

Т. П. ДОВЖЕНКО, канд. техн. наук;

ORCID:0000-0002-0352-8391

М. В. КУКЛІНСЬКИЙ, канд. техн. наук, доцент,

ORCID:0000-0002-2028-9206

Державний університет інформаційно-комунікаційних технологій, Київ

ЗБАЛАНСОВАНЕ ВИЯВЛЕННЯ АНОМАЛЬНИХ ТРАНЗАКЦІЙ МЕТОДОМ HYBRID AWRED НА РЕАЛЬНИХ ДАНИХ В УМОВАХ СИЛЬНОГО ДИСБАЛАНСУ КЛАСІВ

У статті було розглянуто задачу виявлення аномальних транзакцій у сильно незбалансованих фінансових даних. Ці дані включали в себе дуже малу кількість рідкісного класу порівняно з кількістю нормальних операцій. Така постановка виявилась складною не лише через малу кількість аномальних прикладів. Це також обумовлено тим, що частина підозрілих транзакцій не утворювала чітко відокремленої групи у просторі ознак, тому звичайна помилка реконструкції автоенкодера не завжди забезпечувала достатню якість ранжування і надійний вибір порогу. Для розв'язання цієї проблеми використано сьому версію методу HYBRID AWRED. В ньому використано гібридний підхід, у якому автоенкодерна реконструкційна основа поєднується зі стійким профілем каліброваних залишків.

Всі експерименти було виконано на наборі Credit Card Fraud Detection, який включав в себе 284807 транзакцій, 30 ознак і 492 аномальних записів. Частка аномального класу становила лише 0.1727%. Навчальна вибірка була сформована тільки з нормальних транзакцій. Валідаційна вибірка використовувалася для підбору порогу та параметрів HYBRID AWRED v7, а тестова вибірка застосовувалася тільки для фінальної оцінки якості. Такий підхід зменшив ризик витоку інформації з тестових даних у процедуру налаштування порогового рішення.

У методі HYBRID AWRED v7 використано фіксоване калібрування реконструкційних залишків після початкового навчання автоенкодера. Тут для кожної ознаки обчислювався калібрований залишок, після чого бралася його додатна частина і далі застосовувалося м'яке логарифмічне стискування $\ln(1 + r_j^+)$. Це зменшило вплив надто великих значень у правому хвості оцінки аномальності. Підсумкова оцінка була сформована як нормована лінійна комбінація каліброваного профілю залишків із невід'ємними вагами. Навчання AWRED-компоненти виконувалося з використанням псевдоаномалій і ранжувальної функції втрат. Такий підхід дозволив не тільки зменшити помилки реконструкції, а й покращити порядок розміщення нормальних та аномальних транзакцій за рівнем ризику.

За результатами п'яти незалежних прогонів HYBRID AWRED v7 досяг AUC-ROC 0.9432 ± 0.0117 , AUC-PR 0.6639 ± 0.0634 , F1 0.7153 ± 0.0402 , MCC 0.7180 ± 0.0399 , Precision 0.6608 ± 0.0565 і Recall 0.7831 ± 0.0501 . Порівняно з такими конкурентами як AE, AE-finetuned, DAE, SVDD і DAGMM запропонований підхід показав найкращу якість за AUC-PR, F1, MCC і повнотою. Але треба мати на увазі, що цей результат не слід трактувати як повне відділення аномальних транзакцій від нормальних. Частина слабких підозрілих прикладів залишається близькою до нормального профілю. Однак HYBRID AWRED v7 помітно покращив практично важливу порогову якість у задачі з екстремальним дисбалансом класів.

Ключові слова: аномальні транзакції, виявлення аномалій, автоенкодер, HYBRID AWRED v7, AUC-PR, дисбаланс класів, калібровані залишки, порогове рішення.

Вступ

Фінансові сервіси щоденно обробляють великі обсяги транзакцій, серед яких аномальні операції становлять дуже малу частку. Саме через це їх виявлення не є простою задачею класифікації. Більшість записів описує звичайну поведінку, а підозрілі операції часто не мають однієї чіткої ознаки. Вони можуть відрізнитися слабо, через поєднання кількох параметрів. Сильний дисбаланс класів змінює саму логіку оцінювання моделей. За таких умов висока загальна правильність класифікації ще не означає, що метод добре знаходить аномальні транзакції. Модель може майже безпомилково відносити нормальні операції до норми, але пропускати значну частину рідкісного класу. Тому для цієї задачі важливішими стають не лише загальні показники розділення, а й баланс між точністю спрацювання та повнотою виявлення.

У задачах такого типу часто використовують автоенкодери та інші глибокі моделі виявлення аномалій. Їхня перевага полягає в тому, що вони можуть навчатися переважно на нормальних прикладах. Далі аномальність визначається через помилку реконструкції або іншу оцінку відхилення від нормального профілю. Але звичайна середня помилка реконструкції не завжди достатньо чутлива до слабких локальних відхилень, особливо тоді, коли аномальна транзакція не є грубим викидом.

