

УДК 004.8, 004.056, 004.75, 005.21

DOI <https://doi.org/10.35546/kntu2078-4481.2026.1.42>

І. П. СІНЦІН

доктор технічних наук, професор,
член-кореспондент Національної академії наук України, директор
Інститут програмних систем Національної академії наук України
ORCID: 0000-0002-4120-0784

О. П. ІЛЬІНА

кандидат фізико-математичних наук,
провідний науковий співробітник
Інститут програмних систем Національної академії наук України
ORCID: 0000-0002-4073-9649

С. Я. СКИБИК

аспірант
Інститут програмних систем Національної академії наук України
ORCID: 0009-0008-4336-680X

О. А. ЯЦЕНКО

старший науковий співробітник
Інститут програмних систем Національної академії наук України
ORCID: 0000-0002-4700-6704

МОДЕЛЮВАННЯ ТА ОЦІНЮВАННЯ ВЛАСТИВОСТЕЙ АДАПТИВНОГО АНСАМБЛЕВОГО ІНДИКАТОРУ ДІАГНОСТИКИ СТРАТЕГІЧНОЇ РЕСУРСНОЇ БЕЗПЕКИ РОЗПОДІЛЕНОЇ ОРГАНІЗАЦІЙНОЇ СИСТЕМИ

У статті представлено результати дослідження властивостей адаптивного ансамблевого індикатора діагностики стратегічної ресурсної безпеки (ДСРБ) розподіленої організаційної системи, призначеного для підтримки прийняття рішень в умовах високої невизначеності, неповноти інформації та дисбалансу спостережень. Індикатор реалізовано на основі гетерогенного ансамблю бінарних класифікаторів, що поєднує моделі різних парадигм машинного навчання (Naive Bayes, Support Vector Machine, Random Forest, k-Nearest Neighbors, Logistic Regression), та доповнено спеціалізованими механізмами масштабування, калібрування ймовірностей і адаптивного зважування результатів. Актуальність дослідження зумовлена специфікою ресурсних даних у системах стратегічного управління, для яких характерні складні та асиметричні розподіли ознак, часові лаги, висока неоднорідність умов спостереження, домінування благополучних станів та критичність ризику гіподіагностики неблагополучних ситуацій. З урахуванням цих особливостей у роботі запропоновано композитну метрику якості KQ , що інтегрує показники $F1neg$, коефіцієнт Метьюза та каппу Коена з урахуванням їхньої взаємної корельованості та управлінських пріоритетів. Дослідження виконано у формі серії обчислювальних експериментів із використанням спеціально згенерованих синтетичних наборів даних, які імітують доменно обумовлену статистичну поведінку ресурсних спостережень, зокрема наявність важких «хвостів» розподілів, шуму міток, класового дисбалансу та впливу ресурсної передісторії. Для цього розроблено уніфікований алгоритм супроводження обчислювального експерименту, що забезпечує відтворюваність, порівнянність результатів і статистичну валідацію ефектів. У межах трьох експериментів проведено базову верифікацію алгоритму, поетапний аналіз внеску окремих механізмів у показники якості та оцінювання ефективності ансамблю в стресових умовах порівняно з найкращим окремим класифікатором. Отримані результати свідчать про стійку перевагу ансамблевого індикатора ДСРБ за композитною метрикою якості, показниками специфічності та $F1neg$, що підтверджено статистично значущими парними тестами та оцінками розміру ефекту. Показано, що запропонований підхід забезпечує більш надійне виявлення потенційно небезпечних ресурсних станів та формує інтерпретовану основу для підтримки стратегічних управлінських рішень у розподілених організаційних системах.

Ключові слова: класифікація методами машинного навчання, гетерогенний ансамбль класифікаторів, адаптивна метрика якості, дисбаланс класів, специфічність (ризик гіподіагностики), м'яке голосування, підтримка стратегічних рішень.



I. P. SINITSYN

Doctor of Technical Sciences, Professor,
Corresponding Member of the National Academy of Sciences of Ukraine, Director
Institute of Software Systems
of the National Academy of Sciences of Ukraine
ORCID: 0000-0002-4120-0784

O. P. ILINA

PhD (Phys.-Math.), Senior Researcher
Institute of Software Systems
of the National Academy of Sciences of Ukraine
ORCID: 0000-0002-4073-9649

S. Ya. SKYBYK

Postgraduate Student
Institute of Software Systems
of the National Academy of Sciences of Ukraine
ORCID: 0009-0008-4336-680X

O. A. YATSENKO

Senior Researcher
Institute of Software Systems
of the National Academy of Sciences of Ukraine
ORCID: 0000-0002-4700-6704

MODELING AND EVALUATION OF THE PROPERTIES OF AN ADAPTIVE ENSEMBLE INDICATOR FOR DIAGNOSING THE STRATEGIC RESOURCE SECURITY OF A DISTRIBUTED ORGANIZATIONAL SYSTEM

The paper presents the results of a study of the properties of an adaptive ensemble indicator for diagnosing the strategic resource security (DSRS) of a distributed organizational system, intended to support decision-making under conditions of high uncertainty, incomplete information, and class imbalance. The indicator is implemented as a heterogeneous ensemble of binary classifiers combining machine learning models of different paradigms (Naive Bayes, Support Vector Machine, Random Forest, k-Nearest Neighbors, Logistic Regression) and is supplemented with specialized mechanisms for feature scaling, probability calibration, and adaptive weighting of model outputs. The relevance of the study is determined by the specific nature of resource data in strategic management systems, which are characterized by complex and asymmetric feature distributions, temporal lags, high heterogeneity of observation conditions, dominance of satisfactory states, and the critical importance of minimizing the risk of hypodiagnosis of unfavorable situations. Taking these characteristics into account, the paper proposes a composite quality metric KQ that integrates the $F1_{neg}$ score, the Matthews correlation coefficient, and Cohen's kappa, while considering their mutual correlations and domain-specific managerial priorities. The research is conducted through a series of computational experiments using specially generated synthetic datasets that reproduce domain-driven statistical properties of resource observations, including heavy-tailed distributions, label noise, severe class imbalance, and the influence of historical resource dynamics. For this purpose, a unified framework for managing computational experiments is developed, ensuring reproducibility, comparability of results, and statistical validation of observed effects. Within three experimental studies, the algorithm is subjected to baseline verification, a stepwise analysis of the contribution of individual mechanisms to classification quality, and an evaluation of ensemble performance under stress conditions in comparison with the best-performing individual classifier. The results demonstrate a stable advantage of the proposed ensemble DSRS indicator in terms of the composite quality metric, specificity, and $F1_{neg}$ score, which is confirmed by statistically significant paired tests and effect size estimates. It is shown that the proposed approach provides more reliable detection of potentially dangerous resource states and forms an interpretable basis for supporting strategic decision-making in distributed organizational systems.

Key words: machine learning-based classification, heterogeneous ensemble of classifiers, adaptive quality metric, class imbalance, specificity (risk of hypodiagnosis), soft voting, strategic decision support.

