рахунок більшої кількості хибнопозитивних передбачень низький Precision (0,712).

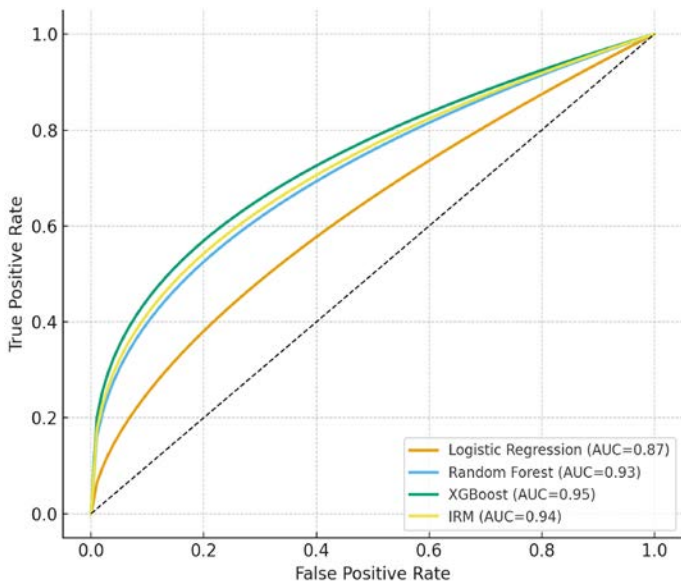


Рис. 3. ROC-криві моделей машинного навчання

соровано підхід Stacked Generalization (Stacking Ensemble), що передбачає використання дворівневої структури моделювання. На першому рівні формуються незалежні передбачення за допомогою різних типів базових алгоритмів: Logistic Regression, Random Forest, XGBoost та Invariant Risk Minimization. На другому рівні використовується метамодель, яка навчається на вихідних даних базових моделей і комбінує їх результати для отримання кінцевого прогнозу. В межах цього дослідження як метамодель застосовано Gradient Boosting Regressor, що приймає на вхід вектор із чотирьох прогнозів і формує інтегральну оцінку ризику мінущання:

$$\hat{y}_{ensemble} = \sigma(w_1 P_{LR} + w_2 P_{RF} + w_3 P_{XGB} + w_4 P_{IRM}),$$

де w — вагові коефіцієнти, отримані під час навчання метамоделі, σ — сигмоїдна функція, що обмежує вихід у діапазоні $[0,1]$.

Для оцінки ефективності ансамблю використовувались ті самі метрики, що і для базових моделей (рис. 4). Всі моделі навчались на ідентичних вибірках, а ансамбль тренувався за допомогою п'ятиразової крос-валідації для уникнення перенавчання.

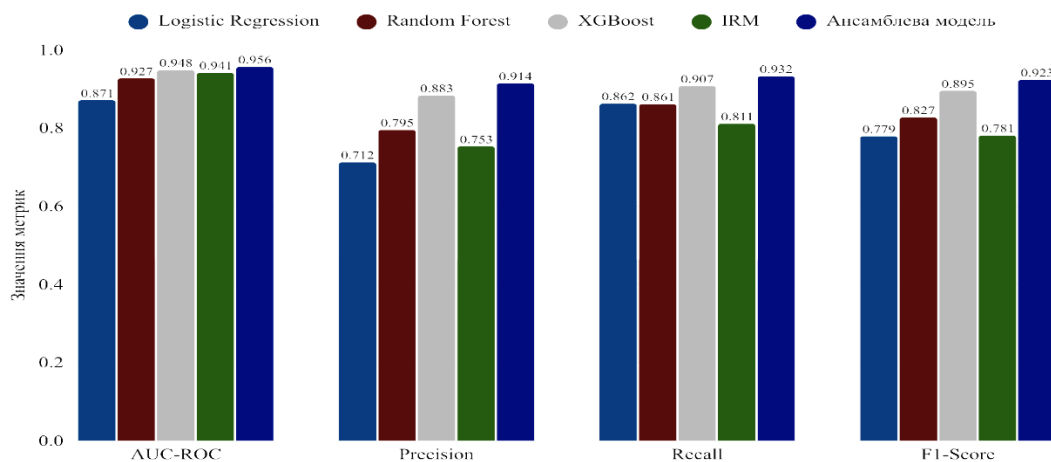


Рис. 4. Результати ансамблювання моделей машинного навчання

Отримані результати показують, що запропонований ансамбль моделей демонструє перевагу над окремими базовими алгоритмами за всіма ключовими метриками ефективності. Збільшення F1-score на 2,8% порівняно з найкращою базовою моделлю XGBoost підтверджує, що метамодель ефективно інтегрує дані, які окремі моделі не здатні повністю узагальнити. Підвищення AUC-PR з 0,684 до 0,712 є особливо важливим для задач з незбалансованими даними, де позитивний клас (наявність мін) становить незначну частку вибірки. Високе значення Recall — 0,932 свідчить про

Моделі Random Forest та XGBoost забезпечують краще співвідношення Precision/Recall. Зокрема, XGBoost досяг найвищого F1-score (0,895), що свідчить про оптимальний баланс між виявленням небезпечних зон і мінімізацією помилок.

