НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
май–июнь 2025 Том 25 № 3 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
May–June 2025 Vol. 25 No 3 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

doi: 10.17586/2226-1494-2025-25-3-466-474

Detecting fraud activities in financial transactions using SMOTENN model

Irfan Syamsuddin[1]✉, Sirajuddin Omsa[2], Andi Rustam[3], Dahsan Hasan[4]

[1,2,4] Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia
[3] Universitas Muhammadiyah Makassar, Makassar 90221, Indonesia

[1] irfans@poliupg.ac.id✉, https://orcid.org/0000-0002-6017-7364
[2] sirajud_om@poliupg.ac.id, https://orcid.org/0000-0003-2776-1531
[3] a.rustam@unismuh.ac.id, https://orcid.org/0009-0000-2082-4339
[4] dahsan.hasan@poliupg.ac.id, https://orcid.org/0009-0002-6612-0104

Abstract

The financial industry plays an important role in national economic growth. Because of their critical function, banks have become prime targets for numerous financial crimes. Among these, fraudulent financial transactions are regarded as a severe issue in the financial industry. Conventional approaches are frequently criticized for being ineffective in dealing with fraud in finance; therefore, machine learning approaches have a potential answer to deal with this problem. The goal of this research is to introduce a novel SMOTENN model to establish early detection of cyber fraudulent activities in financial transactions accurately. Two methods are used in this study: first, the Neural Network algorithm is applied to a dataset that contains unbalanced classes; second, the dataset is balanced using the SMOTE (Synthetic Minority Over-sampling Technique) algorithm first, followed by the Neural Network algorithm which we refer to as SMOTENN. The both models are assessed using evaluation metrics of Area Under the Curve, F1-score, precision, recall, specificity, accuracy, and processing time. The comparative analysis shows that the performance of the new SMOTENN model with a balanced dataset is significantly better than that of the neural network approach with an imbalanced dataset, implying that the new SMOTENN model is effective in detecting fraud activities in financial transactions.

Keywords

financial industry, bank, financial fraud, imbalanced dataset, SMOTE, neural network

Acknowledgements

The author would like to thank supports from Politeknik Negeri Ujung Pandang, Indonesia and Universitas Muhammadiyah Makassar, Indonesia.

For citation: Syamsuddin I., Omsa S., Rustam A., Hasan D. Detecting fraud activities in financial transactions using SMOTENN model. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 3, pp. 466–474. doi: 10.17586/2226-1494-2025-25-3-466-474

УДК 004.89

Обнаружение мошенничества при финансовых транзакциях с использованием модели SMOTENN

Ирфан Сямсуддин[1]✉, Сираджуддин Омса[2], Анди Рустам[3], Дахсан Хасан[4]

[1,2,4] Государственный политехнический институт Уджунг Панданга, Макассар, 90245, Индонезия
[3] Университет Мухаммадии в Макассаре, Макассар, 90221, Индонезия

[1] irfans@poliupg.ac.id✉, https://orcid.org/0000-0002-6017-7364
[2] sirajud_om@poliupg.ac.id, https://orcid.org/0000-0003-2776-1531
[3] a.rustam@unismuh.ac.id, https://orcid.org/0009-0000-2082-4339
[4] dahsan.hasan@poliupg.ac.id, https://orcid.org/0009-0002-6612-0104

Аннотация

Финансовая индустрия играет важную роль в национальном экономическом росте. Из-за своей критической функции банки стали главными целями для многочисленных финансовых преступлений. Среди них

мошеннические финансовые транзакции считаются серьезной проблемой в финансовой индустрии. Традиционные подходы часто критикуются за неэффективность в борьбе с мошенничеством в сфере финансов; поэтому подходы машинного обучения имеют потенциальный ответ для решения этой проблемы. Целью данного исследования является внедрение новой модели Synthetic Minority Over-sampling Technique с Neural Network (SMOTENN) для точного раннего обнаружения кибермошеннических действий в финансовых транзакциях. В работе используются два метода: алгоритм нейронной сети применяется к набору данных, содержащему несбалансированные классы; набор данных первоначально балансируется с помощью алгоритма SMOTE, а затем алгоритма нейронной сети SMOTENN. Обе модели оцениваются с использованием метрик Area Under the Curve, F1-score, точность, полнота, специфичность, достоверность и время обработки. Сравнительный анализ показывает, что эффективность новой модели SMOTENN со сбалансированным набором данных значительно выше, чем у подхода на основе нейронной сети с несбалансированным набором данных. Это свидетельствует об эффективности новой модели SMOTENN в обнаружении мошеннических действий в финансовых транзакциях.

Ключевые слова
финансовая индустрия, банк, финансовое мошенничество, несбалансированный набор данных, SMOTE, нейронная сеть

Благодарности
Автор благодарит за поддержку Политехнический институт Негери Уджунг Панданг, Индонезия, и Университет Мухаммадии Макассар, Индонезия.

Ссылка для цитирования: Сямсуддин И., Омса С., Рустам А., Хасан Д. Обнаружение мошенничества при финансовых транзакциях с использованием модели SMOTENN // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 3. С. 466–474 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-3-466-474

Introduction

The worldwide financial industry is critical in enabling transactions, managing investments, and assisting firms and individuals with their sophisticated operations. However, the rising sophistication of fraudulent actions has infused the need for powerful fraud detection techniques within the intricacies of this interconnected ecosystem [1]. Financial institutions are continually threatened by a wide range of fraudulent actions, ranging from identity theft to complex cyber attacks.

The potential for fraud to disrupt the financial sector smooth operation has grown significantly in the last decades. In an interconnected world where financial transactions happen at the speed of light, the possibility of fraud wreaking havoc has increased, demanding proactive and advanced fraud detection systems [1, 2].