Постановка проблеми

Призначенням індикатору діагностики стратегічної ресурсної безпеки (ДСРБ) є виявлення таких поточних станів ресурсної ситуації в розподіленій організаційній системі (РОС), які становлять потенційну небезпеку для її стратегічних цілей.

У попередній роботі [1] було розглянуто доцільний спосіб формалізації такого індикатору, оснований на застосуванні великого обсягу ресурсних спостережень за елементами організаційної системи та експертними оцінками

стану стратегічних цілей. Пропонована формалізація використовує машинне навчання ансамблю бінарних класифікаторів. Він є максимально диверсифікованим за парадигмами репрезентованих моделей та адаптивно налаштовуваним на динамічну актуалізацію знань про статистичну поведінку ресурсних даних. Такий підхід орієнтований на використання індикатора за умов глибокої невизначеності стратегічних управлінських процесів [2].

На основі аналізу структурних та процесних аспектів предметної області, попередньо здійсненого в [3], були виявлені статистичні особливості даних та пріоритети їх аналізу, що визначають актуальні вимоги до властивостей індикатора.

Метою даного дослідження була організація, проведення та аналіз результатів обчислювальних експериментів, спрямованих на дослідження актуальних властивостей індикатора ДСРБ в умовах можливих форм статистичної поведінки вхідних даних та наявних управлінських пріоритетів.

Алгоритм індикатора ДСРБ проєктувався як відповідь на низку викликів, характерних для ресурсних даних індикатора та контексту прийняття рішень. Ці виклики були систематизовані у вигляді множини факторів Ф1–Ф11:

Ф1. Різноманітність ознак за типами (ресурсні, лагові, службові) та шкалами вимірювання.

Ф2. Складність розподілів (обмеженість області, скупченість, правосторонні «хвости»).

Ф3. Залежність поточного стану від ресурсної передісторії (інерційність процесів, часові лаги).

Ф4. Високий дисбаланс на користь позитивного (нормального) класу.

Ф5. Критичність гіподіагностики незадовільних станів (пропуск негативних станів є набагато небезпечнішим, ніж хибна тривога).

Ф6. Наявність нефіксованих зовнішніх факторів впливу, які можуть змінювати розподіли без прямого спостереження.

Ф7. Неоднорідність умов спостережень у часі та між різними підмножинами даних.

Ф8. Епізоди оцінювання цільової змінної за відмінними експертними моделями та шкалами.

Ф9. Обмежений обсяг доступних даних у порівнянні з потенційною складністю простору ознак.

Ф10. Використання різних типів класифікаційних моделей (імовірнісних, геометричних, деревоподібних) в одному ансамблі.

Ф11. Вимога придатності результатів класифікації для практичної підтримки рішень організаційної системи.

Призначення ДСРБ для процесів підтримки прийняття рішень в розподіленій організаційній системі розглядуваного типу акцентується на таких задачах:

- розпізнання ситуації, в якій ДСРБ може бути достатнім підґрунтям для прийняття рішення, та здійснення обґрунтованого прогнозу її стратегічної безпечності;
- отримання обґрунтованих гіпотез щодо потенційних ресурсних ризиків за умов високого рівня невизначеності;
- надання критерію для порівняльного аналізу послідовності періодів оперативного управління організаційною системою в аспекті стратегічної ресурсної небезпеки.

Аналіз останніх досліджень і публікацій

Використання ансамблів класифікаторів та дослідження їх ефективності для заданих типів задач та певної предметної області становить тематику багатьох сучасних досліджень [4]. Такі дослідження стосуються ансамблів різної структурної організації (гомогенних та гетерогенних), а також різних способів їх побудови (статичного або динамічного) [5]. Серед предметних областей, для яких незмінно актуальними є дослідження в області ансамблевої класифікації, привертають увагу медико-біологічна сфера [6], проблематика управління й підтримки прийняття рішень [7], а також кібербезпеки [8]. Задачі, в яких використовується такий інструментарій, стосуються більшості діагностики станів систем та визначення пріоритетних об'єктів вибору або втручання [9].

Серед властивостей ансамблю, які привертали інтерес дослідників, найчастіше представлені стандартні характеристики результатів в різних аспектах матриці помилок та показники стабільності [10]. Для дослідження валідності та перспективності наданого ансамблем інструментарію діагностики за умов, характерних для предметної області використання, в якості інформаційних об'єктів дослідження використовувалися як стандартні набори даних, представлені у відкритому доступі, так і синтетичні набори даних [11].

Окремий аспект методичного та інструментального ракурсу досліджень складає організація та підтримка відповідних обчислювальних експериментів [12].

Таким чином, має місце доменна орієнтованість досліджень в області ефективного ансамблювання класифікаторів. Це повною мірою стосується вивчення властивостей ансамблю за допомогою обчислювальних експериментів із генерацією даних, що імітують статистичну поведінку спостережень за об'єктами й процесами певної предметної області [13].

Об'єкти та процеси предметної області управління ресурсною безпекою стратегії розподілених організаційних систем були розглянуті раніше в [1]. Одну із найважливіших репрезентацій відповідної організаційної системи становить організаційна інфраструктура воєнної сфери національної безпеки. Дані ресурсних спостережень у цій області, які можуть використовуватися індикатором ДСРБ, мають спеціалізований статистичний характер

у зв'язку з процесами предметної області [14], [15]. Додатковий внесок до такої специфіки роблять використовувані регламенти спостережень та система пріоритетів, прийнята в процесах прийняття рішень [3].

Основними чинниками впливу виступають:

- різноманітність та динамічність поточних станів ресурсів різних типів;
- оцінювання задовільності ресурсного стану в долях від нормативів різної природи (за типом об'єкту управління, за делегованими спроможностями, за завданнями поточної ситуативної та антикризової діяльності);
- значна різниця в інтервалах спостережень за станом стратегічних цілей та за ресурсним станом об'єктів;
- високий дисбаланс класів з домінуванням благополучних станів;
- переважаюча ціна гіподіагностики міноритарного неблагополучного класу.

Все це обґрунтовує актуальність задачі дослідження ефективності пропонованого ансамблю класифікаторів за умов, близьких до практик його використання, за допомогою відповідних обчислювальних експериментів.

Формулювання мети дослідження

Метою роботи є визначення параметрів та властивостей запропонованого в [1] алгоритму індикатору ДСРБ, актуальних для його відповідного використання, та проведення обчислювальних експериментів щодо результативності алгоритму в аспекті цих характеристик.

Викладення основного матеріалу дослідження

Опис об'єкту експериментальних досліджень. Об'єктом досліджень є алгоритм індикатору ДСРБ, що здійснює підтримку машинного навчання та прогнозу стану стратегічних цілей РОС на основі даних про ресурсну забезпеченість її елементів. Алгоритм характеризується наступними рисами:

- 1) базується на ансамблі бінарних класифікаторів у складі: Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (kNN) та Logistic Regression (LR);
- 2) використовує синтетичні набори даних спеціальної структури.