У цій роботі розглядається застосування HYBRID AWRED v7 для виявлення аномальних транзакцій у реальних фінансових даних. Основний акцент зроблено на формуванні більш придатної для порогового рішення оцінки аномальності. Йдеться не про повне відділення всіх підозрілих операцій від нормальних. Такий результат для реальних даних був би надто сильним припущенням. Завдання полягає в іншому: покращити співвідношення між виявленими аномаліями та хибними спрацюваннями в умовах дуже малої частки аномального класу.

Постановка задачі

Розглядається задача виявлення аномальних транзакцій у реальних фінансових даних, де нормальні операції суттєво переважають аномальні. У такій постановці модель має не тільки знайти якомога більше підозрілих операцій, а й не створити надмірної кількості помилкових спрацювань. Це складний компроміс особливо для практичних систем фінансового моніторингу.

Класичний автоенкодерний підхід зазвичай використовує загальну помилку реконструкції як оцінку аномальності. Це просте і зрозуміле рішення, але воно має обмеження. Якщо відхилення розподілене між кількома ознаками або має невелику амплітуду, середнє значення помилки може згладити важливу інформацію. До того ж різні ознаки мають різний масштаб залишків і без окремого калібрування вони по-різному впливають на підсумкову оцінку.

Проблема полягає в недостатній збалансованості наявних реконструкційних і глибоких методів виявлення аномалій у випадку рідкісного аномального класу. Частина методів добре ранжує записи в цілому, але після вибору порогу дає недостатню повноту або занадто багато хибних спрацювань. Тому потрібна оцінка аномальності, яка краще враховує структуру реконструкційних залишків і водночас залишається придатною для порогового рішення на валідаційній вибірці. У цій роботі така задача розв'язується на основі HYBRID AWRED v7.

Аналіз останніх досліджень і публікацій

Сучасні роботи з виявлення аномальних фінансових операцій показують, що головна складність цієї задачі пов'язана не тільки з вибором моделі. Важливими залишаються природний дисбаланс класів, спосіб підготовки даних, вибір порогу та метрики оцінювання. В оглядах останніх років наголошується, що штучне вирівнювання класів може змінювати властивості задачі й давати надто оптимістичну оцінку якості у реальних умовах [1], [2].

Окремо досліджується вплив виділення ознак і вибіркового балансування даних. У роботі [3] показано, що такі процедури справді можуть покращувати результат, але їхній ефект залежить від конкретного протоколу. Не менш важливим є поріг прийняття рішення. У статті [4] було звернуто увагу на те, що навіть за близького ранжування прикладів різний вибір порогу може суттєво змінити співвідношення між пропущеними аномаліями та хибними спрацюваннями.

У багатьох наукових дослідженнях порівнювалися бінарні та однокласові підходи до цієї задачі [5]. Так бінарні моделі часто показували вищі результати за наявності достатньої розмітності. Але в реальних умовах аномальні приклади часто бувають неповними або рідкісними. Тому реконструкційні моделі, такі як автоенкодери, залишаються важливими, оскільки вони можуть використовуватися як самостійна основа для оцінки аномальності або як проміжний засіб формування ознак [6], [7].

Для сильно незбалансованих даних все частіше використовують не одну метрику, а набір показників: AUC-PR, F1, MCC, Precision, Recall і AUC-ROC [8]. Ця стратегія виправдана, бо загальна здатність до розділення ще не гарантує прийнятної порогової якості. В інших роботах розвиваються складніші схеми з балансуванням, ансамблями та глибокими моделями [9]. Це може давати високі результати, але і ускладнювати пояснення того, який саме компонент забезпечив приріст.

Ще досить цікавими є підходи, де виявлення підозрілих операцій виконується без повного покладання на розмічені дані. У роботі [10] було розглянуто саме таку постановку на банківських транзакціях, що наближає задачу до реальних умов. Все це показує, що наявні дослідження добре охоплюють балансування, ансамблеві моделі, автоенкодерні схеми та однокласову класифікацію. Водночас недостатньо розкритою залишається побудова оцінки аномальності, яка не переносить остаточне рішення на зовнішній класифікатор, враховує структуру реконструкційних залишків і перевіряється на реальних фінансових даних із сильним природним дисбалансом класів.

Мета і задачі дослідження

З урахуванням поставленої проблеми подальше дослідження спрямоване не на збільшення складності моделі саме по собі, а на покращення якості порогового виявлення рідкісного аномального класу. Це важливо для фінансових даних, де навіть добра загальна здатність до ранжування не завжди означає прийнятний баланс між виявленими аномаліями та хибними спрацюваннями.

Метою дослідження є підвищення збалансованості виявлення аномальних транзакцій у реальних фінансових даних з сильним дисбалансом класів на основі методу HYBRID AWRED v7.