The significant financial losses experienced by both institutions and their clients are one of the key reasons for the urgency in fraud detection in the financial industry. Fraudulent acts not only deplete financial institutions assets, but also destroy clients' trust in these institutions [3]. The consequences go beyond monetary losses, impacting financial institutions reputation and trustworthiness, and ultimately compromising the stability of the entire financial system. As a result, the requirement for real-time and predictive fraud detection systems has become critical [3–5].

Furthermore, since the global financial system is interrelated, fraud in one region of the world may have repercussions that across international borders. As a result, any solutions that improve the collective resilience of the whole financial ecosystem against the transnational character of financial fraud are urgently needed to identify and trace patterns of suspicious behavior in global financial industries [6].

Machine Learning (ML) emerges as a highly viable method to address the issue. The expanding financial crime scene necessitates a dynamic and adaptive approach to fraud detection, and traditional methods fall short of modern sophistication [7]. ML algorithms, which are powered by artificial intelligence, can analyze massive volumes of data at previously unheard-of rates, discovering patterns and abnormalities that may signal fraudulent behavior. This adaptive technology keeps financial institutions one step ahead of fraudsters by constantly learning and developing to combat emerging threats [7, 8]. In turn, financial policy makers may obtain correct and quick information regarding any potential fraud in the future.

On one hand, the effectiveness of applying ML algorithms to combat financial fraud is highly depending on the availability of appropriate dataset. The so called imbalanced dataset is prune to results over-fitting or inaccurate results due to very small number of fraud data in comparison to normal data [9].

This phenomenon occurs when the number of transactions is significantly more than the number of fraudulent transactions. This will have a negative impact on the performance of ML models [10]. The so called imbalance data commonly found financial industry in which significant imbalance data between normal and abnormal transactions occurs in the dataset, which is the main concern of our study.

The objective of this study is to present a new ML model based on Synthetic Minority Over-sampling Technique (SMOTE) technique in combination with Neural Network algorithm. An experimental study was also given to compare imbalance dataset and balance dataset after applying SMOTE technique and followed by measuring their performance based on evaluation metrics. It is also important to note that, the study specifically uses Orange data mining software. It is an open source data mining software with many ML algorithms features and it has intuitive graphical user interface [11, 12] suitable for whole analysis in the study in order to showcase the advantages of visual data mining software for non-technical or science audiences.

The whole processes are conducted within Orange data mining software, and open source data mining software with intuitive graphical user interface [11, 12].

Literature Review

The banking sector is a fundamental pillar in supporting global economic dynamics, providing a vital contribution to the circulation of funds and the growth of the business sector. However, the sustainability of this role is faced with serious challenges, especially related to the increase in incidents of fraud in financial transactions. According to Indonesia Fraud Report in 2019 [13], the level of financial crime in Indonesia has increased significantly, posing a serious threat to the stability of the financial sector.

One of the most detrimental forms of threat is financial fraud which includes various activities, such as identity theft, credit card fraud, and transaction manipulation. According to study by Ikbal et.al [14], Indonesia experienced a jump in financial fraud cases in the past year, highlighting the need for a new and effective approach to fighting financial crime.

The importance of involving cutting-edge technology in efforts to prevent banking fraud is becoming increasingly urgent. Reference [15] shows that the application of ML can be a leading solution in mitigating financial risks. ML is able to process and analyze large data volumes quickly, recognize suspicious transaction patterns, and provide early detection of fraudulent activities. Thus, this research puts forward the hypothesis that the integration of ML in banking security systems will pave the way towards more efficient protection of customer finances [16].

The importance of applying ML approach is emphasized in [17] considering bank fraud case is rising significantly, which ends in many damages for the banks. In addition, understanding the underlying patterns in the dataset is important for detecting fraud effectively.

Real time application of ML is studied in [18]. They also realize that datasets in such domain are mostly imbalance. Therefore, techniques to deal with imbalance problems, such as SMOTE, might improve the performance of the classification model after making balance labels for detecting fraud cases in banking sector [19–21].

Another research on imbalance dataset in financial industry is presented in [22]. There is a problem of overfitting in ML results by using datasets that are not directly proportional. Making all labels in appropriately balanced is important for further ML analysis [23]. Using the SMOTE technique, the quality of the dataset processed through ML is very good, so it is recommended to balance data that is still not proportional before further processing.

In [24] it is stated that when the fraud ratio is very small compared to the normal ratio, it is very difficult for classification algorithms to detect fraud cases validly. For this reason, experts must carefully pay attention to this imbalance problem from the beginning of processing the dataset. Apart from that, if imbalanced data is not handled optimally, it is very likely that false positive detection errors will occur, especially if the data is very unbalanced. In fact, in some cases, false warnings are generated, which can also be detrimental to financial institutions.

According to [25], traditional statistical approaches are not suitable to deal with imbalanced data usually occur in bank institutions. Therefore, they propose a model for credit default prediction by employing various credit-related datasets. In addition by considering a significant difference between the minimum and maximum values in different features, they employ Min-Max normalization to scale the features into specific range. Using this approach, imbalance dataset could be handled and solved.

Considering previous studies, the issue of imbalanced dataset in financial industry is still a serious problem. Therefore, handling the issue of imbalance dataset has become the objective in this study through the application of oversampling technique and then applying the results to ML algorithm.

Methodology

The research methodology in conducting this study consists of five steps. First of all, it begins with dataset collection. The dataset collected for the study is Bank Marketing Dataset from UCI Machine Learning Repository website[1].

The second step is pre-processing the dataset. Although the dataset has been validated in the study by Moro [26], cleaning the data, solving missing data or other possible errors are still important step to do in any ML analysis in order to ensure readiness of the dataset for further analysis.

Then, once the dataset are clearly free from any errors, we can go through previewing the dataset class or target. In this step, proportion between target classes will be previewed clearly in order to identify the existing of class imbalance as the problem to be tackled in the following steps.