Кожен з таких наборів даних:

- включає множину ресурсних ознак, вибірки значень яких імітують один із варіантів доменно обумовленої статистичної картини результатів спостережень за станом ресурсів елементів РОС [13];
- надає значення бінарної мітки благополучності стратегічного стану РОС, кожне з яких відповідає певному підперіоду зі складу періоду спостережень;
- містить лагові ознаки, які забезпечують синхронізацію даних за приналежністю до підперіодів;
- репрезентує як ознаки, корельовані з міткою, так і іррелевантні до неї.

Алгоритм ДСРБ визначає процес, до якого входять етапи:

E1. Структуризація системи ознак та масштабування.

E2. Навчання моделей ансамблю.

E3. Статичне зважування внесків моделей.

E4. Побудова динамічної метрики якості класифікації.

E5. Прогнозна інтегрована класифікація для спостережень тестової вибірки.

Докладний розгляд алгоритму було здійснено в [1].

Специфічні механізми, включені до складу алгоритму для відповіді на виклики Ф1–Ф11, розглянуті у постановці проблеми, охарактеризовані в табл. 1.

На основі призначення індикатору, розглянутого в постановці проблеми, до властивостей ДСРБ, що потребують вивчення за допомогою обчислювального експерименту, доцільно віднести:

- досягнуте значення інтегрального показника якості класифікації, оснований в алгоритмі ДСРБ на трьох взаємодоповнюючих критеріях [1];
- внесок спеціалізованих механізмів зі складу алгоритму ДСРБ, оцінюваний впливом на показники якості прогнозованої класифікації;
- критерій ризику недовиявлення негативних ресурсних ситуацій – показник Specificity

$$Spec = TN / (TN + FP) \quad (1)$$

обчислюється на основі елементів матриці помилок, являючи собою відношення числа істинно негативних класифікацій до суми істинно негативних та хибно позитивних;

- оцінку відносної ефективності ансамблю за інтегральним показником та критерієм ризику (1) в зіставленні з результатами окремих класифікаторів із його складу (особливо за умов невизначеності, при недостатніх для категоричного прийняття рішень значеннях критерію якості).

Вхідний набір даних алгоритму має структуру, модель якої ґрунтується на наступних параметрах та ознаках. Приймемо ряд позначень:

- TP – довжина періоду планової актуальності поточної стратегії РОС;
- DT – довжина підперіодів, наприкінці яких здійснюється оцінювання благополучності стану стратегічних цілей;

Таблиця 1

Втілені в алгоритмі ДСРБ методичні рішення

Виклики	Втілені в алгоритмі методичні рішення	Аргументація доцільності
Ф1	Розрізання ресурсних та лагових ознак	Різний препроцесинг для запобігання мультиколінеарності
Ф1, Ф3	Видалення спостережень з пропуском лагів та kNN доповнення ресурсних	Збереження послідовностей
Ф1, Ф2, Ф6, Ф10	Робастне масштабування ресурсних ознак в LR, kNN, SVM з використанням медіани та інтерквартильного розмаху (IQR) замість середнього та дисперсії (без зважування класів та видалення корельованих ознак)	Забезпечення однорідності ознак за шкалами та, що найважливіше, зменшення деструктивного впливу аномальних викидів, характерних для «важких хвостів» логнормальних розподілів
Ф4, Ф6, Ф7, Ф8	5 фолдів з почерговим використанням в ролі тестового	Стабільність оцінки якості
Ф4, Ф5	Апріорна метрика якості KQ_i для моделі M_i : зважена сума адаптованих критеріїв F1 (з інверсією класів), MCC та Каппа (перенормованої)	Акцент на виявленні негативних станів, баланс аспектів якості, зіставність
Ф10	Навчання M_i з крос-валідаційним калібруванням ймовірностей за методом Платта [16]	Отримання достовірних та уніфікованих імовірностей
Ф2, Ф3, Ф6, Ф10	Визначення коефіцієнтів довіри (KDR)	Регуляція впливу моделей згідно вразливості до властивостей спостережень, ще до етапу оцінки їхньої емпіричної точності
Ф3, Ф4, Ф5	Урахування фактичної корельованості елементів KQ_i для динамічної вагової параметризації з отриманням KQD_i для M_i	Адаптивна та об'єктивна оцінка якості, що пріоритезує меноритарний клас і запобігає домінуванню колінеарних метрик
Ф2, Ф3, Ф5, Ф10	Вагові коефіцієнти моделей M_i для наступного використання в голосуванні	Інтеграція вразливості M_i з продемонстрованою якістю
Ф10	Прогноз індивідуальних (від M_i) імовірностей класів для тестових точок	Зіставність M_i – прогнозів завдяки попереднім крокам
Ф2, Ф4, Ф6, Ф7, Ф8, Ф9	Інтеграція індивідуальних прогнозів для точок тестової вибірки зваженим м'яким голосуванням [16]	Підвищення впливу переваг, запобігання переоцінці більшого класу, агрегація розмитих рішень при складних розподілах ознак
Ф11	Оптимальна порогова дискретизація результатів класифікації (границя між класами встановлюється як значення ймовірності, яке максимізує KQD в точках тестової вибірки)	Створення можливості порівняння прогнозу класу з фактичними значеннями для тестової вибірки з метою наступної оцінки властивостей індикатору

- NT – номер такого підперіоду;
 - $L(NT) \in \{0,1\}$ – значення бінарної мітки благополучності стану стратегічних цілей для підперіоду з номером NT ;
 - NR – число планованих ресурсних спостережень в межах кожного з підперіодів.
- Тоді модель набору даних має наступний вигляд

$$DS = \{(FNT_i, L_i(NT_i), FH_i(NT_i), \{FR_{ij}\}_{j=1, \dots, R})\}_{i=1, \dots, NO}, \tag{2}$$

де FNT_i – номер підперіоду, до якого належить i -те спостереження; $L(NT)$ – мітка відповідного підперіоду; $FH_i(NT_i)$ – число підперіодів, що безпосередньо передують NT_i -му, в яких значення мітки співпадають з L_i ; FR_{ij} – доля нормативного значення ресурсу j -го виду, наявна на момент спостереження; R – число видів ресурсу; NO – кількість ресурсних спостережень.

Дві лагові ознаки FNT та FH забезпечують можливість впорядкування спостережень та врахування передісторії.

Алгоритм та засоби підтримки обчислювального експерименту. Для проведення обчислювальних експериментів із дослідження властивостей індикатору ДСРБ за умов різних статистичних моделей ресурсної ситуації був розроблений уніфікований алгоритм супроводження обчислювальних експериментів (СОЕ). До його задач належать:

- створення єдиного інформаційного середовища проведення експериментів;
- формування вхідних наборів даних, що відповідають моделі імітованої ресурсної ситуації та її стратегічних наслідків;
- відповідна параметризація звернень до бібліотечних процедур класифікаційних моделей елементів ансамблю;
- вибіркове включення до складу досліджуваного алгоритму ДСРБ механізмів, охарактеризованих в табл. 1, та параметризація процедур їх реалізації;
- управління експериментом (запуск, контроль, логування);
- систематизація, узагальнення та аналіз результатів;
- візуалізація результатів.

