DOI 10.31891/2307-5732-2024-333-2-30
УДК 004.8

МЕЛЬНИК КАТЕРИНА
Луцький національний технічний університет
https://orcid.org/0000-0002-9991-582X
e-mail: nataliya@gmail.com
ЛАВРЕНЧУК СВІТЛАНА
Луцький національний технічний університет
https://orcid.org/0000-0002-5453-3924
e-mail: s.lavrenchuk@lntu.edu.ua
ХРИСТИНЕЦЬ НАТАЛІЯ
Луцький національний технічний університет
https://orcid.org/0000-0002-4836-7632
e-mail: hrystynets.at.ua@gmail.com

ВИЯВЛЕННЯ ШАХРАЙСТВА З КРЕДИТНИМИ КАРТКАМИ МЕТОДАМИ МАШИННОГО НАВЧАННЯ

В роботі проведено дослідження великих обсягів транзакцій кредитних карток з метою виявлення шахрайських операцій. Здійснено первинний аналіз дослідних даних та застосовано дві техніки усунення дисбалансу класів: випадкова недостатня вибірка, техніка SMOTE. Побудовано ряд класифікаторів для визначення шахрайських операцій, проведена статистична обробка отриманих результатів, що дозволило оцінити адекватність побудованих класифікаторів, визначено їх оптимальні параметри для максимальної ефективності. Для недостатньої та надмірної вибірки побудовані повноз'язні нейронні мережі з одним прихованим шаром та проведено порівняння їх точності.

Ключові слова: класифікатори, показники оцінки ефективності, надмірна вибірка, випадкова недостатня вибірка, логістична регресія, повнозв'язна нейронна мережа.

MELNYK KATERYNA, LAVRENCHUK SVITLANA, KHRYSTYNETS NATALIIA
Lutsk National Technical University

CREDIT CARD FRAUD DETECTION METHODS OF MACHINE LEARNING

In the work, a study of large volumes of transactions was conducted. The problem of fraud detection is unique because it is necessary to take into account: data imbalance (ie, fraudulent transactions are usually less than 1% compared to normal ones); fraud scenarios change over time and need to be detected quickly; transactions usually contain numerous categorical characteristics; as a result of the confidentiality of transactions, there are no publicly available datasets. All this creates problems with the development of classification methods and with the selection of performance evaluation metrics. Threshold indicators for evaluating the effectiveness of classifiers were studied and the expediency of using thresholdless metrics was substantiated. There is currently no consensus on which set of performance indicators should be used. The primary analysis of research data was carried out and two techniques for eliminating class imbalance were applied: random undersampling, SMOTE technique. Removal of outliers was shown to improve the accuracy of the classification methods by more than 3%. A number of classifiers were built to determine fraudulent operations, statistical processing of the obtained results was carried out, which allowed to assess the adequacy of the built classifiers, to determine their optimal parameters, at which the classifiers work with maximum efficiency. It should be noted that all classifiers were tested on real data. Determined: The logistic regression classifier performs best on both the training and cross-validation sets. The Precision-Recall indicator was used to assess the effectiveness of the logistic regression model. For undersampling and oversampling, fully connected neural networks with one hidden layer were constructed and their accuracy was compared. It should be noted that the neural network on the oversampled data set predicts fewer correct fraud transactions than the model using the undersampled data set.

Keywords: classifiers, performance metrics, dimensionality reduction, oversampling, random undersampling, logistic regression, fully connected layer neural network.

Постановка проблеми

На основі звіту Європейського центрального банку [1] за 2020-2021 рік проаналізовано загальний рівень шахрайства в карткових платіжних системах, що має тенденцію до зменшення, роглянуто основні категорії шахрайства. Загальний рівень шахрайства в карткових платіжках свідчить про важливість постійного моніторингу та заходів безпеки з боку наглядачів карткових систем.

Проблема виявлення шахрайства є унікальною, оскільки дані про транзакції кредитної картки дуже незбалансовані, тобто шахрайських транзакцій зазвичай становить менше 1%, що створює проблеми і з розробкою методів класифікації і з вибором метрики оцінки продуктивності. В роботі застосовується дві техніки усунення дисбалансу класів: випадкова недостатня вибірка, техніка SMOTE.