In the fourth step, the dataset will be applied firstly to Neural Network algorithm directly. Then, subsequently the dataset will firstly undergo class balancing using SMOTE technique before Neural Network algorithm. SMOTE technique is a method to handle imbalance dataset by increasing the number of minority class to be equal with majority class [19–21].

Neural Networks is an artificial intelligence algorithms that analyses large datasets by mimicking the functionality of an animal brain such as emulating the synaptic and neural connections found in biological systems [27, 28]. The selection of Neural Network in Orange is based on the requirement of the study to apply algorithm with ability to deal with nonlinear relationships as found in the dataset, in addition to Neural Network powerful capabilities for tasks, such as classification, regression, and pattern recognition. According to the Orange documentation, type of Neural Network used by Orange is a Multi-Layer Perceptron (MLP) algorithm with back propagation[2]. By having the

[1] Bank Marketing Dataset. UCI Repository. Available at: https://archive.ics.uci.edu/dataset/222/bank+marketing (accessed: 01.09.2024).

[2] Neural Network. Orange Documentation. Available at: https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/model/neuralnetwork.html (accessed: 12.02.2025).

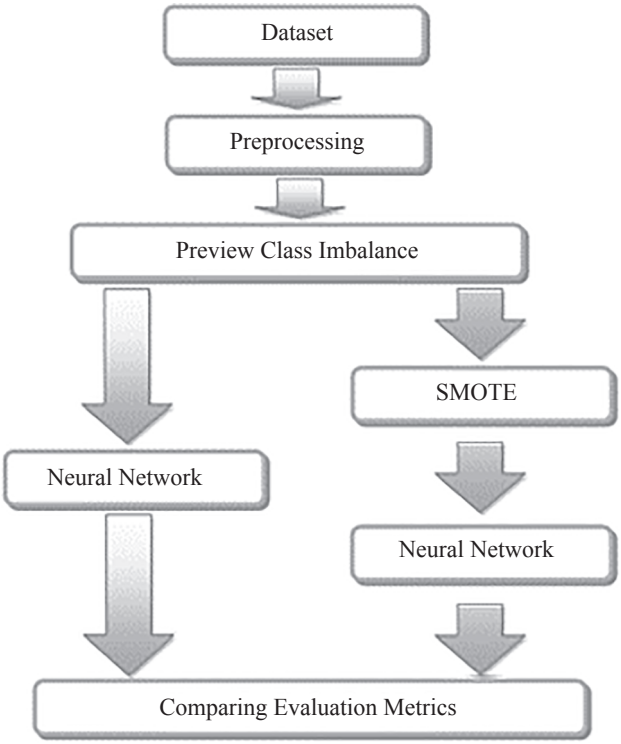


Fig. 1. Research method

basic structure of a Neural Network, the MLP algorithm uses back propagation for increasing the model accuracy which is crucial in such fraud detection case [29].

Finally, both models (Neural Network and SMOTE Neural Network) are evaluated to compare their performance according to accuracy, precision, recall, specifity, Area Under the Curve (AUC), and F1-score. Fig. 1 illustrates the whole steps of research methodology.

Results and Discussion

Dataset collection

This research uses a dataset from an open source or public website (UCI ML Repository: Bank Marketing Data Set, n.d.) with a file size of 3,664 kilobytes. The dataset represents many financial transactions data from a Portuguese banking institution. The data has two labels, namely normal and fraud. Normal label represents normal financial transaction while fraud label means fraudulent financial transactions.

The dataset (Table 1) takes original data obtained from May 2008 to June 2013 with a total of 45,211 rows of data. The dataset has been corrected by undergoing a pre-processing stage to select the data that is really needed so that the data is not too large for research [26].

Data Pre-processing

The study is conducted using Orange Data Mining software installed on Windows 11 machine of AMD A10-9600P RADEON R5, 10 COMPUTE CORES 4C + 6G 2.40 GHz. In Orange, the pre-processing step could be done by applying the Pre-Processing widget which represents the whole pre-processing such as data cleaning and transformation.

The results give a clean and error free dataset as depicted in Fig. 2. It is mentioned that there is no more missing data, 16 features, and a target with 2 values (fraud and normal) and so on. In this clear form, the dataset will be ready for further ML analysis in the next steps.

ML Application with Imbalance Dataset

In the case of fraud detection, we put our concern to the column 17 of the dataset which is Class as the target for analysis with two categories, normal transaction and fraud transaction. Fig. 3 shows the proportion of both class in which there is a significant difference between normal transactions and fraud transactions. This phenomenon is

Table 1. Description of the Dataset

Feature	Type	Description
Age	Numerical	Age
Job	Categorical	Job
Marital	Numerical	Marital status
Education	Categorical	Highest education
Default	Categorical	Credit card status
Balance	Numerical	Amount of transaction
Housing	Categorical	Housing loan status
Loan	Categorical	Bank loan status
Contact	Categorical	Contact
Day	Categorical	Last weekly call log
Month	Categorical	Last monthly call log
Duration	Numerical	Duration of last call
Campaign	Numerical	Number of Bank Campaigns
Pdays	Numerical	Number of Days after Campaigns Campaign
Previous	Numerical	Number of Contact before Campaigns
Poutcome	Categorical	Campaign result
Class	Categorical	Fraud indication (*target*)