Алгоритм СОЕ реалізує наступну етапність процесу проведення експерименту.

Етап 1. Генерація батьківських вибірок.

Визначення структурних і об'ємних параметрів набору даних, видів і параметрів розподілу, характеристик шуму для міток та ознак, параметрів генерації. Виконання генерації.

Етап 2. Формування вхідних наборів даних.

Генерація послідовності міток благополучності підперіодів. Формування вибірок лагових ознак. Вибір та параметризація методів наповнення вибірок PT_i ресурсних ознак для i -го підперіоду добором значень із відповідних батьківських вибірок.

А) Пороговий метод, при якому умовою добору елементів до формованої вибірки має вид

$$((L_i = k) \wedge (|H| = m)) \Rightarrow \forall x \in PT_i : F_k(x, p, m), \quad (3)$$

де L_i – значення мітки в підперіоді; H – множина безпосередньо попередніх підперіодів з таким самим її значенням; $|H| = m$ – довжина періоду стабільності мітки; PT_i – множина елементів (спостережень), що належать підперіоду i ; F_0, F_1 – відповідно, предикати приналежності до періодів із міткою 0 або 1, що визначають порогові умови співвідношення між елементом x , параметрами розподілу ознаки p та довжиною періоду стабільності m .

Б) Зважений імовірнісний відбір значень, оснований на використанні контекстно залежної функції $S(x)$ корисності елементів x , взятих із батьківської вибірки.

До підперіодів із $L = 1$ відбір здійснюється з імовірністю p , а для $L = 0$ – з імовірністю $1 - p$. При цьому

$$p = 1 + \exp(-k \cdot S(x)), \quad (4)$$

$$S(x) = (\alpha c(x) + \beta t(x) + \gamma(2Lpr - 1))u(r) - \delta t(x)(1 - c(x)) + b, \quad (5)$$

де складові функції $S(x)$ відповідають метрикам корисності елемента x : $x(c)$ – міра близькості значення ознаки до 1; $t(x)$ – міра належності елемента до правого «хвоста» розподілу; $u(r)$ – функція значимості довжини періоду стабільності H (визначена в (3)).

Інші параметри в (4), (5) відображують пріоритет аспектів корисності.

Етап 3. Здійснення відбору механізмів.

До механізмів з табл. 1, які можуть бути чи не бути включені до складу алгоритму ДСРЗ, належать:

- використання в ролі базового показника якості класифікації критерію Каппа Коена;
- запровадження композитної метрики якості KQ , ґрунтованої на критеріях $F1_{neg}$ ($F1$, інвертованого щодо основного класу [17]), MCC та Каппа Коена (з апіорними вагами критеріїв або з динамічно налаштовуваними);
- форма інтеграції результатів різних моделей (за допомогою дискретного зваженого голосування або засобами ймовірнісного голосування – зі статичними апіорними вагами або адаптованими, а також із апіорними порогоми дискретизації результату чи з оптимізованими);
- калібрування моделей класифікаторів при навчанні.

Етап 4. Визначення складу результатів експерименту.

У межах експерименту заздалегідь визначається уніфікований склад вихідних артефактів, що забезпечує порівнянність між етапами та відтворюваність результатів. До нього відносяться:

- підсумкові метрики якості класифікації для ДСРБ та базових моделей (зокрема, за критеріями – Карра, MCC, $F1_{neg}$), а також показник ризику гіподіагностики негативного класу Specificity та композитна метрика якості KQ ;
- матриці помилок (TN/FP/FN/TP) та похідні показники;
- метадані запуску (параметри генерації даних і тренування, початкові значення генератору, версії, ідентифікатор серії);
- агреговані табличні представлення результатів серії прогонів (run-level і summary-level) для статистичної валідації, включно з метриками різниці показників якості для пари моделей (ДСРБ – Random Forest).

За потреби до складу результатів включаються візуалізації (графіки розподілів синтетичних ознак, діаграми стабільності/варіативності метрик), що інтерпретують властивості згенерованих даних і поведінку алгоритму.

Етап 5. Проведення експерименту.

Експеримент виконується як відтворюваний конвеєр запусків, у якому фіксуються конфігурації (YAML), seed-значення та версії артефактів, після чого послідовно запускаються процедури генерації наборів даних і навчання ДСРБ (переважно через *.bat-скрипти як основний механізм запуску). У процесі забезпечується журналювання (log files), контроль успішності за вихідним кодом та наявністю вихідних артефактів (результати, метадані, матриці помилок), а також повторюваність запусків шляхом централізованого збереження даних і результатів у стандартизованій структурі директорій. За потреби підтримується пакетний режим виконання (серії прогонів/запусків) із подальшим автоматизованим парсингом та агрегацією результатів.

Етап 6. Аналіз результатів експерименту.

Аналіз результатів експерименту здійснюється як послідовний процес інтерпретації якості та стійкості ДСРБ в контрольних і стресових умовах: виконується узагальнення метрик (зокрема Карра, MCC, $F1_{neg}$, Specificity, KQ) і матриць помилок із застосуванням базових характеристик описової статистики (середні значення, медіани та

стандартні відхилення), оцінюється варіативність результатів між запусками/початковим значенням генератору та зіставляються конфігурації «до/після» увімкнення ключових механізмів (KQ , імовірнісне голосування, калібрування, робастне масштабування) через зміну інтегральних показників. Додатково аналізується неоднорідність отриманих ефектів між наборами даних/умовами генерації та чутливість результатів до параметрів моделі даних (кореляційні та регресійні зіставлення). Окремо виконується формалізоване парне зіставлення ДСРБ з базовими класифікаторами з аналізом метрик різниці значень їх властивостей та статистичним підтвердженням значущості й практичної прийнятності величини ефекту.

Для статистичного аналізу достовірності різниці результатів моделей використовуються:

- t -критерій, в тому числі в варіанті [18], зі статистикою, скоригованою для випадку залежних підвибірок;
- критерій Вілксона (пріоритетний для асиметричних розподілів та при наявності «хвостів»);
- пермутаційний тест [19], оснований на випадкових перестановках результатів між двома моделями, в його

односторонньому варіанті (гіпотеза: $M1$ краща за $M2$).

Опис проведених експериментів та аналіз отриманих результатів. З метою валідації алгоритму ДСРБ та дослідження його властивостей було виконано три обчислювальні експерименти.

Кожен із них був підпорядкованим певній цілі дослідження, включав генерацію та формування необхідного й достатньо репрезентативного набору даних, а також використовував підмножину функцій алгоритму СОЕ, відповідну поставленій цілі.

Експеримент 1 мав на меті базову валідацію алгоритму ДСРБ на контрольній синтетичній моделі з нормальним розподілом.