Для досягнення цієї мети було поставлено такі задачі:

1. Побудувати оцінку аномальності HYBRID AWRED v7 на основі профілю каліброваних реконструкційних залишків і м'якого логарифмічного стиснення їх додатної частини.
2. Виконати підбір порогу виявлення на валідаційній вибірці без використання тестових даних під час налаштування порогового рішення.
3. Порівняти HYBRID AWRED v7 з AE, AE-finetuned, DAE, SVDD і DAGMM за метриками AUC-ROC, AUC-PR, F1, MCC, Precision і Recall та визначити, чи забезпечує запропонований підхід кращу збалансованість виявлення аномальних транзакцій за дуже малої частки аномального класу.

Результати дослідження

1. Дані та експериментальний протокол

Експериментальну перевірку виконано на наборі даних Credit Card Fraud Detection, який містить реальні фінансові транзакції, подані в анонімізованому вигляді [11]. У вибірці наявні 284807 записів, 30 ознак і 492 аномальні транзакції. Частка аномального класу становить лише 0.1727%. Це створює досить жорстку умову для порогового виявлення, оскільки навіть невелика кількість хибних спрацювань може помітно знизити точність.

Навчальна вибірка формувалася тільки з нормальних транзакцій. Валідаційна вибірка використовувалася для підбору порогу та параметрів HYBRID AWRED v7, а тестова - лише для фінального оцінювання. У серії виконано п'ять незалежних прогонів. Якість оцінювалася за AUC-ROC, AUC-PR, F1, MCC, Precision і Recall, оскільки ці метрики краще відображають поведінку моделі за дуже малої частки аномального класу.

2. Побудова оцінки аномальності HYBRID AWRED v7

Поточна версія HYBRID AWRED v7 є розвитком попередньої методологічної лінії HYBRID AWRED. У ранній версії метод було перевірено на складному синтетичному фінансовому сценарії, де основний акцент робився на адаптивній реконструкційній оцінці та стійкості до складної структури даних [12]. Надалі було показано можливість формування оцінки аномальності без зовнішнього однокласового класифікатора, безпосередньо в просторі реконструкційних помилок [13]. У цій роботі ці ідеї перенесено на реальні фінансові транзакції з фактичним сильним дисбалансом класів.

Метод HYBRID AWRED v7 у цій роботі реалізовано на основі автоенкодерної реконструкції. Автоенкодер навчається на нормальних транзакціях і формує відновлений век-

тор ознак. Далі аналізується не лише загальна помилка реконструкції, а її профіль за окремими ознаками. Це важливо, бо в аномальній транзакції відхилення може бути не дуже великим у середньому, але помітним за певною групою координат.

Для вхідного вектора x та його реконструкції $\hat{x}$ покоординатна помилка задається як:

$$e_j(x) = (x_j - \hat{x}_j)^2, j = 1, \dots, d, \quad (1)$$

де d — кількість ознак, x_j — значення j -ї ознаки, $\hat{x}_j$ — її реконструйоване значення.

Після початкового навчання автоенкодера виконується калібрування залишків на нормальних транзакціях. Для кожної ознаки визначаються медіана та масштаб. Далі формується калібрований залишок:

$$rCal_j(x) = \frac{e_j(x) - med_j}{scale_j + \varepsilon}, \quad (2)$$

де med_j і $scale_j$ — робастні параметри калібрування для j -ї ознаки, ε — мала стала для числової стійкості.

У HYBRID AWRED v7 враховується тільки додатна частина каліброваного залишку:

$$rPos_j(x) = \max(0, rCal_j(x)). \quad (3)$$

Після цього застосовується м'яке логарифмічне стискання:

$$u_j(x) = \ln(1 + rPos_j(x)). \quad (4)$$

Цей крок є важливим для поточної версії методу. Він не відкидає великі залишки жорстко, але зменшує вплив надто великих значень у правому хвості. Тому оцінка аномальності не насичується так різко і краще зберігає ранжування транзакцій.

Підсумкова оцінка аномальності формується як лінійна комбінація каліброваного профілю:

$$s(x) = \sum_{j=1}^d w_j u_j(x), w_j \geq 0, \sum_{j=1}^d w_j = 1. \quad (5)$$

Така форма має дві переваги. По-перше, вона не зводить усі ознаки до простої середньої помилки. По-друге, ваги залишаються інтерпретованими, бо є невід'ємними і нормованими.

Схематично побудова HYBRID AWRED v7 представлена на рис. 1.