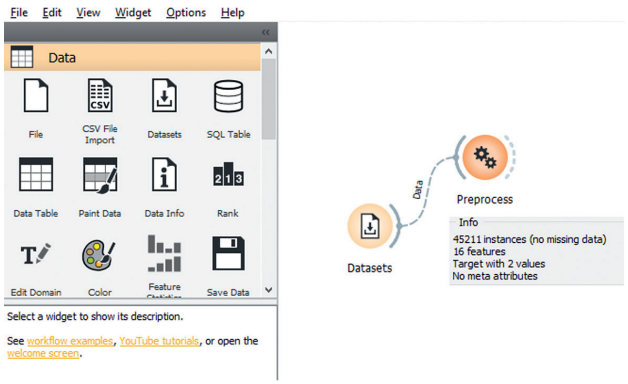
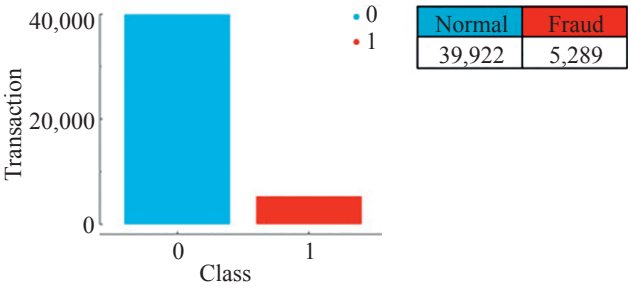


Fig. 2. Pre-processing function in Orange



Normal	Fraud
39,922	5,289

Fig. 3. Imbalanced dataset

commonly found in several cases and is commonly known as imbalance issue of dataset.

Fig. 4 shows that process of ML application in Orange data mining. It begins with loading the dataset, and connect it with ML widget which representing Neural Network algorithm. Furthermore, we perform model evaluation to assess accuracy, precision, recall, specifity, AUC, and F1-score of the applied model.

The Orange Neural Network utilizes sklearn library MLP algorithm by default setting as follows, using Rectified Linear Unit (ReLu) activation, using Adaptive Moment Estimation (Adam) solver, using 100 neurons in hidden layer and 200 iterations. In this step, this default setting is applied to the Imbalanced Dataset as represented in Fig. 4.

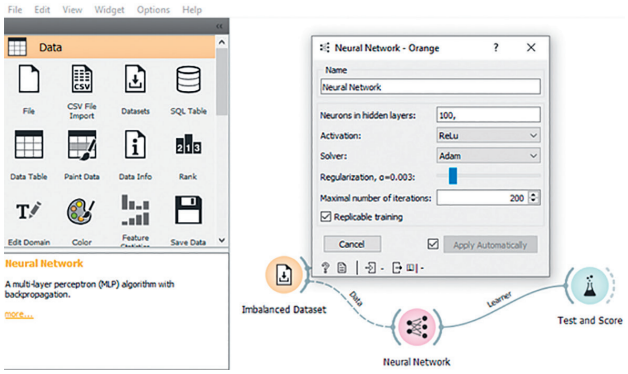


Fig. 4. Implementation of Neural Network on imbalance dataset

Furthermore evaluation is performed using Test and Score widget, we obtain Table 2 which represents the results of evaluation metrics for Neural Network algorithm on the given dataset. These results indicate a robust binary classification model with excellent discrimination (AUC = 90.6 %) and a strong balance between correctly identifying positive instances and avoiding false positives (F1-score = 89.5 %, precision = 89.1 %, recall = 90 %), leading to a high overall accuracy of 90 %. However, only 51.2 % of specificity indicates a notable weakness in correctly identifying negative instances, which is due to class imbalance of the dataset. In addition, the processing time of 1.07 seconds suggests reasonable efficiency of the model.

Implementation of SMOTE

At this step, the dataset that displays imbalance proportion between legitimate and fraudulent financial activities will be corrected by implementing an oversampling technique. As mentioned previously, making the proportional dataset is the objective of oversampling.

The process of oversampling is done in Orange data mining software. The Python code of SMOTE algorithm is supplied to the widget of Python Script in Orange as depicted in Fig. 5.

By using the script feature in Orange software, the SMOTE code in Python could be supplied to perform oversampling process for minority labels. The process is carried out in several steps by adding new fraud data to the dataset repeatedly until a balanced value is obtained. Finally, we obtain a proportional percentage between labels (0 represents normal data, while 1 represents fraud data) as seen in Fig. 6. This new dataset is further called SMOTE Balanced Dataset.

ML Application with Balance Dataset

Previous step has produced a new dataset from the application of SMOTE algorithm to the default dataset called SMOTE Balanced Dataset. In this step, we apply the Neural Network algorithm to the SMOTE Balanced Dataset in Orange data mining.

Similar to the first analysis with imbalance dataset, we apply the default setting of Orange Neural Network as follows by using ReLu activation, Adam solver, 100 neurons in hidden layer and 200 iterations. In this step, this default setting is applied to the SMOTE Balanced Dataset as shown in Fig. 7.

Based on the approach, we produce a new model named SMOTENN which stands for SMOTE Neural Network model and is a combination of SMOTE algorithm and Neural Network. Furthermore, assessment on the SMOTENN model is also conducted to evaluate its accuracy, precision, recall, specifity, AUC, and F1-score for comparison purpose. Table 3 represents the results of

Table 2. Metrics evaluation of Neural Network on imbalance dataset, %

Model	AUC	F1-score	Precision	Recall	Specifity	Accuracy	Processing time, s
Neural Network	90.6	89.5	89.1	90	51.2	90	1.07