Задачею експерименту було встановлення опорної «контрольної» оцінки якості прогнозування значення індикатору ДСРБ за умови високої роздільності класів, де очікується близька до ідеальної якість (див. рис. 1). Синтетичний набір даних був сформованим із структурою (2) на основі нормального батьківського розподілу з обрізанням у нулі (гарантування невід’ємності ознак). Для відбору значень ресурсних ознак був використаний пороговий підхід (3). Імітована структура періоду спостережень (див. підрозд. 4.1) включала 100 підперіодів із 30 елементами в кожному. Баланс класів, заданий значеннями мітки, становив 0.3 стосовно негативного класу. Лагові ознаки були сформовані на періодному рівні як детерміновані функції згенерованої послідовності 100 значень мітки L . При цьому для підперіоду з номером j $FNT(j) = L(j - 1)$, а $FH(j)$ відображає довжину поточної стабільної послідовності станів. Далі L , FNT і FH розгортаються в 30 спостережень реплікацією в межах періоду. Об’єм батьківських вибірок масштабується коефіцієнтом 10, що забезпечує резерв значень для фільтрації та стабільність генерації.

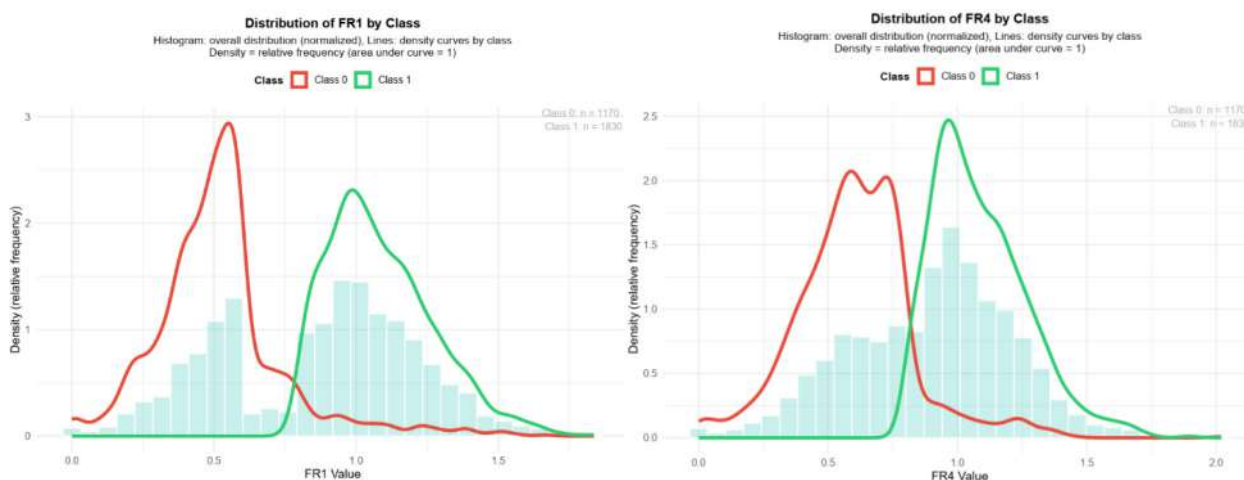


Рис. 1. Ілюстрація розподілів для ресурсних ознак в Експерименті 1

В якості сукупності механізмів, використаних в алгоритмі ДСРБ, був репрезентований мінімальний набір: каппа Коена як критерій якості класифікації; інтеграція результатів методом зваженого бінарного голосування з ваговими коефіцієнтами, що враховують критерій якості та рівень апріорної довіри до кожної моделі [1].

Верифікація алгоритму в 10 імітаційних повтореннях надала середні результати

$$Каппа = 0.9803; \quad Спец = 1.$$

Це дозволяє судити про успішність верифікації в аспекті коректної реалізації алгоритму. Але водночас такі результати вказують, що використані дані є «надто легкими» для алгоритму ДСРБ, що не дозволяє повноцінно дослідити внесок його спеціальних механізмів.

Таким чином, проведення подальших досліджень вимагало перегляду вхідного набору даних.

Експеримент 2 здійснювався з метою дослідження внеску окремих механізмів алгоритму ДСРБ до показників якості прогнозу класифікації.

Задачею дослідження було виявлення змін у показниках якості, що виникають при еволюції алгоритму ДСРБ, здійснюваній поетапним введенням додаткових механізмів, що належать до обґрунтованих у Таблиці 1 та врахованих у складі Етапу 3 алгоритму СОЕ.

Таким чином етапність Експерименту 2 передбачає наступне.

Етап 1. Додавання композитного критерію якості KQ [1] із налаштуванням використання KQ для оптимізації гіперпараметрів базових моделей.

Етап 2. Додавання механізму ймовірнісного голосування з оптимізацією порогу дискретизації для 4 моделей із збереженням дискретного для SVM.

Етап 3. Додавання до ймовірнісного голосування SVM із псевдо-ймовірностями.

Етап 4. Додавання крос-валідаційного уніфікованого калібрування всіх моделей за методом Платта.

На основі досвіду Експерименту 1 було сформовано більш реалістичні умови для індикатора ДСРБ – в аспектах дисбалансу класів, перекриття розподілів ознак, підвищеної варіативності (див. рис. 2). Набір даних генерується за тією самою структурною схемою і пороговим механізмом формування сигнальних ознак, лагові ознаки обчислюються аналогічно Етапу 1 як похідні від періодних міток L (без додаткового шуму міток), забезпечуючи часовий контекст для моделей за збереження незалежної генерації L по періодах. Обрана параметризація навмисно ускладнює задачу класифікації шляхом посилення класового дисбалансу (до 0.1 за негативним класом), збільшення варіативності батьківського розподілу ознак та зсуву порогів (предикати F , F_1 з (3)) у напрямку системного перекриття класів. У такий спосіб реалізується кероване «зашумлення» даних як стрес-модель для перевірки стійкості алгоритму ДСРБ та оцінювання ефекту запропонованих алгоритмічних покращень.

За використання базової версії алгоритму ДСРБ, ідентичної застосованій в Експерименті 1, для сформованих наборів даних було отримано такі значення метрик: $Kappa = 0.8112$; $Spec = 0.7232$.

Отже, подальші експерименти виконувалися в умовах зниженої інформативності вибірок порівняно з Експериментом 1.

Результати поетапної еволюції алгоритму наведено у вигляді значень метрик (KQ , $Spec$):

- Етап 1: (0.8452; 0.7321);
- Етап 2: (0.8970; 0.8482);
- Етап 3: (0.8820; 0.8661);
- Етап 4: (0.9071; 0.8661).

Загалом спостерігається узгоджене зростання обох метрик у процесі еволюції алгоритму. Тимчасове зниження значення KQ на Етапі 3 може бути пов'язане з проміжною фазою інтеграції моделі SVM, що передують досягненню стабільного балансу між узагальнюючою здатністю класифікатора та його специфічністю. Отримані результати свідчать про доцільність прийнятих при розробці алгоритму ДСРБ рішень щодо подолання базових викликів предметної області використання індикатора.