Крім незбалансованості вибірки, задача виявлення шахрайства має ще ряд особливостей [2]:

– дрейф концепції: моделі транзакцій і шахрайства змінюються з часом;

– вимоги майже до реального часу: побудовані системи повинні мати можливість швидко виявляти шахрайські транзакції;

– дані транзакцій зазвичай містять численні категоріальні характеристики, такі як ідентифікатор клієнта, термінал, тип картки тощо. Загальні стратегії трансформації категоріальних ознак на числові включають агрегацію функцій, трансформацію на основі графів або підходи глибокого навчання;

– кожна кінцева точка та/або клієнт генерує послідовний потік даних з унікальними характеристиками. Здійснюється моделювання цих потоків для кращого опису їх очікуваної поведінки та виявлення випадків

аномальної поведінки. Моделювання можна виконувати шляхом агрегування функцій у часі (наприклад, відстеження середньої частоти або суми транзакцій клієнтів) або покладаючись на послідовні прогнозуючі моделі (наприклад, приховані моделі Маркова або рекурентні нейронні мережі);

– перекриття класів: майже неможливо відрізнити шахрайські транзакції від справжніх лише на основі необробленої інформації про транзакції. Цю проблему вирішують за допомогою методів розробки функцій, які додають контекстну інформацію до необробленої платіжної інформації;

– показники ефективності: стандартні показники для систем класифікації, такі як середня помилка неправильної класифікації або AUC ROC, не дуже підходять для вирішення проблем виявлення шахрайства. Для оцінки загальної ефективності системи виявлення шахрайства розглядають кілька методів. Наразі немає консенсусу, який набір показників ефективності слід використовувати;

– відсутність загальнодоступних наборів даних: оскільки конфіденційні транзакції реальних кредитних карток не можна оприлюднити. Існує лише один загальнодоступний набір даних на Kaggle (хоча він обмежений двома днями) [3], який був опублікований у 2016 році.

Аналіз останніх джерел

В роботі [4] наведено дані про програмні інструменти для виявлення шаблонів і аномалій у даних транзакціях, які використовують алгоритми машинного навчання та розширену аналітику. За їх допомогою фінансові установи у реальному часі відстежують та виявляють підозрілі дії, що запобігає обробці шахрайських транзакцій. При виявлені шахрайських дій в основному використовують три різні системи:

– Системи на основі правил (мають попередньо визначені правила для підозрілих транзакцій, і будь-які транзакції, які відповідають цим критеріям, позначаються для перевірки) [5];

– Системи виявлення аномалій (на основі алгоритмів машинного навчання виявляють підозрілі дії, аналізуючі дані минулих транзакцій, позначаючи будь-яке відхилення) [6-8];

– Системи прогнозного моделювання (створюються у режимі реального часу, використовуючи історичні дані та передову аналітику) [10-11].

При наявності великого обсягу даних про транзакції з позначками, можливо тренувати класифікатори на основі машинного навчання. Це дає такі переваги як: здатність ідентифікувати нові закономірності та адаптуватися до змін у сценаріях шахрайства, висока точність та зменшення ручної роботи, зменшення помилок при класифікаціях. Для виявлення шахрайства використовуються різні методи контрольованого навчання, такі як дерева рішень [12], нейронні мережі зворотного поширення [13], опорні векторні машини [14], випадкові ліси [15] та байєсовські мережі [16]. Але такий підхід ефективний лише для виявлення шахрайства за подібними моделями, що були ідентифіковані як шахрайство в минулому і не придатні для виявлення нових моделей. Неконтрольовані методи виявлення шахрайства більш гнучкі [17]. Вони не потребують позначених даних для навчання, тому можуть адаптуватися до нових типів шахрайства. Інший клас методів аналізує поведінку окремого користувача [18], враховуючи індивідуальні особливості, і дає можливість виявити нові типи шахрайства. Цей підхід може бути використаний, як доповнення до традиційних методів машинного навчання.