Модель XGBoost виявилась найефективнішою для прогнозування ризику мінущання територій за сукупністю метрик точності, AUC-ROC 0.948 та F1-score 0.895. IRM-модель показала найкращу стабільність результатів у разі перенесення на нові регіони, тому її доцільно використовувати для інтеграції у геоінформаційні системи. Random Forest є компромісним варіантом між точністю та швидкістю обчислень. Logistic Regression поступається за точністю, але може бути корисною як базовий метод швидкої оцінки ризику.

Для побудови ансамблю моделей засто-

здатність ансамблю виявляти більшість небезпечних ділянок, а Precision — 0,914 вказує на низьку частку хибнопозитивних передбачень. Порівняльний аналіз ROC-кривих демонструє, що ансамбль має найбільшу площу під кривою, що підтверджує якість розмежування позитивного та негативного класів. Додатково, аналіз Precision-Recall кривих показує стабільну перевагу ансамблю у діапазоні порогових значень 0,3...0,7, що забезпечує оптимальний компроміс між точністю та повнотою прогнозів.

Для оцінювання впливу окремих ознак на прогноз ризику мінування проведено аналіз важливості ознак на основі ансамблевої моделі та алгоритмів градієнтного бустингу. Результати показали, що найбільший внесок у формування прогнозу мають такі фактори: відстань до лінії фронту, тип землекористування, індекс інтенсивності бойових дій та історія зміни контролю над територією. Отримані висновки підтверджують релевантність вибраних ознак для задач прогнозування мінної небезпеки.

Висновки

Розроблено підхід до прогнозування ризику мінування територій з використанням методів машинного навчання. Запропонована модель поєднує результати декількох алгоритмів у рамках ансамблевої архітектури типу Stacked Generalization, що забезпечує вищу точність у порівнянні з базовими моделями. Сформовано багатокомпонентну вибірку даних з відкритих джерел, що охоплюють географічні, соціально-економічні та історичні параметри, які впливають на ризик мінування територій. Проведено порівняльний аналіз моделей машинного навчання: Logistic Regression, Random Forest, XGBoost, Invariant Risk Minimization. Реалізовано ансамблеву модель, яка об'єднує базові алгоритми, отриманий ансамбль показав покращення F1-score на 2,8 % та AUC на 4,1 % порівняно з найкращою базовою моделлю XGBoost, що свідчить про здатність метамоделі об'єднувати різні аспекти даних та зменшувати похибки окремих моделей. На основі аналізу важливості ознак визначено ключові фактори ризику замінування територій. Отримані результати підтверджують ефективність поєднання ML і GIS-підходів у задачах гуманітарного розмінування, що відкриває перспективи для створення інтелектуальних систем підтримки рішень під час планування розмінування та управління ризиками.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189-215, 2020. <https://doi.org/10.1016/j.neucom.2019.10.118>.
- [2] P. Zhang, Y. Jia, and Y. Shang, "Research and application of XGBoost in imbalanced data," *International Journal of Distributed Sensor Networks*, vol. 18, no. 6, 2022. <https://doi.org/10.1177/15501329221106935>.
- [3] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz, "Invariant Risk Minimization," *arXiv preprint arXiv:1907.02893*, 2019. <https://doi.org/10.48550/arXiv.1907.02893>.
- [4] E. Richardson, et al., "The receiver operating characteristic curve accurately assesses imbalanced datasets," *Patterns*, vol. 5, no. 6, 2024. <https://doi.org/10.1016/j.patter.2024.100994>.
- [5] P. Kosmas, W. Rafique, J. Barras, S. P. Joglekar, and D. Zheng, "Predictive Analysis of Landmine Risk," *IEEE Access*, no. 7, pp.107259-107269, 2019. <https://doi.org/10.1109/ACCESS.2019.2929677>.
- [6] R. Cirillo, et al., "Desk-AID: Humanitarian Aid Desk Assessment with Geospatial AI for Predicting Landmine Areas," *arXiv preprint arXiv:2405.09444*, 2024. <https://doi.org/10.48550/arXiv.2405.09444>.
- [7] M. D. Rubio, S. Zeng, Q. Wang, H. Heidari, and F. Fang, "RELand: Risk Estimation of Landmines via Interpretable Invariant Risk Minimization," *ACM Journal on Computing and Sustainable Societies*, vol. 2, no. 2, pp. 1-29, 2024. <https://doi.org/10.1145/3648437>.
- [8] L. Sun, "Landmine Classification and Prediction Using Machine Learning Techniques," in *Proc. IEEE ICIPCA*, 2025, pp. 407-414. <https://doi.org/10.1109/ICIPCA65645.2025.11138774>.
- [9] E. Vivoli, M. Bertini, and L. Capineri, "Deep learning-based real-time detection of surface landmines using optical imaging," *Remote Sensing*, vol. 16, no. 4, p. 677, 2024. <https://doi.org/10.3390/rs16040677>.
- [10] S. Lameri, F. Lombardi, P. Bestagini, M. Lualdi, and S. Tubaro, "Landmine detection from GPR data using convolutional neural networks," in *Proc. 25th European Signal Processing Conf.*, 2017, pp. 508-512. <https://doi.org/10.23919/EUSIPCO.2017.8081259>.
- [11] Y. Lin, H. Dong, H. Wang, and T. Zhang, "Bayesian Invariant Risk Minimization" in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 16000-16009. <https://doi.org/10.1109/CVPR52688.2022.01555>.
- [12] A. Krtalić, and M. Bajić, "Development of the TIRAMISU Advanced Intelligence Decision Support System," *European Journal of Remote Sensing*, vol. 52, no. 1, pp. 40-55, 2019. <https://doi.org/10.1080/22797254.2018.1550351>.
- [13] O. A. Pryshchenko, et al., "Implementation of an artificial intelligence approach to GPR systems for landmine detection," *Remote Sensing*, vol. 14, no. 17, p. 4421, 2022. <https://doi.org/10.3390/rs14174421>.
- [14] DeepStateMap, Frontline dynamics and military activity 2025. [Electronic resource]. Available: <https://deepstatemap.live>.
- [15] State Emergency Service of Ukraine (DSNS), "Interactive map of areas that may be contaminated with explosive objects," 2025. [Electronic resource]. Available: <https://mine.dsns.gov.ua/>.