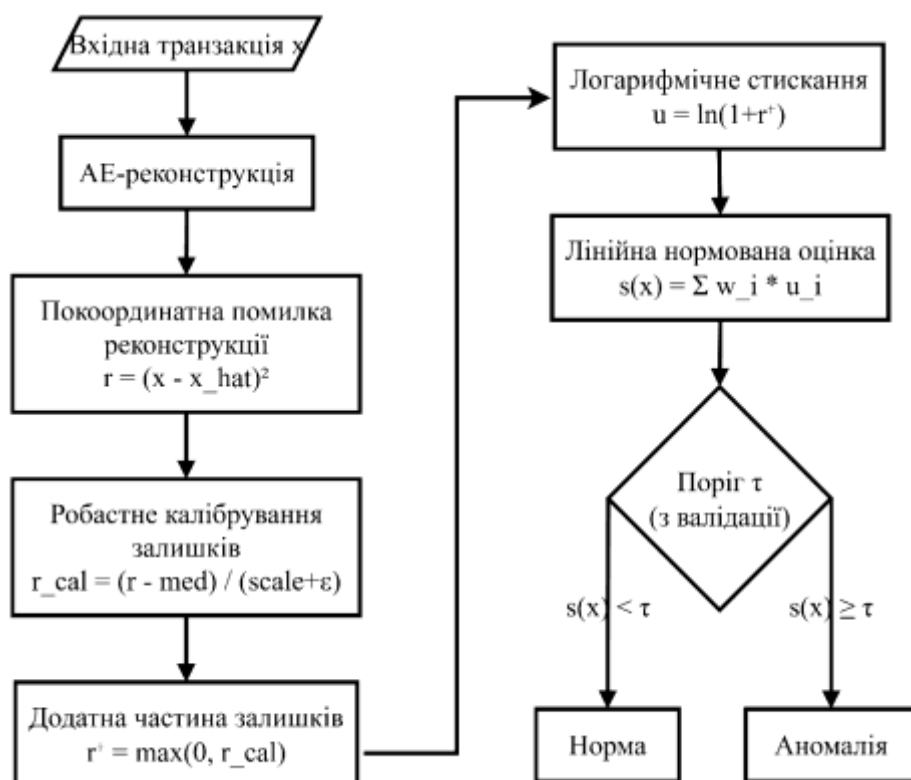


Рис. 1. Структурна схема формування оцінки аномальності методом HYBRID AWRED v7

3. Навчання та підбір параметрів

Після початкового навчання автоенкодера параметри калібрування залишків фіксуються. Далі HYBRID AWRED v7 навчається з використанням псевдоаномалій. Вони потрібні не для заміни реальних аномальних прикладів, а для створення контрасту між нормальним профілем і штучно порушеним профілем ознак.

Для навчання використовується ранжувальна функція втрат:

$$L_{rank} = \frac{1}{N} \sum_{i=1}^N \ln (1 + \exp (m + s(x_i^n) - s(x_i^p))) , \quad (6)$$

де x_i^n — нормальний приклад, x_i^p — псевдоаномальний приклад, m — параметр відступу, $s(\cdot)$ — оцінка аномальності HYBRID AWRED v7.

Параметри та коефіцієнт ваги ранжувального штрафу підбиралися за допомогою байєсівської оптимізації на валідаційній вибірці [14]. Для кожного прогону виконувалося 15 оцінювань. Оптимальні значення параметрів відрізнялися між прогонами, що є очікуваним за малої кількості аномальних прикладів у валідаційній частині. Водночас фінальна якість HYBRID AWRED v7 залишалася достатньо стабільною за F1, MCC і Recall.

4. Порівняння з базовими методами

Для порівняння використано п'ять методів: AE, AE-finetuned, DAE, SVDD і DAGMM. Автоенкодерні моделі є базовим класом реконструкційних підходів, де аномальність визначається через відхилення між вхідним і відновленим вектором [15]. DAE розширює цю логіку за рахунок навчання на зашумлених входах і відновлення стійкішого подання [16]. SVDD та його глибокі розширення представляють однокласову логіку, у якій нормальні об'єкти описуються через компактну область у прихованому просторі [17]. DAGMM-подібні моделі поєднують автоенкодерне подання з імовірнісною оцінкою через модель гауссових сумішей [18].

Результати за п'ятьма незалежними прогонами наведено у табл. 1. Значення подано як середнє $\pm$ СКВ.

Таблиця 1

Порівняння методів виявлення аномальних транзакцій за п'ятьма прогонами

Модель	AUC-ROC	AUC-PR	F1	MCC	Precision	Recall
AE	0.9389 $\pm$ 0.0118	0.4716 $\pm$ 0.1220	0.5379 $\pm$ 0.1211	0.5384 $\pm$ 0.1204	0.5510 $\pm$ 0.0925	0.5302 $\pm$ 0.1513
AE-finetuned	0.9415 $\pm$ 0.0111	0.4787 $\pm$ 0.1030	0.5429 $\pm$ 0.1014	0.5437 $\pm$ 0.0994	0.5659 $\pm$ 0.0667	0.5265 $\pm$ 0.1310
DAE	0.9362 $\pm$ 0.0124	0.5156 $\pm$ 0.0693	0.5827 $\pm$ 0.1024	0.5829 $\pm$ 0.1026	0.5752 $\pm$ 0.1039	0.5942 $\pm$ 0.1136
SVDD	0.9305 $\pm$ 0.0143	0.5040 $\pm$ 0.1811	0.5514 $\pm$ 0.1663	0.5555 $\pm$ 0.1595	0.5150 $\pm$ 0.1900	0.6088 $\pm$ 0.1196
DAGMM	0.9392 $\pm$ 0.0168	0.2013 $\pm$ 0.0885	0.3275 $\pm$ 0.1105	0.3671 $\pm$ 0.0969	0.2333 $\pm$ 0.0957	0.6053 $\pm$ 0.1031
HYBRID AWRED v7	0.9432 $\pm$ 0.0117	0.6639 $\pm$ 0.0634	0.7153 $\pm$ 0.0402	0.7180 $\pm$ 0.0399	0.6608 $\pm$ 0.0565	0.7831 $\pm$ 0.0501