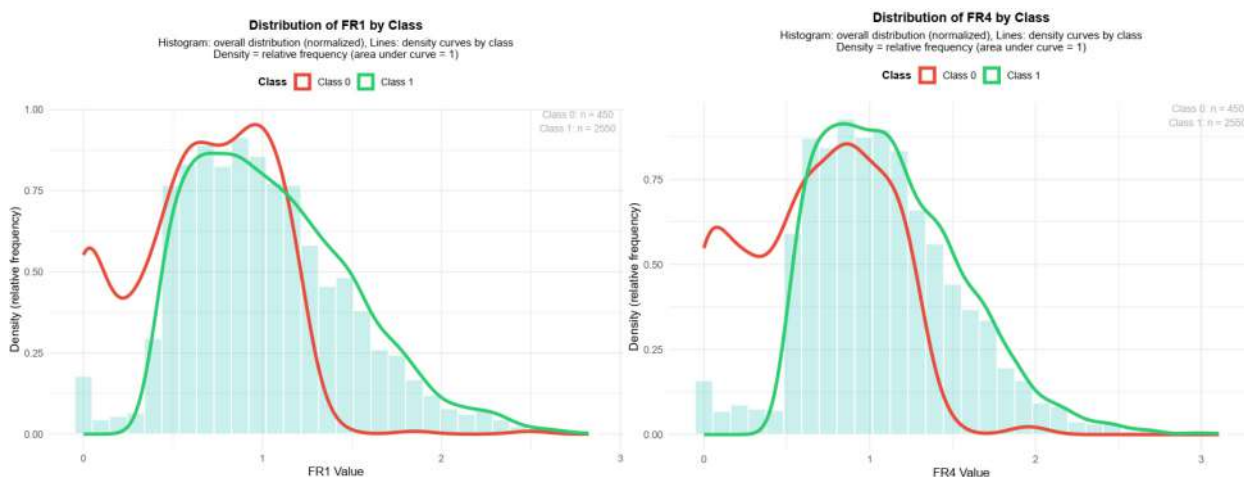


Рис. 2. Ілюстрація розподілів для ресурсних ознак Експерименту 2

Експеримент 3 було проведено з метою дослідження властивостей алгоритму ДСРБ в умовах значного впливу викликів, окреслених у постановці проблеми.

Основними завданнями експерименту були:

- генерація та наповнення відповідних наборів даних;
- навчання моделі та виконання прогнозування на заздалегідь виділених підмножинах наборів даних із використанням найбільш розвиненої версії алгоритму ДСРБ, що відповідає конфігурації Етапу 4 Експерименту 2;
- аналіз якості та стійкості результатів роботи алгоритму, зокрема в аспекті порівняння ефективності ансамблю з його найбільш успішним окремим елементом.

Для імітації ресурсної картини, характерної для систем, що розглядаються [3], було сформовано набір даних із чотирма сигнальними ресурсними ознаками FR_1-FR_4 , кожна з яких має логнормальний розподіл, де параметри розподілу задано таким чином:

$$\begin{aligned} (\mu_{in}^{(1)}, \sigma_{in}^{(1)}) &= (-0.4476, 0.67), & (\mu_{in}^{(2)}, \sigma_{in}^{(2)}) &= (-0.7090, 0.84), \\ (\mu_{in}^{(3)}, \sigma_{in}^{(3)}) &= (-0.5760, 0.76), & (\mu_{in}^{(4)}, \sigma_{in}^{(4)}) &= (-0.2110, 0.46). \end{aligned}$$

Інші 6 ресурсних ознак відіграють шумову роль та генеруються з довільними параметрами.

З метою ускладнення поверхні розділу між класами, посилення акценту на правому «хвості» розподілу та збільшення скупченості значень поблизу 1 було застосовано м'який відбір із батьківських вибірок (див. п. Б етапу 2) із параметрами $\alpha = 1, \beta = 0.9, \delta = 1.2, k = 4, \lambda = 1.5$ (див. (3)).

Частка негативного класу в наборі даних становила 0.2, при цьому зашумлення міток здійснювалося з імовірністю 0.03.

Поведінку сформованих розподілів ознак проілюстровано на рис. 3.