Метою роботи є розробка ефективної прогнозуючої моделі для визначення шахрайства з кредитними картками на основі алгоритмів машинного навчання з врахуванням незбалансованої навчальної виборки.

Виклад основного матеріалу

Опис набору даних, на яких проводились дослідження. Через конфіденційність інформації можливі два варіанти представлення даних: симулятор даних транзакції, або реальні дані, частина яких перетворена методом головних компонентів (PCA). Дослідження проводилося на реальних даних, що включають інформацію про транзакції, здійснені європейськими власниками кредитних карток у вересні 2013 року. Цей набір даних зосереджений на двох днях, під час яких було здійснено 284,807 транзакцій, з яких 492 виявилися шахрайськими. Спостерігається значна незбалансованість у розподілі класів, де позитивний клас (шахрайство) складає лише 0,172% від усіх транзакцій.

Отже, шість основних характеристик, які підсумовують транзакцію: ідентифікатор транзакції; дата й час здійснення транзакції; ідентифікатор клієнта; ідентифікатор терміналу; сума транзакції; мітка шахрайства (двійкова змінна зі значенням 0 для законної транзакції та 1 для шахрайської операції).

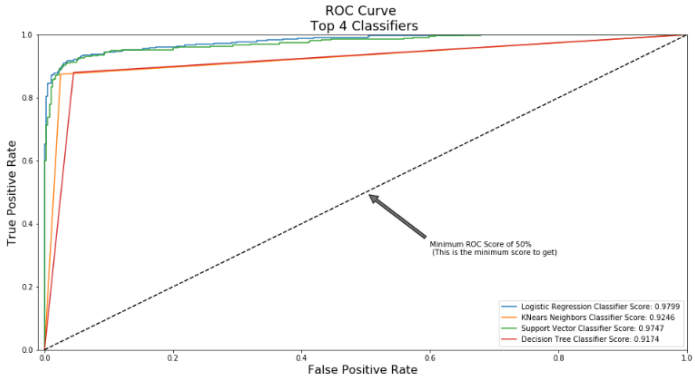
Основні етапи дослідження: створення підвибірки з однаковою кількістю шахрайських та не шахрайських операцій (або або надлишкової вибірки), визначити класифікатори, які мають найвищу точність визначення шахрайства, створити нейронну мережу та здійснити порівняння точності з найкращим класифікатором. Дослідити типові помилки, які виникають при роботі з незбалансованими наборами даних.

Дослідження проводиться в декілька етапів. Перед використанням техніки випадкової недостатньої вибірки або надлишкової вибірки (для отримання більш збалансованого набору даних та для уникнення перенавчання моделей), виділяється підвибірка, що містить оригінальні дані для тестування. А також використована стратифікована перехресна перевірка kfold, яка використовується для незбалансованих даних та коли розмір даних невеликий. Основна проблема з «випадковою недостатньою вибіркою» полягає в ризику втрати точності моделі класифікації, оскільки втрачається значна кількість інформації.

Зі збалансованими даними проведена попередня обробка даних та кореляційний аналіз для виявлення істотної залежності між даними і які з них істотно впливають на визначення, що певна транзакція є шахрайством. Наступним кроком видалялися «надзвичайні викиди» з функцій, які мають високу кореляцію з нашими класами, використовується метод інтерквартильного діапазону. Показано, що впровадження

зменшення викидів покращило точність методів класифікації більш ніж на 3%.

Побудовано чотири типи класифікаторів для виявлення шахрайських операцій і проведена оцінка їхньої ефективності (рис. 1, а). Як видно, логістична регресія досить точно проводить класифікацію. На рис.1, б проведено порівняння показника акуратності (accuracy_score) для класифікатора логістичної регресії з технікою надмірної вибірки (SMOTE) та випадкової недостатньої вибірки (Random UnderSampling). Як видно з рис.1, техніка надмірної вибірки показує вищий показник акуратності.