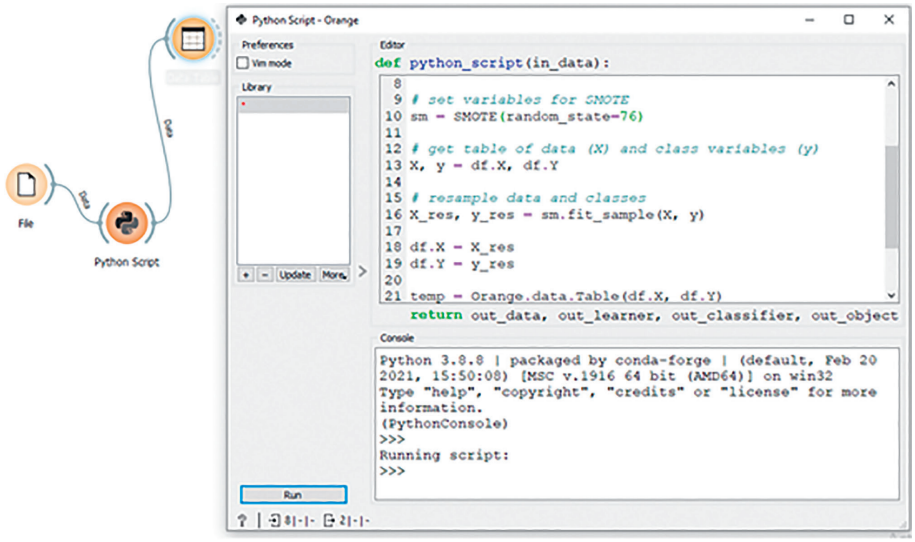


Fig. 5. SMOTE process in Orange

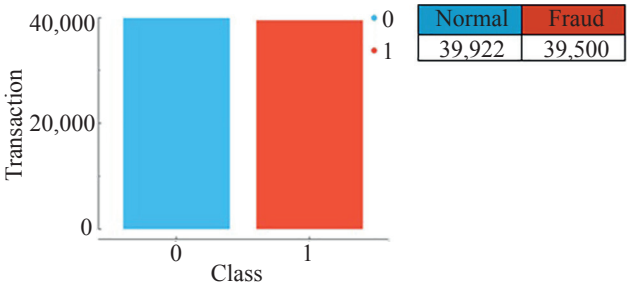


Fig. 6. Balanced dataset after SMOTE

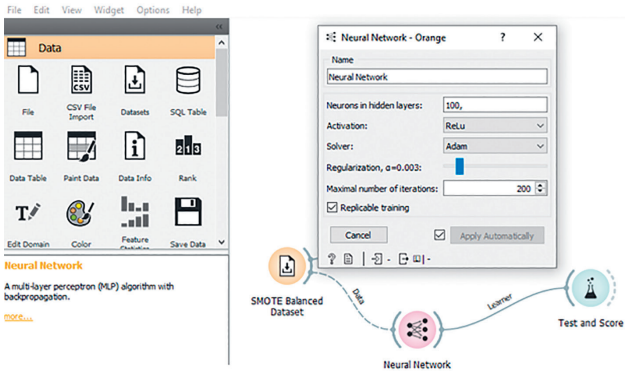


Fig. 7. Implementation of Neural Network on SMOTE balanced dataset

evaluation metrics for SMOTENN model. These results clearly demonstrate a high-performing binary classification model, characterized by excellent discrimination (AUC = 95.7 %) and a strong, well-balanced ability to correctly classify both positive and negative instances, as evidenced by the consistently high F1-score (92.4 %), precision (92.7 %), and recall (92.4 %). A significant improvement to 92.4 % on specifity is potentially due to class balance impact after applying SMOTE technique. Then, increased accuracy to 92.4 % suggesting effective classification across both groups. Finally, the processing time of 1.44 seconds also imply efficient performance of the model.

Comparative Analysis

In this section, we conduct comparative analysis of the results between the first approach (using imbalance dataset) and the second one using SMOTE based balance dataset.

First of all, we conducted the Kolmogorov–Smirnov (KS) test to compare the distributions of the two datasets, before and after the application of SMOTE (Fig. 8). We obtained KS statistics of 0.3 and a P-value of 0.786. The KS statistics of 0.3 indicates a moderate difference between the two datasets, but the high P-value of 0.7869 shows this difference is not statistically significant as results and suggests the datasets are likely from the same distribution.

So, Fig. 8 illustrates the detail comparison between Neural Network models performance with and without SMOTE across six evaluation metrics, namely AUC, F1-score, Precision, Recall, Specificity, Accuracy. Overall, the SMOTENN model demonstrates consistent improvement in all metrics in comparison to the first model, with detailed comparisons as follows:

From AUC perspective, before applying SMOTE, the AUC was 0.906. After applying SMOTE, it increased to 0.957, representing a relative improvement of approximately 5.62 %. This highlights better discrimination between classes in the new SMOTE model.

Based on F1-score perspective, its F1-score accounts for 0.895, then it rose to 0.924 with SMOTE, reflecting a 3.24 % increase. This indicates enhanced balance between precision and recall in the new SMOTENN model.

Table 3. Metrics evaluation of SMOTE with Neural Network on balance dataset, %

Model	AUC	F1-score	Precision	Recall	Specifity	Accuracy	Processing Time, s
SMOTE Neural Network	95.7	92.4	92.7	92.4	92.4	92.4	1.44

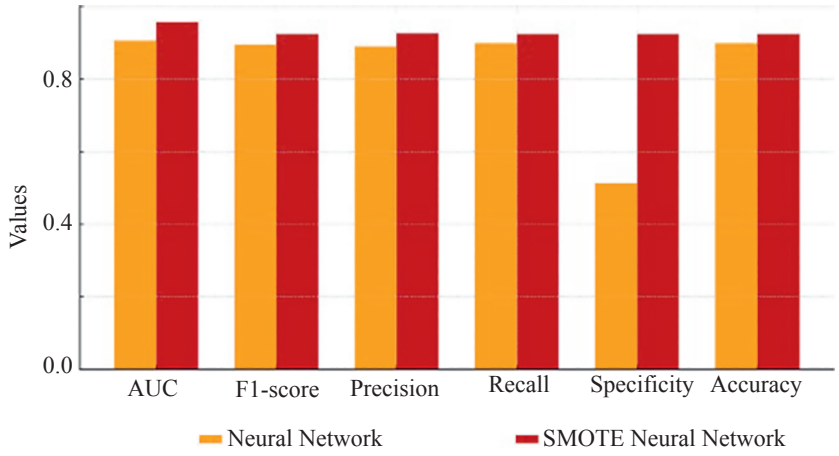


Fig. 8. Overall comparison metrics of Neural Network and SMOTENN

According to Precision metric, it is found that the Precision value climbed from 0.891 to 0.927, marking a 4.04 % improvement. This shows the SMOTENN model increased accuracy in correctly predicting positive instances.