Рекомендована кафедрою системного аналізу та інформаційних технологій ВНТУ

Дата надходження: 18.11.2025
Дата прийняття до публікації: 27.03.2026
Дата публікації: 8.04.2026

Братасюк Денис Іванович — аспірант кафедри програмного забезпечення, e-mail: denys.i.bratasiuk@lpnu.ua ;
Федасюк Дмитро Васильович — д-р техн. наук, професор, завідувач кафедри програмного забезпечення,
e-mail: dmytro.v.fedasyuk@lpnu.ua . <https://orcid.org/0000-0003-3552-7454>.
Національний університет «Львівська політехніка», Львів

D. I. Bratasiuk¹
D. V. Fedasyuk¹

Prediction of Landmine Contamination Risk Using a Machine Learning Model Ensemble

Lviv Polytechnic National University

Landmine contamination remains one of the most serious humanitarian challenges in modern Ukraine. As a result of the full-scale war, vast areas have been mined, and according to international organizations, nearly one quarter of the country's territory is considered potentially dangerous. Traditional demining methods require substantial time and resources, creating the need for intelligent systems capable of predicting landmine risk, automatically identifying high-hazard areas, and improving demining planning. This study proposes a method for predicting landmine contamination risk using machine learning and geospatial technologies. A multilayer dataset was constructed, incorporating geospatial, socio-economic, and military features. The spatial grid covers more than 0.55 million cells sized 500x500 m, ensuring a high level of detail in mapping contaminated areas. Four baseline algorithms: Logistic Regression, Random Forest, XGBoost, and Invariant Risk Minimization were employed for data analysis. Based on these models, an ensemble architecture using the Stacked Generalization approach was implemented, where a meta-model integrates the outputs of the base learners into a unified prediction. Model performance was assessed using AUC, Precision, Recall, and F1-score metrics. Experimental results show that the ensemble outperformed the best individual model (XGBoost), confirming the effectiveness of combining heterogeneous algorithms to improve prediction accuracy. The study demonstrates that ensemble machine learning methods enable accurate and automated prediction of landmine contamination risk, enhancing demining planning and contributing to improved civilian safety.

Keywords: logistic regression, machine learning, model ensemble, prediction.

Bratasiuk Denys I. — Post-Graduate Student, of the Chair of Software Engineering, e-mail: denys.i.bratasiuk@lpnu.ua ;
Fedasyuk Dmytro V. — Dr. Sc. (Eng.), Professor, Head of the Chair of Software Engineering, e-mail:
dmytro.v.fedasyuk@lpnu.ua . <https://orcid.org/0000-0003-3552-7454>