а)

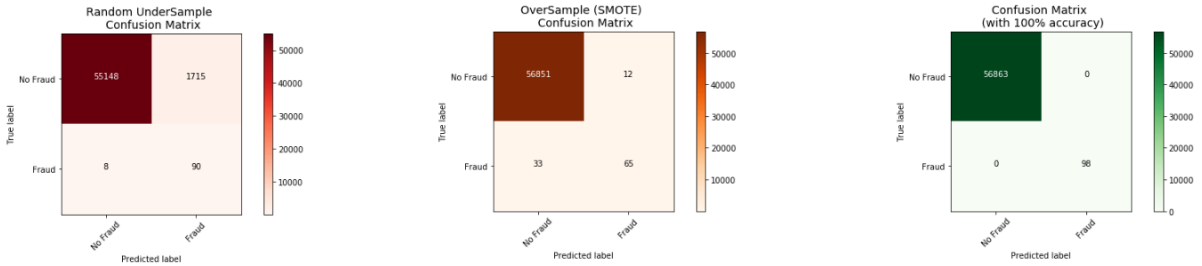
Техніка¤	Score¤
Random·UnderSampling¤	0.942105¤
Oversampling·(SMOTE)¤	0.987079¤

б)

Рис. 1. Оцінки ефективності класифікаторів: а) ROC криві для класифікаторів; б) Показники акуратності для логістичної регресії для випадкової недостатньої та надлишкової вибірки

Побудовані класифікатори на основі повнозв'зної нейронної мережі, що складається з одного вхідного шару, одного прихований шар (32 вузла) та один вихідний вузол, що складається з двох можливих результатів 0 або 1 (звичайна та шахрайська транзакція) [19]. Інші характеристики: швидкість навчання 0,001, оптимізатор AdamOptimizer, функція активації Relu, а для кінцевих результатів використовується розріджена категоріальна перехресна ентропія, яка дає ймовірність того, чи є випадок шахрайством чи ні.

На рис. 2 наведено підсумок роботи нейронних мереж, які побудовані на недостатній виборці (рис. 2, а) та на надмірній виборці (рис. 2, б) у вигляді матриці плутанини без нормалізації даних. Проведено порівняння з фактичними даними (рис. 2, в).



а) **б)** **в)**

Рис. 2. Матриця плутанин для прогнозованих нейронних значень: а) недостатня вибірка; б) надмірна вибірка); в) оригінальних значень

Порівнюючи результати роботи нейронної мережі на наборі даних із надмірною та недостатньою вибіркою, отримаємо, що в першому випадку отримані менш правильні транзакції шахрайства, ніж модель, яка використовує набір даних із недостатньою вибіркою.

Висновки

Проблема виявлення шахрайства є унікальною, оскільки необхідно враховувати: незбалансованість даних (тобто шахрайських транзакцій зазвичай становить менше 1% в порівнянні з звичайними); шахрайства змінюються у часі; необхідність швидко виявляти шахрайство; транзакцій зазвичай містять численні категоріальні характеристики. Все це створює проблеми із розробкою методів класифікації, і з вибором метрики оцінки продуктивності. Точність матриці плутанини може бути нерепрезентативною на таких даних. Досліджені порогові показники оцінки ефективності класифікаторів та обґрунтовано доцільність використання безпорогових метрик, яка вимірюється за допомогою площі під кривою «точність – повнота».

Проведено первинний аналіз дослідних даних та застосовано дві техніки усунення дисбалансу класів: випадкова недостатня вибірка, техніка SMOTE. Методика видалення викидів було здійснена лише у наборі даних із недостатньою вибіркою. Проведено масштабування та розподіл даних.

Досліджено ефективність роботи класифікаторів на випадковій недостатній вибірці: логістичної регресії, K-найближчих сусідів, машини опорних векторів, дерева рішень. Проведено тестування та порівняльний аналіз. Визначено: класифікатор логістичної регресії показує найкращий результат як на наборах навчання, так і на наборах перехресної перевірки. При дослідженні надійності (точності) моделі

логістичної регресії використовувався показник «точність-повнота» (Precision-Recall). Отриманий компроміс: при зниженні точності, модель зможе виявити більше випадків шахрайства. Отже, при зниженні точності між 0,90 і 0,92 (показник точності високий), отримаємо все ще високий показник повноти.