Based on Recall metric, we see the Recall metric before SMOTE was 0.900 and after applying SMOTE it become 0.924. It is a 2.67 % increase for new SMOTENN model enhanced ability to capture positive cases.

Specificity shows the most significant improvement among other metrics. Before applying SMOTE, specificity was only 0.512. After SMOTE, it surged to 0.924, representing a dramatic improvement of approximately 80.47 %. This significant increase underscores SMOTENN model effectiveness in identifying fraudulent activities.

Then the accuracy aspect also shows significant rise. Before, the accuracy was only 0.900, then it jumps to 0.924 or gives a modest 2.67 % increase after applying SMOTE in tackling imbalance dataset. This specific metric result indicates better overall performance for our new SMOTENN model in detecting fraudulent activities in financial transactions.

Lastly, there is a slight increase of processing time before and after SMOTE implementation. The processing time required for SMOTENN model increased slightly, from 1.07 s to 1.44 s, reflecting an expected computational cost due to the augmented dataset size after SMOTE application. However, this difference, a mere 0.37 s, is negligible and does not pose a significant issue, particularly when weighed against the benefits SMOTE offers in addressing class imbalance.

Overall, these results confirm that our new SMOTENN model is effective to be applied in financial fraud detection where minority class identification is crucial.

In the future, next research could explore under-sampling method for balancing the dataset, or applying other hybrid approaches such as ensemble methods or cost-sensitive learning to further enhance these improvements while maintaining robustness across diverse datasets.

Conclusion

Frauds in financial institutions are considered serious problems by the whole financial industry. Sophisticated technologies must be incorporated to deal with the problem since traditional approaches are incapable of catching up with such emerging crimes. Therefore, artificial intelligence approaches such as Machine Learning (ML) are relevant to be proposed in dealing with fraud issues in the recent interconnected financial industry. The availability of historical dataset is prominent to enable ML techniques, showing off their capabilities. However, the proportion of fraud data is often very minimal in comparison to normal data of financial transactions. As a result, an imbalanced dataset issue occurs, which results in poor performance of ML.

In order to accurately identify any possible fraudulent financial activities, this study suggests a new ML model, namely SMOTENN. The model is integration between the Synthetic Minority Over-sampling Technique (SMOTE) dealing with imbalanced dataset and the Neural Network algorithm, a supervised ML algorithm. The new SMOTENN model clarifies that by fixing the imbalanced dataset, the performance could be significantly improved as seen in all evaluation metrics, particularly the specificity metric which accounted for a sharp 80.47 % increase.

References

1. Hidajat T. Rural banks fraud: a story from Indonesia. *Journal of Financial Crime*, 2020, vol. 27, no. 3, pp. 933–943. https://doi.org/10.1108/jfc-01-2020-0010
2. Syahria R. Detecting financial statement fraud using fraud diamond (A study on banking companies listed on the Indonesia Stock Exchange period 2012-2016). *Asia Pacific Fraud Journal*, 2019, vol. 4, no. 2, pp. 183–190. https://doi.org/10.21532/apfjournal.v4i2.114
3. Nabila P., Omsa S., Bravely I. Financial performance analysis of KSP Berkat Bulukumba in year 2018-2022. *AKUNSIKA: Jurnal Akuntansi dan Keuangan*, 2025, vol. 6, no. 1, pp. 10–20. https://doi.org/10.31963/akunsika.v6i1.4902
4. Tjambolang T.A., Radjab E., Hamid A. Financial technology dan gaya hidup dalam perilaku keuangan mahasiswa Politeknik Negeri Ujung Pandang. *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*, 2023, vol. 9, no. 2, pp. 355–362. (in Indonesian)
5. Omsa S., Hasan D., Bravely I., Suryadi A.S., Ischika A.P. Peningkatan kinerja keuangan "Anisah Catering" melalui peningkatan manajemen usaha mikro. *Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV)*, 2023, vol. 9, no. 3, pp. 66–73. (in Indonesian)
6. Lim K.S., Lee L.H., Sim Y.W. A review of machine learning algorithms for fraud detection in credit card transaction. *International Journal of Computer Science & Network Security*, 2021, vol. 21, no. 9, pp. 31–40.
7. Ali A., Abd Razak S., Othman S.H., Eisa T.A.E., Al-Dhaqm A., Nasser M., Elhassan T., Elshafie H., Saif A. Financial fraud detection based on machine learning: a systematic literature review. *Applied Sciences*, 2022, vol. 12, no. 19, pp. 9637. https://doi.org/10.3390/app12199637
8. Syamsuddin I., Hwang J. The application of AHP model to guide decision makers: a case study of e-banking security. *Proc. of the Fourth International Conference on Computer Sciences and Convergence Information Technology*, 2009, pp. 1469–1473. https://doi.org/10.1109/iccit.2009.251
9. Gupta A., Lohani M.C., Manchanda M. Financial fraud detection using naive bayes algorithm in highly imbalance data set. *Journal of Discrete Mathematical Sciences and Cryptography*, 2021, vol. 24, no. 5, pp. 1559–1572. https://doi.org/10.1080/09720529.2021.1969733
10. El-Naby A., Hemdan E.E.D., El-Sayed A. An efficient fraud detection framework with credit card imbalanced data in financial services. *Multimedia Tools and Applications*, 2023, vol. 82, no. 3, pp. 4139–4160. https://doi.org/10.1007/s11042-022-13434-6
11. Tebala D., Marino D. Companies and artificial intelligence: an example of clustering with orange. *Studies in Systems, Decision and Control*, 2023, vol. 222, pp. 1–12. https://doi.org/10.1007/978-3-031-33461-0_1
12. Raqib F., Dunne M., Gurney J., Harle D.E., Sivapalan T., Sabokbar N., Bhogal-Bhamra G.K. Translational learning with orange data mining. *Proc. of the 11th International Conference on Research Advancement Resilience in the Pandemic Era: A Drive for Innovative Transformation*, 2021.
13. Ariyanto R., Bone H. Fraud awareness in Indonesian governmental sector: Multi-agency responses. *Review of Integrative Business and Economics Research*, 2020, vol. 9, no. 2, pp. 209–222.
14. Ikbal M., Irwansyah I., Paminto A., Ulfah Y., Darma D.C. Financial intelligence: Financial statement fraud in Indonesia. *Journal of Intelligence Studies in Business*, 2020, vol. 10, no. 3, pp. 80–95. https://doi.org/10.37380/jisib.v10i3.640
15. Taneja S., Suri B., Kothari C. Application of balancing techniques with ensemble approach for credit card fraud detection. *Proc. of the International Conference on Computing, Power and Communication Technologies (GUCON)*, 2019, pp. 753–758.
16. Saragih M.G., Chin J., Setyawasih R., Nguyen P.T., Shankar K. Machine learning methods for analysis fraud credit card transaction. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2019, vol. 8, no. 6 S, pp. 870–874.
17. Hashemi S.K., Mirtaheri S.L., Greco S. Fraud detection in banking data by machine learning techniques. *IEEE Access*, 2023, vol. 11, pp. 3034–3043. https://doi.org/10.1109/ACCESS.2022.3232287
18. Tadvi F., Shinde S., Patil D., Dmello S. Real time credit card fraud detection. *International Research Journal of Engineering and Technology*, 2021, vol. 8, no. 5, pp. 2177–2180.
19. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial*