Порівняння показує, що AUC-ROC у всіх сильніших моделей є досить близьким. Для сильно незбалансованих задач доцільно не обмежуватися ROC-кривою, а використовувати кілька метрик, які по-різному відображають поведінку моделі щодо рідкісного класу [19]. Тому в цій роботі AUC-PR, F1, MCC, Precision і Recall розглядаються як основні показники порогової якості. MCC додатково використано як збалансовану метрику, чутливу до всіх елементів матриці помилок [20]. Це очікувано для такого набору даних: ROC-крива не завжди чітко показує різницю між методами за дуже малої частки аномального класу. HYBRID AWRED v7 має AUC-ROC 0.9432 $\pm$ 0.0117, що лише трохи вище за AE-finetuned і DAGMM. Тобто головна перевага методу не в цьому.

Суттєвіша різниця спостерігається за AUC-PR. Для HYBRID AWRED v7 цей показник становить 0.6639 $\pm$ 0.0634. Найближчий конкурент, DAE, має 0.5156 $\pm$ 0.0693. Різниця є помітною і саме вона краще відображає поведінку моделі щодо рідкісного аномального класу. Так само HYBRID AWRED v7 має найвищі F1 і MCC: 0.7153 $\pm$ 0.0402 та 0.7180 $\pm$ 0.0399 відповідно.

Порівняння за ключовими метриками подано в табл. 1. Графічно поведінку методів додатково проілюстровано ROC- та PR-кривими для репрезентативного прогону на рис. 2.

5. Аналіз ROC- та PR-кривих

Для додаткової інтерпретації було побудовано ROC- та PR-криві для окремих прогонів. Як репрезентативний можна використати Run_01, оскільки його результат для HYBRID AWRED v7 є близьким до середнього профілю серії: AUC-ROC 0.9485, AUC-PR 0.6434, F1 0.7092, MCC 0.7121, Precision 0.6449 і Recall 0.7876.

ROC-криві показують, що кілька методів мають близьку глобальну здатність до розділення нормальних і аномальних транзакцій. Проте PR-криві краще виявляють різницю. Для HYBRID AWRED v7 крива точність–повнота утримується вище в тій області, де збільшення повноти ще не призводить до різкого падіння точності. Саме ця поведінка є важливою для практичного порогового рішення.