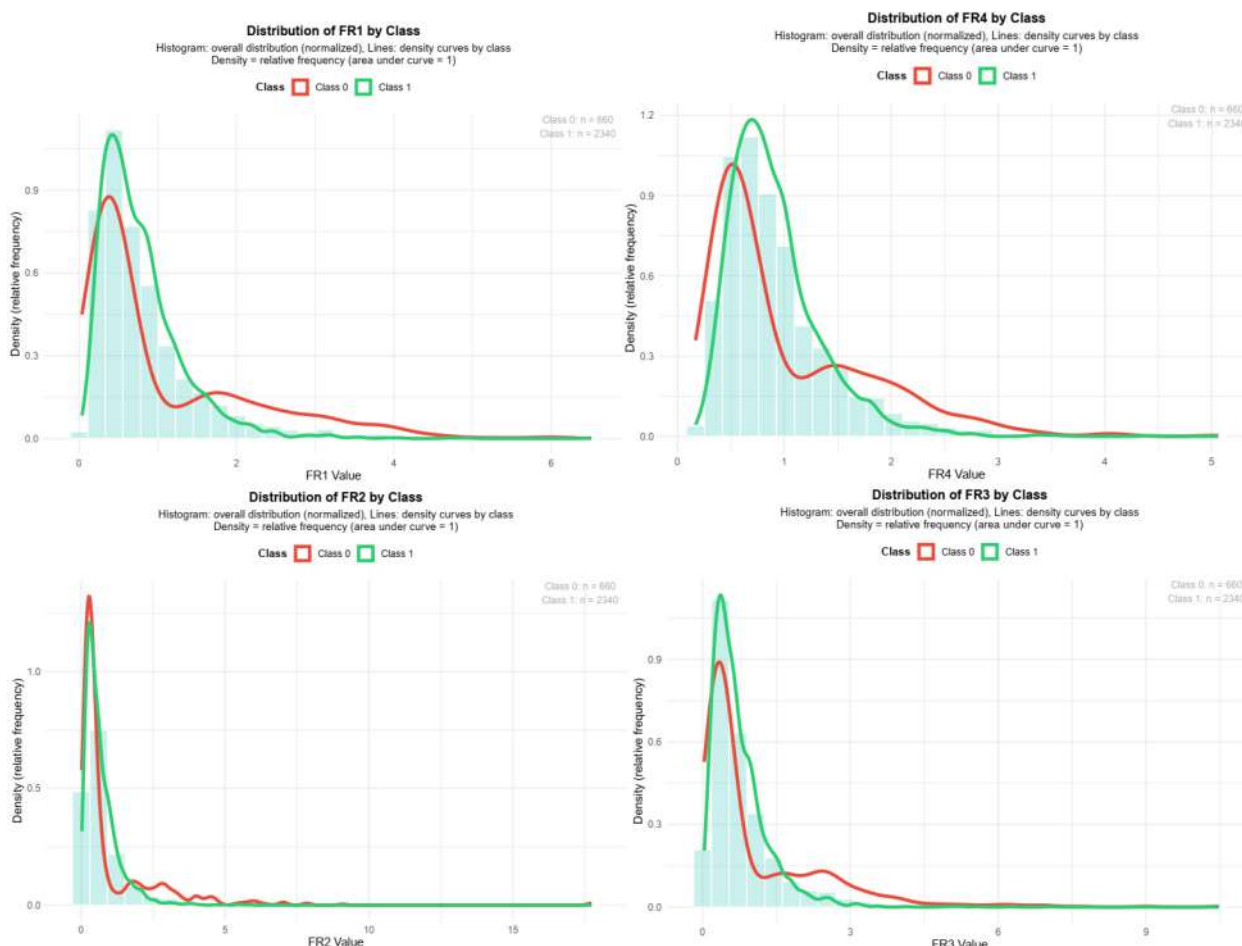


Рис. 3. Ілюстрація розподілів для чотирьох сигнальних ресурсних ознак Експерименту 3

Було згенеровано 5 наборів даних, для яких відрізнявся об'єм вибірок (кількість спостережень): 3000, 1500, 840, 364, 140. На кожному наборі даних було виконано 50 прогонів.

У табл. 2 наведено узагальнену оцінку переваги ансамблю ДСРБ над Random Forest для всіх 250 прогонів у парному дизайні (для кожного прогону обидві моделі оцінюються на тих самих даних). Таблиця одночасно

містить описову статистику різниць $\Delta = (RSI - RF)$ та результати статистичних тестів достовірності переваги (одностороння гіпотеза: алгоритм ДРСБ кращий за RF). Звертає увагу на себе перевага композитного індикатору порівняно з кращим компонентом – RF, яка має місце за всіма метриками якості.

Водночас середні значення KQ для обох індикаторів лежать у зоні розмитих рішень (ДРСБ: 0.6260; RF: 0.5578). Однак при сильному дисбалансі класів інтервал 0.6–0.7 для комбінованої метрики якості вважається прийнятним [20], [21].

Таблиця 2

Перевага ансамблю ДРСБ над RF у парному дизайні: описова статистика різниць та статистична значущість

Показник ($n = 250, \Delta = ДРСБ - RF$)	ΔKQ	$\Delta F1_{neg}$	$\Delta Spec$
Середнє значення Δ	0.0682	0.0823	0.2462
Медіана Δ	0.0499	0.0647	0.2143
Стандартне відхилення Δ	0.0683	0.0726	0.1559
Мінімум Δ	-0.0533	-0.0441	-0.0909
Максимум Δ	0.3715	0.3857	0.6818
Критерій Вілкоксона: p-value ($H1: \Delta > 0$)	2.06×10^{-38} (<0.001)	3.40×10^{-41} (<0.001)	1.28×10^{-41} (<0.001)
Критерій Вілкоксона: 95 % ДІ	[0.0528, ∞]	[0.0653, ∞]	[0.2222, ∞]
Парний t-критерій: p-value ($H1: \Delta > 0$)	1.20×10^{-39} (<0.001)	4.63×10^{-47} (<0.001)	4.49×10^{-70} (<0.001)
Парний t-критерій: 95 % ДІ	[0.0610, ∞]	[0.0610, ∞]	[0.0610, ∞]
Розмір ефекту (d Коена)	1.4112	1.6046	2.2341

Висновки

Використання додаткових механізмів адаптивної гармонізації та інтеграції (калібрування моделей, композитна метрика якості, зважене ймовірнісне голосування з оптимізацією порогів дискретизації) надає вагомий внесок до збільшення ефективності моделі гетерогенного ансамблю класифікаторів для індикатору ДРСБ. Відповідні ефекти підтверджені в обчислювальних експериментах на основі динаміки зростання досягнутої якості класифікації за трикомпонентним критерієм при покроковому еволюціонуванні алгоритму.

Досліджена ансамблева модель демонструє приблизно однаковий рівень якості з її кращим (і вже також ансамблевим за своєю структурою) компонентом-моделлю Random forest на статистичних даних помірної складності, коли комбінований показник якості класифікації $KQ > 0.8$ (область впевненого прийняття рішень).

При використанні в ситуаціях, складних для класифікації (високий дисбаланс класів, складна границя розділення класів, нелінійний вплив передісторії, асиметричний розподіл із важким «хвостом» та концентрацією біля точки $x = 1$), при значенні $0.6 < KQ < 0.7$ (прийнятна якість класифікації) ансамблева модель достовірно переважає Random forest як за значенням KQ , так і за показником специфічності (ризик гіподіагностики ресурсної небезпеки).

Це робить модель індикатору перспективною для знаходження зон підвищеної уваги та для порівняння стратегічної благополучності різних часових періодів.

Розроблений інструментарій підтримки обчислювальних експериментів надає засоби для проведення подальших досліджень для докладного визначення зон пріоритетності ансамблевої моделі індикатору ДРСБ.

Список використаної літератури

1. Skybyk S., Doroshenko A., Ilyina O., Sinitsyn I. Machine-learning-based model for indicators of the resource-based security of interests in high-level organizational systems. In: V. Snytyuk, V. Pleskach, I. Javorskyj, M. Dyvak, N. Goranin, V. Levashenko, E. Zaitseva, & O. Pursky (Eds.). Proc. 9th International Scientific and Practical Conference "Applied Information Systems and Technologies in the Digital Society" (AISTDS 2025) (CEUR Workshop Proceedings, Vol. 4133, pp. 153–169). 2025. CEUR-WS.
2. Ільїна О., Сініцин І., Слабоспицька О. Принципи та моделі експертно-аналітичної методології підтримки формування адаптивних організаційних рішень за умов глибокої невизначеності. *Проблеми програмування*. 2022. № 3–4. С. 364–375. DOI: <https://doi.org/10.15407/pp2022.03-04.364>
3. Сініцин І. П., Шевченко В. Л., Дорошенко А. Ю., Федоренко Р. М. *Моделі та програмні системи управління оборонними ресурсами: монографія*. Київ : ІПС НАНУ, 2024.
4. Dietterich T. G. Ensemble methods in machine learning. *Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science*. 2000. Vol. 1857. P. 1–15. DOI: https://doi.org/10.1007/3-540-45014-9_1
5. Kuncheva L. I. *Combining pattern classifiers: methods and algorithms*. New York : Wiley-Interscience, 2004.
6. Polikar R. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*. 2006. Vol. 6, N 3. P. 21–45. DOI: <https://doi.org/10.1109/MCAS.2006.1688199>