Литература

1. Hidajat T. Rural banks fraud: a story from Indonesia // Journal of Financial Crime. 2020. V. 27. N 3. P. 933–943. https://doi.org/10.1108/jfc-01-2020-0010
2. Syahria R. Detecting financial statement fraud using fraud diamond (A study on banking companies listed on the Indonesia Stock Exchange period 2012-2016) // Asia Pacific Fraud Journal. 2019. V. 4. N 2. P. 183–190. https://doi.org/10.21532/apfjournal.v4i2.114
3. Nabila P., Omsa S., Bravely I. Financial performance analysis of KSP Berkat Bulukumba in year 2018-2022 // AKUNSIKA: Jurnal Akuntansi dan Keuangan. 2025. V. 6. N 1. P. 10–20. https://doi.org/10.31963/akunsika.v6i1.4902
4. Tjambolang T.A., Radjab E., Hamid A. Financial technology dan gaya hidup dalam perilaku keuangan mahasiswa Politeknik Negeri Ujung Pandang // Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV). 2023. V. 9. N 2. P. 355–362.
5. Omsa S., Hasan D., Bravely I., Suryadi A.S., Ischika A.P. Peningkatan kinerja keuangan "Anisah Catering" melalui peningkatan manajemen usaha mikro // Proc. of the Seminar Nasional Terapan Riset Inovatif (SENTRINOV). 2023. V. 9. N 3. P. 66–73.
6. Lim K.S., Lee L.H., Sim Y.W. A review of machine learning algorithms for fraud detection in credit card transaction // International Journal of Computer Science & Network Security. 2021. V. 21. N 9. P. 31–40.
7. Ali A., Abd Razak S., Othman S.H., Eisa T.A.E., Al-Dhaqm A., Nasser M., Elhassan T., Elshafie H., Saif A. Financial fraud detection based on machine learning: a systematic literature review // Applied Sciences. 2022. V. 12. N 19. P. 9637. https://doi.org/10.3390/app12199637
8. Syamsuddin I., Hwang J. The application of AHP model to guide decision makers: a case study of e-banking security // Proc. of the Fourth International Conference on Computer Sciences and Convergence Information Technology. 2009. P. 1469–1473. https://doi.org/10.1109/iccit.2009.251
9. Gupta A., Lohani M.C., Manchanda M. Financial fraud detection using naive bayes algorithm in highly imbalance data set // Journal of Discrete Mathematical Sciences and Cryptography. 2021. V. 24. N 5. P. 1559–1572. https://doi.org/10.1080/09720529.2021.1969733
10. El-Naby A., Hemdan E.E.D., El-Sayed A. An efficient fraud detection framework with credit card imbalanced data in financial services // Multimedia Tools and Applications. 2023. V. 82. N 3. P. 4139–4160. https://doi.org/10.1007/s11042-022-13434-6
11. Tebala D., Marino D. Companies and artificial intelligence: an example of clustering with orange // Studies in Systems, Decision and Control. 2023. V. 222. P. 1–12. https://doi.org/10.1007/978-3-031-33461-0_1
12. Raqib F., Dunne M., Gurney J., Harle D.E., Sivapalan T., Sabokbar N., Bhogal-Bhamra G.K. Translational learning with orange data mining // Proc. of the 11th International Conference on Research Advancement Resilience in the Pandemic Era: A Drive for Innovative Transformation. 2021.
13. Ariyanto R., Bone H. Fraud awareness in Indonesian governmental sector: Multi-agency responses // Review of Integrative Business and Economics Research. 2020. V. 9. N 2. P. 209–222.
14. Ikbal M., Irwansyah I., Paminto A., Ulfah Y., Darma D.C. Financial intelligence: Financial statement fraud in Indonesia // Journal of Intelligence Studies in Business. 2020. V. 10. N 3. P. 80–95. https://doi.org/10.37380/jisib.v10i3.640
15. Taneja S., Suri B., Kothari C. Application of balancing techniques with ensemble approach for credit card fraud detection // Proc. of the International Conference on Computing, Power and Communication Technologies (GUCON). 2019. P. 753–758.
16. Saragih M.G., Chin J., Setyawasih R., Nguyen P.T., Shankar K. Machine learning methods for analysis fraud credit card transaction // International Journal of Engineering and Advanced Technology (IJEAT). 2019. V. 8. N 6 S. P. 870–874.
17. Hashemi S.K., Mirtaheri S.L., Greco S. Fraud detection in banking data by machine learning techniques // IEEE Access. 2023. V. 11. P. 3034–3043. https://doi.org/10.1109/ACCESS.2022.3232287
18. Tadvi F., Shinde S., Patil D., Dmello S. Real time credit card fraud detection // International Research Journal of Engineering and Technology. 2021. V. 8. N 5. P. 2177–2180.
19. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. SMOTE: synthetic minority over-sampling technique // Journal of Artificial Intelligence Research. 2002. V. 16. P. 321–357. https://doi.org/10.1613/jair.953