Досліджено роботу класифікатора логістичної регресії з даними надмірної вибірки (SMOTE). Слід зазначити, що тестування всіх класифікаторів проводилось на реальних даних, і не тестувалися на наборі даних із надмірною чи недостатньою вибіркою. Проведено порівняння показника акуратності (accuracy_score) для класифікатора логістичної регресії з технікою надмірної вибірки SMOTE та випадкової недостатньої вибірки Random UnderSampling. Техніка надмірної вибірки показує вищий показник акуратності.

Для недостатньої та надмірної вибірки побудовані повноз'язні нейронні мережі з одним прихованим шаром та проведено порівняння їх точності. Слід зазначити, що нейронна мережа на наборі даних із надмірною вибіркою передбачає менш правильні транзакції шахрайства, ніж модель, яка використовує набір даних із недостатньою вибіркою. Крім того, для даних із недостатньою вибіркою нейронна мережа не може правильно виявити велику кількість випадків нешахрайських транзакцій і натомість неправильно класифікує ці нешахрайські транзакції, як випадки шахрайства. Подальші дослідження передбачають видалення викидів із набору даних надмірної вибірки та перевірка чи покращиться точність на тестовому наборі.

Література

1. Report on card fraud in 2020 and 2021. European Central Bank. URL: http://surl.li/lycns (date of access: 08.10.2023).

2. Лян П. Credit Card Fraud Detection: A Note to CCFD Analysis. https://medium.com/codex/credit-card-fraud-detection-a-concise-note-to-ccfd-analysis-6ad521f5366d (дата звернення: 07.11.2023).

3. Credit Card Fraud Detection. URL: http://surl.li/fmdgz (date of access: 13.11.2023).

4. Top 10 Credit Card Fraud Detection Solutions in 2023. CybeReady. URL: https://cybeready.com/comprehensive-guide-to-fraud-detection-management-and-analysis/top-10-credit-card-fraud-detection-solutions-in-2023 (date of access: 04.03.2024).

5. Building A Strong Defense: Strengthening Compliance With Transaction Monitoring Rules. Financial Crime Academy. URL: http://surl.li/rgujo (date of access: 06.03.2024).

6. Narayana, M. S., Prasad, B. V. V. S., Srividhya, A., & Reddy, K. P. R. (2011). Data mining machine learning techniques–A study on abnormal anomaly detection system. International Journal of Computer Science and Telecommunications, 2(6).

7. Rao, K. H., Srinivas, G., Damodhar, A., & Krishna, M. V. (2011). Implementation of anomaly detection technique using machine learning algorithms. International journal of computer science and telecommunications, 2(3), 25-31.

8. Mokhtari, S., Abbaspour, A., Yen, K. K., & Sargolzaei, A. (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. Electronics, 10(4), 407.

9. Tertychnyi, P., Godgildieva, M., Dumas, M., & Ollikainen, M. (2022). Time-aware and interpretable predictive monitoring system for Anti-Money Laundering. Machine Learning with Applications, 8, 100306.

10. Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., & Kuruwitaarachchi, N. (2019, January). Real-time credit card fraud detection using machine learning. In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 488-493). IEEE.

11. Mittal, S., & Tyagi, S. (2020). Computational techniques for real-time credit card fraud detection. Handbook of Computer Networks and Cyber Security: Principles and Paradigms, 653-681.

12. Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised machine learning algorithms for credit card fraud detection: a comparison. In 2020 10th international conference on cloud computing, data science & engineering (confluence) (pp. 680-683). IEEE.

13. Ghosh, R. (1994). Credit card fraud detection with a neural-network. In Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, Wailea, HI (pp. 621–630).

14. Gyamfi, N. K., & Abdulai, J. D. (2018, November). Bank fraud detection using support vector machine. In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 37-41). IEEE.

15. Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018, March). Random forest for credit card fraud detection. In 2018 IEEE 15th international conference on networking, sensing and control (ICNSC) (pp. 1-6). IEEE.

16. Yee, O. S., Sagadevan, S., & Malim, N. H. A. H. (2018). Credit card fraud detection using machine learning as data mining technique. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(1-4), 23-27.

17. Bolton, R. J., & Hand, D. J. (2001) Unsupervised profiling methods for fraud detection. In Conference on Credit Scoring and Credit Control.