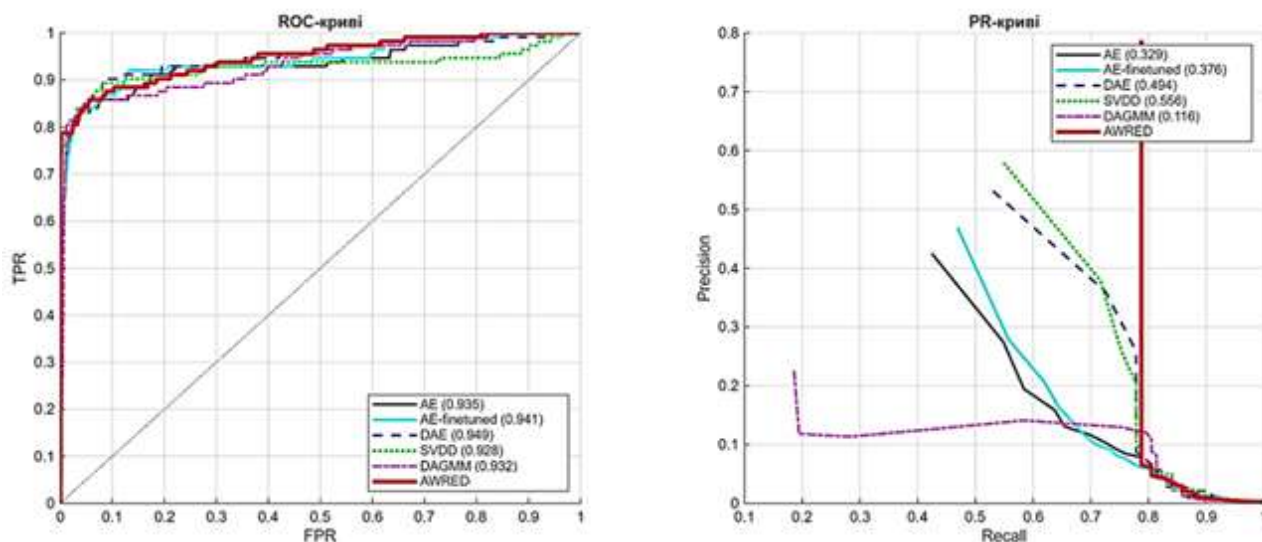


Рис. 2. ROC- та PR-криві методів для репрезентативного прогону

6. Розподіл оцінки аномальності

Окремо було проаналізовано розподіл оцінки аномальності HYBRID AWRED v7 для нормальних та аномальних транзакцій. Такий графік не замінює числові метрики, але добре показує характер роботи методу. У більшості прогонів аномальні транзакції зміщені в бік більших значень оцінки. Це означає, що метод справді підсилює відмінності між класами.

Важливо підкреслити, що разом із тим повного розділення аномальних прикладів немає. Частина їх залишається близькою до нормального класу. Реальні фінансові дані не формують дві повністю відокремлені групи. Тому результат HYBRID AWRED v7 слід трактувати не як абсолютне відділення аномального класу, а як покращення ранжування і порогової якості.

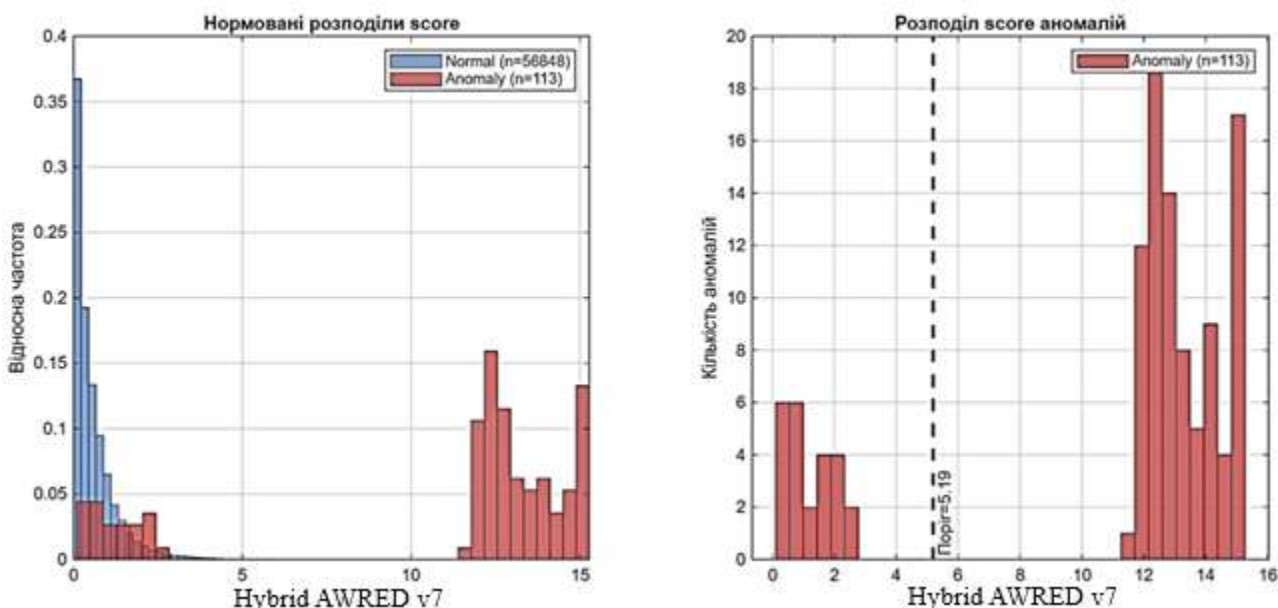


Рис. 3. Розподіл оцінки аномальності HYBRID AWRED v7 для нормальних і аномальних транзакцій

Висновки та перспективи подальших досліджень

У роботі побудовано оцінку аномальності методом HYBRID AWRED v7 для реальних фінансових даних із сильним дисбалансом класів. На відміну від звичайної середньої помилки реконструкції, дана оцінка використовує профіль каліброваних реконструкційних залишків. Додатна частина цих залишків стискається за допомогою натурального логарифмічного перетворення $\ln(1 + r_j^+)$, після чого формується нормована лінійна оцінка аномальності. Такий підхід дає змогу краще зберегти локальні відхилення за окремими ознаками. Це дає змогу не допустити надмірного домінування великих залишків у правому хвості розподілу.

На валідаційній вибірці проводився підбір порогу виявлення аномалій. Тестові дані при цьому не використовувалися для налаштування порогового рішення. Це важливо для задач з дуже малою часткою аномального класу, оскільки навіть незначне зміщення порогу може суттєво змінити співвідношення між хибними спрацюваннями та пропущеними аномаліями. Це дозволило оцінити метод без витоку інформації з тестової вибірки.