7. Hnatchuk Y., Lebedovska M. Decision support system for project resource planning based on the random forest method. *Computer Systems and Information Technologies*. 2025. Vol. 4. P. 35–42. DOI: <https://doi.org/10.31891/csit-2025-4-4>
8. Sommer R., Paxson V. Outside the closed world: on using machine learning for network intrusion detection. *Proc. of the 2010 IEEE Symposium on Security and Privacy (SP'10)*. Washington : IEEE, 2010. P. 305–316. DOI: <https://doi.org/10.1109/SP.2010.25>
9. Fawcett T., Provost F. Combining data mining and machine learning for effective user profiling. *Proc. 2nd International Conference on Knowledge Discovery and Data Mining (KDD'96)*. 1996. P. 8–13.
10. Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. 2006. Vol. 27, N. 8. P. 861–874. DOI: <https://doi.org/10.1016/j.patrec.2005.10.010>
11. Hastie T., Tibshirani R., Friedman J. *The elements of statistical learning*. New York : Springer, 2009.
12. Gentleman R., Lang D. T. Statistical analyses and reproducible research. *Journal of Computational and Graphical Statistics*. 2007. Vol. 16, N 1. P. 1–23. DOI: <https://doi.org/10.1198/106186007X178663>
13. Santner T. J., Williams B. J., Notz W. I. *The design and analysis of computer experiments*. New York : Springer, 2018.
14. Kress M. *Operational logistics: the art and science of sustaining military operations*. 2nd ed. New York : Springer International Publishing, 2016.
15. Степанюк М. Ю., Сініцин І. П., Котеля О. В. Проблема створення інформаційної системи логістики в Збройних силах України, що відповідає стандартам НАТО. *Проблеми програмування*. 2018. № 4. С. 101–110. DOI: <https://doi.org/10.15407/pp2018.04.101>
16. Large J., Lines J., Bagnall A. A probabilistic classifier ensemble weighting scheme based on exponentially weighting the probability estimates. *Data Mining and Knowledge Discovery*. 2019. Vol. 33. P. 1674–1709. DOI: <https://doi.org/10.1007/s10618-019-00638-y>
17. Sokolova M., Lapalme G. *A systematic analysis of performance measures for classification tasks*. *Information Processing & Management*. 2009. Vol. 45, N 4. P. 427–437. DOI: <https://doi.org/10.1016/j.ipm.2009.03.002>
18. Nadeau C., Bengio Y. Inference for the generalization error. *Machine Learning*. 2003. Vol. 52, N 3. P. 239–281. DOI: <https://doi.org/10.1023/A:1024068626366>
19. Ojala M., Garriga G. C. Permutation tests for studying classifier performance. *Journal of Machine Learning Research*. 2010. Vol. 11. P. 1833–1863. DOI: <https://doi.org/10.1109/ICDM.2009.108>
20. Baldi P., Brunak S., Chauvin Y., Andersen C. A. F., Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*. 2000. Vol. 16, N 5. P. 412–424. DOI: <https://doi.org/10.1093/bioinformatics/16.5.412>
21. Witten I. H., Frank E., Hall M. A., Pal C. J. *Data Mining: Practical Machine Learning Tools and Techniques*. 4th ed. Burlington : Morgan Kaufmann, 2016.

References

1. Skybyk, S., Doroshenko, A., Ilyina, O., & Sinitsyn, I. (2025). Machine-learning-based model for indicators of the resource-based security of interests in high-level organizational systems. In: V. Snytyuk, V. Pleskach, I. Javorskyj, M. Dyvak, N. Goranin, V. Levashenko, E. Zaitseva, & O. Pursky (Eds.). *Proc. 9th International Scientific and Practical Conference “Applied Information Systems and Technologies in the Digital Society” (AISTDS 2025) (CEUR Workshop Proceedings)*, 4133, 153–169. CEUR-WS.
2. Illina, O. P., Sinitsyn, I. P., & Slabospitska, O. O. (2022). Principles and models of expert-analytical methodology for adaptive organizational decisions forming under deep uncertainty. *Problems in programming*, (3–4) 364–375. <https://doi.org/10.15407/pp2022.03-04.364> [in Ukrainian].
3. Sinitsyn, I. P., Shevchenko, V. L., Doroshenko, A. Yu., & Fedorenko, R. M. (2024). *Models and software systems for defense resource management: monograph*. Institute of Software Systems of the National Academy of Sciences of Ukraine. [in Ukrainian].
4. Dietterich, T. G. (2000). Ensemble methods in machine learning. *Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science, 1857*, 1–15. https://doi.org/10.1007/3-540-45014-9_1
5. Kuncheva, L. I. (2004). *Combining pattern classifiers: methods and algorithms*. Wiley-Interscience.
6. Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, 6(3) 21–45. <https://doi.org/10.1109/MCAS.2006.1688199>
7. Hnatchuk, Y., & Lebedovska, M. (2025). Decision support system for project resource planning based on the random forest method. *Computer Systems and Information Technologies*, 4, 35–42. <https://doi.org/10.31891/csit-2025-4-4>
8. Sommer, R., & Paxson, V. (2010). Outside the closed world: on using machine learning for network intrusion detection. *Proceedings of the 2010 IEEE Symposium on Security and Privacy (SP'10)*, 305–316. <https://doi.org/10.1109/SP.2010.25>

9. Fawcett, T., & Provost, F. (1996). Combining data mining and machine learning for effective user profiling. *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD '96)*, 8–13.
10. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
11. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer.
12. Gentleman, R., & Lang, D. T. (2007). Statistical analyses and reproducible research. *Journal of Computational and Graphical Statistics*, 16(1), 1–23. <https://doi.org/10.1198/106186007X178663>
13. Santner, T. J., Williams, B. J., & Notz, W. I. (2018). *The design and analysis of computer experiments*. Springer.
14. Kress, M. (2016). *Operational logistics: the art and science of sustaining military operations*. 2nd ed. Springer International Publishing.
15. Stepaniuk, M. Y., Sinitsyn, I. P., & Kotelia, O. V. (2018). About applicability of NATO logistics information systems in Ukraine. *Problems in programming*, 4, 101–110. <https://doi.org/10.15407/pp2018.04.101> [in Ukrainian].
16. Large, J., Lines, J., & Bagnall, A. (2019). A probabilistic classifier ensemble weighting scheme based on exponentially weighting the probability estimates. *Data Mining and Knowledge Discovery*, 33, 1674–1709. <https://doi.org/10.1007/s10618-019-00638-y>
17. Sokolova, M., & Lapalme, G. (2009). *A systematic analysis of performance measures for classification tasks*. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
18. Nadeau, C., & Bengio, Y. (2003). Inference for the generalization error. *Machine Learning* 52(3), 239–281. <https://doi.org/10.1023/A:1024068626366>
19. Ojala, M., & Garriga, G. C. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research*, 11, 1833–1863. <https://doi.org/10.1109/ICDM.2009.108>
20. Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5), 412–424. <https://doi.org/10.1093/bioinformatics/16.5.412>
21. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: practical machine learning tools and techniques*. 4th ed. Morgan Kaufmann.

Дата першого надходження статті до видання: 08.01.2026

Дата прийняття статті до друку після рецензування: 12.02.2026

Дата публікації (оприлюднення) статті: 30.04.2026