18. Zheng, L., Liu, G., Luan, W., Li, Z., Zhang, Y., Yan, C., et al. (2018). A new credit card fraud detecting method based on behavior certificate. In IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai (pp. 1–6).

19. Міскевич, О., Багнюк, Н., Христинець, Н., & Марчевська, О. (2020). Автоматизація виявлення

бракованої продукції методами машинного навчання. Комп'ютерно-інтегровані технології: освіта, наука, виробництво , (39), 175-180.

References

1. Report on card fraud in 2020 and 2021. European Central Bank. URL: http://surl.li/lycns (date of access: 08.10.2023).

2. Лян П. Credit Card Fraud Detection: A Note to CCFD Analysis. https://medium.com/codex/credit-card-fraud-detection-a-concise-note-to-ccfd-analysis-6ad521f5366d (дата звернення: 07.11.2023).

3. Credit Card Fraud Detection. URL: http://surl.li/fmdgz (date of access: 13.11.2023).

4. Top 10 Credit Card Fraud Detection Solutions in 2023. CybeReady. URL: https://cybeready.com/comprehensive-guide-to-fraud-detection-management-and-analysis/top-10-credit-card-fraud-detection-solutions-in-2023 (date of access: 04.03.2024).

5. Building A Strong Defense: Strengthening Compliance With Transaction Monitoring Rules. Financial Crime Academy. URL: http://surl.li/rgujo (date of access: 06.03.2024).

6. Narayana, M. S., Prasad, B. V. V. S., Srividhya, A., & Reddy, K. P. R. (2011). Data mining machine learning techniques–A study on abnormal anomaly detection system. International Journal of Computer Science and Telecommunications, 2(6).

7. Rao, K. H., Srinivas, G., Damodhar, A., & Krishna, M. V. (2011). Implementation of anomaly detection technique using machine learning algorithms. International journal of computer science and telecommunications, 2(3), 25-31.

8. Mokhtari, S., Abbaspour, A., Yen, K. K., & Sargolzaei, A. (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. Electronics, 10(4), 407.

9. Tertychnyi, P., Godgildieva, M., Dumas, M., & Ollikainen, M. (2022). Time-aware and interpretable predictive monitoring system for Anti-Money Laundering. Machine Learning with Applications, 8, 100306.

10. Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., & Kuruwitaarachchi, N. (2019, January). Real-time credit card fraud detection using machine learning. In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 488-493). IEEE.

11. Mittal, S., & Tyagi, S. (2020). Computational techniques for real-time credit card fraud detection. Handbook of Computer Networks and Cyber Security: Principles and Paradigms, 653-681.

12. Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised machine learning algorithms for credit card fraud detection: a comparison. In 2020 10th international conference on cloud computing, data science & engineering (confluence) (pp. 680-683). IEEE.

13. Ghosh, R. (1994). Credit card fraud detection with a neural-network. In Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, Wailea, HI (pp. 621–630).

14. Gyamfi, N. K., & Abdulai, J. D. (2018, November). Bank fraud detection using support vector machine. In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 37-41). IEEE.

15. Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018, March). Random forest for credit card fraud detection. In 2018 IEEE 15th international conference on networking, sensing and control (ICNSC) (pp. 1-6). IEEE.

16. Yee, O. S., Sagadevan, S., & Malim, N. H. A. H. (2018). Credit card fraud detection using machine learning as data mining technique. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(1-4), 23-27.

17. Bolton, R. J., & Hand, D. J. (2001) Unsupervised profiling methods for fraud detection. In Conference on Credit Scoring and Credit Control.

18. Zheng, L., Liu, G., Luan, W., Li, Z., Zhang, Y., Yan, C., et al. (2018). A new credit card fraud detecting method based on behavior certificate. In IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai (pp. 1–6).

19. Miskevych, O., Bahniuk, N., Khrystynets, N., & Marchevska, O. (2020). Avtomatyzatsiia vyiavlennia brakovanoi produktsii metodamy mashynnoho navchannia. Kompiuterno-intehrovani tekhnolohii: osvita, nauka, vyrobnytstvo , (39), 175-18