При порівнянні з AE, AE-finetuned, DAE, SVDD і DAGMM було виявлено, що HYBRID AWRED v7 забезпечує кращу збалансованість знаходження аномальних транзакцій. За п'ятьма незалежними прогонами отримано AUC-ROC 0.9432 ± 0.0117 , AUC-PR 0.6639 ± 0.0634 , F1 0.7153 ± 0.0402 , MCC 0.7180 ± 0.0399 , Precision 0.6608 ± 0.0565 та Recall 0.7831 ± 0.0501 . Основна перевага методу проявилася не стільки в AUC-ROC, скільки в AUC-PR, F1, MCC і Recall. Саме ці показники краще відображають якість роботи моделі при наявності дуже малої частки аномального класу.

Отримані результати не означають повного відділення аномальних транзакцій від нормальних. Деяка їх частина залишається близькою до нормального профілю. Тому HYBRID AWRED v7 слід розглядати як метод, що покращує ранжування і порогову якість виявлення, але не усуває повністю складність реальних фінансових даних.

Подальші дослідження доцільно спрямувати на перевірку методу на інших фінансових наборах, аналіз стабільності порогу за зміни розподілу транзакцій і підвищення інтерпретованості ваг у нормованій лінійній оцінці аномальності.

Декларація про штучний інтелект

Автори не використовували штучний інтелект при створенні матеріалів статті.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів.

Список використаної літератури

1. Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025). A systematic review of machine learning in credit card fraud detection under original class imbalance. *Computers*, 14(10), 437. <https://doi.org/10.3390/computers14100437>
2. Mienye, I. D., & Jere, N. (2024). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. *IEEE Access*, 12, 96893–96910. <https://doi.org/10.1109/ACCESS.2024.3426955>
3. Salekshahrezaee, Z., Leevy, J. L., & Khoshgoftaar, T. M. (2023). The effect of feature extraction and data sampling on credit card fraud detection. *Journal of Big Data*, 10, Article 6. <https://doi.org/10.1186/s40537-023-00684-w>
4. Leevy, J. L., Johnson, J. M., Hancock, J., & Khoshgoftaar, T. M. (2023). Threshold optimization and random undersampling for imbalanced credit card data. *Journal of Big Data*, 10, Article 58. <https://doi.org/10.1186/s40537-023-00738-z>
5. Leevy, J. L., Hancock, J., & Khoshgoftaar, T. M. (2023). Comparative analysis of binary and one-class classification techniques for credit card fraud data. *Journal of Big Data*, 10, Article 118. <https://doi.org/10.1186/s40537-023-00794-5>
6. Kennedy, R. K. L., Villanustre, F., Khoshgoftaar, T. M., & Salekshahrezaee, Z. (2024). Synthesizing class labels for highly imbalanced credit card fraud detection data. *Journal of Big Data*, 11, Article 38. <https://doi.org/10.1186/s40537-024-00897-7>

7. Ding, L., Liu, L., Wang, Y., Shi, P., & Yu, J. (2024). An AutoEncoder enhanced light gradient boosting machine method for credit card fraud detection. *PeerJ Computer Science*, 10, e2323. <https://doi.org/10.7717/peerj-cs.2323>
8. Breskuvienė, D., & Dzemyda, G. (2024). Enhancing credit card fraud detection: Highly imbalanced data case. *Journal of Big Data*, 11, Article 182. <https://doi.org/10.1186/s40537-024-01059-5>
9. Albalawi, T., & Dardouri, S. (2025). Enhancing credit card fraud detection using traditional and deep learning models with class imbalance mitigation. *Frontiers in Artificial Intelligence*, 8, Article 1643292. <https://doi.org/10.3389/frai.2025.1643292>
10. Karnavou, E., Cascavilla, G., Marcelino, G., & Geradts, Z. (2025). I know you're a fraud: Uncovering illicit activity in Greek bank transactions with unsupervised learning. *Expert Systems with Applications*, 288, Article 128148. <https://doi.org/10.1016/j.eswa.2025.128148>
11. Lim, Z. Y., Pang, Y. H., Kamarudin, K. Z. B., Ooi, S. Y., & Hiew, F. S. (2024). Bayesian optimization driven strategy for detecting credit card fraud with Extremely Randomized Trees. *MethodsX*, 13, Article 103055. <https://doi.org/10.1016/j.mex.2024.103055>
12. Довженко, Т. П. (2026). HYBRID AWRED: Синергія адаптивної реконструкції та топологічної кластеризації для виявлення аномалій у мультимодальних даних. *Зв'язок*, (1), 80–88. <https://doi.org/10.31673/2412-9070.2026.017405>
13. Dovzhenko, T., & Storchak, K. (2026). Advantages of the end-to-end Hybrid AWRED architecture in terms of the efficiency of visual anomaly detection under conditions of training data deficiency compared with the classical CNN + One-Class SVM ensemble. *Technology Audit and Production Reserves*, 3(2(89)), 53–59. <https://doi.org/10.15587/2706-5448.2026.362237>
14. Morales-Hernández, A., Van Nieuwenhuyse, I., & Rojas Gonzalez, S. (2023). A survey on multi-objective hyperparameter optimization algorithms for machine learning. *Artificial Intelligence Review*, 56(8), 8043–8093. <https://doi.org/10.1007/s10462-022-10359-2>
15. Nelay, A. A., & Turgeon, M. (2024). A comprehensive study of auto-encoders for anomaly detection: Efficiency and trade-offs. *Machine Learning with Applications*, 17, Article 100572. <https://doi.org/10.1016/j.mlwa.2024.100572>
16. Kascenas, A., Sanchez, P., Schrempf, P., Wang, C., Clackett, W., Mikhael, S. S., Voisey, J. P., Goatman, K., Weir, A., Pugeault, N., Tsiftaris, S. A., & O'Neil, A. Q. (2023). The role of noise in denoising models for anomaly detection in medical images. *Medical Image Analysis*, 90, Article 102963. <https://doi.org/10.1016/j.media.2023.102963>
17. Hojjati, H., & Armanfard, N. (2024). DASVDD: Deep autoencoding support vector data descriptor for anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, 36(8), 3739–3750. <https://doi.org/10.1109/TKDE.2023.3328882>
18. Tian, Y., Li, J., Song, Q., Li, Z., & Huang, X. (2023). Pyramid reconstruction assisted deep autoencoding Gaussian mixture model for industrial fault detection. *Information Sciences*, 649, Article 119682. <https://doi.org/10.1016/j.ins.2023.119682>
19. Gaudreault, J.-G., & Branco, P. (2024). Empirical analysis of performance assessment for imbalanced classification. *Machine Learning*, 113, 5533–5575. <https://doi.org/10.1007/s10994-023-06497-5>
20. Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, 16, Article 4. <https://doi.org/10.1186/s13040-023-00322-4>

T. Dovzhenko, M. Kuklinskyi

BALANCED DETECTION OF ANOMALOUS TRANSACTIONS USING HYBRID AWRED ON REAL DATA UNDER STRONG CLASS IMBALANCE

The article considered the problem of detecting anomalous transactions in strongly imbalanced financial data. These data included a very small amount of the rare class compared with the number of normal operations. Such formulation was difficult not only because of the small number of anomalous examples. It was also caused by the fact that part of suspicious transactions did not form a clearly separated group in the feature space, therefore the usual reconstruction error of the autoencoder

did not always provide sufficient ranking quality and reliable threshold selection. To solve this problem, the seventh version of the HYBRID AWRED method was used. It applies a hybrid approach in which the autoencoder reconstruction basis is combined with a stable profile of calibrated residuals.

All experiments were carried out on the Credit Card Fraud Detection dataset, which included 284807 transactions, 30 features and 492 anomalous records. The share of the anomalous class was only 0.1727%. The training sample was formed only from normal transactions. The validation sample was used for threshold selection and parameter tuning of HYBRID AWRED v7, and the test sample was used only for the final quality evaluation. This approach reduced the risk of information leakage from the test data into the threshold tuning procedure.

In the HYBRID AWRED v7 method, fixed calibration of reconstruction residuals was used after the initial training of the autoencoder. For each feature, a calibrated residual was calculated, after which its positive part was taken and then soft logarithmic compression $\ln(1 + r_j^+)$ was applied. This reduced the influence of too large values in the right tail of the anomaly score. The final score was formed as a normalized linear combination of the calibrated residual profile with non-negative weights. The AWRED component was trained using pseudo-anomalies and a ranking loss function. Such approach made it possible not only to reduce reconstruction errors, but also to improve the ordering of normal and anomalous transactions by risk level.

According to the results of five independent runs, HYBRID AWRED v7 achieved AUC-ROC 0.9432 ± 0.0117 , AUC-PR 0.6639 ± 0.0634 , F1 0.7153 ± 0.0402 , MCC 0.7180 ± 0.0399 , Precision 0.6608 ± 0.0565 and Recall 0.7831 ± 0.0501 . Compared with such competitors as AE, AE-finetuned, DAE, SVDD and DAGMM, the proposed approach showed the best quality by AUC-PR, F1, MCC and Recall. But it should be taken into account that this result should not be interpreted as a complete separation of anomalous transactions from normal ones. Part of weak suspicious transactions remains close to the normal profile. However, HYBRID AWRED v7 noticeably improved the practically important threshold quality in the task with extreme class imbalance.

Keywords: anomalous transactions, anomaly detection, autoencoder, HYBRID AWRED v7, AUC-PR, class imbalance, calibrated residuals, threshold decision.

Надійшла до редакції: 30.06.2026

Прийнята до друку: 14.07.2026

Опубліковано: 17.08.2026

© 2026 Довженко Т. П., Куклінський М. В.

Цей матеріал ліцензовано за умовами CC BY 4.0. <https://creativecommons.org/licenses/by/4.0/>