Intelligence Research, 2002, vol. 16, pp. 321–357. https://doi.org/10.1613/jair.953
20. Fernández A., Garcia S., Herrera F., Chawla N.V. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, 2018, vol. 61, pp. 863–905. https://doi.org/10.1613/jair.1.11192
21. Elreedy D., Atiya A.F., Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, 2024, vol. 113, no. 7, pp. 4903–4923. https://doi.org/10.1007/s10994-022-06296-4
22. Veigas K.C., Regulagadda D.S., Kokatnoor S.A. Optimized stacking ensemble (OSE) for credit card fraud detection using synthetic minority oversampling model. *Indian Journal of Science and Technology*, 2021, vol. 14, no. 32, pp. 2607–2615. https://doi.org/10.17485/ijst/v14i32.807
23. Syamsuddin I., Barukab O.M. SUKRY: Suricata IDS with enhanced kNN algorithm on Raspberry Pi for classifying IoT botnet attacks. *Electronics*, 2022, vol. 11, no. 5, pp. 737. https://doi.org/10.3390/electronics11050737
24. Makki S., Assaghir Z., Taher Y., Haque R., Hacid M.S., Zeineddine H. An experimental study with imbalanced classification approaches for credit card fraud detection. *IEEE Access*, 2019, vol. 7, pp. 93010–93022. https://doi.org/10.1109/ACCESS.2019.2927266
25. Alam T.M., Shaukat K., Hameed I.A., Luo S.H., Sarwar M.U., Shabbir S., Li J.M., Khushi M. An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access*, 2020, vol. 8, pp. 201173–201198. https://doi.org/10.1109/access.2020.3033784
26. Moro S., Cortez P., Rita P. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 2014, vol. 62, pp. 22–31. https://doi.org/10.1016/j.dss.2014.03.001
27. Kawam A.A., Mansour N. Metaheuristic optimization algorithms for training artificial neural networks. *International Journal of Computer and Information Technology*, 2012, vol. 1, no. 2, pp. 156–161.
28. Quiroga F.M. Invariance and same-equivariance measures for convolutional neural networks. *CLEI Electronic Journal*, 2021, vol. 24, no. 1 SI. https://doi.org/10.19153/cleiej.24.1.8
29. Alla H., Moumoun L., Balouki Y. A multilayer perceptron neural network with selective-data training for flight arrival delay prediction. *Scientific Programming*, 2021, vol. 2021, pp. 5558918. https://doi.org/10.1155/2021/5558918

20. Fernández A., Garcia S., Herrera F., Chawla N.V. SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary // Journal of Artificial Intelligence Research. 2018. V. 61. P. 863–905. https://doi.org/10.1613/jair.1.11192
21. Elreedy D., Atiya A.F., Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning // Machine Learning. 2024. V. 113. N 7. P. 4903–4923. https://doi.org/10.1007/s10994-022-06296-4
22. Veigas K.C., Regulagadda D.S., Kokatnoor S.A. Optimized stacking ensemble (OSE) for credit card fraud detection using synthetic minority oversampling model // Indian Journal of Science and Technology. 2021. V. 14. N 32. P. 2607–2615. https://doi.org/10.17485/ijst/v14i32.807
23. Syamsuddin I., Barukab O.M. SUKRY: Suricata IDS with enhanced kNN algorithm on Raspberry Pi for classifying IoT botnet attacks // Electronics. 2022. V. 11. N 5. P. 737. https://doi.org/10.3390/electronics11050737
24. Makki S., Assaghir Z., Taher Y., Haque R., Hacid M.S., Zeineddine H. An experimental study with imbalanced classification approaches for credit card fraud detection // IEEE Access. 2019. V. 7. P. 93010–93022. https://doi.org/10.1109/ACCESS.2019.2927266
25. Alam T.M., Shaukat K., Hameed I.A., Luo S.H., Sarwar M.U., Shabbir S., Li J.M., Khushi M. An investigation of credit card default prediction in the imbalanced datasets // IEEE Access. 2020. V. 8. P. 201173–201198. https://doi.org/10.1109/access.2020.3033784
26. Moro S., Cortez P., Rita P. A data-driven approach to predict the success of bank telemarketing // Decision Support Systems. 2014. V. 62. P. 22–31. https://doi.org/10.1016/j.dss.2014.03.001
27. Kawam A.A., Mansour N. Metaheuristic optimization algorithms for training artificial neural networks // International Journal of Computer and Information Technology. 2012. V. 1. N 2. P. 156–161.
28. Quiroga F.M. Invariance and same-equivariance measures for convolutional neural networks // CLEI Electronic Journal. 2021. V. 24. N 1 SI. https://doi.org/10.19153/cleiej.24.1.8
29. Alla H., Moumoun L., Balouki Y. A multilayer perceptron neural network with selective-data training for flight arrival delay prediction // Scientific Programming. 2021. V. 2021. P. 5558918. https://doi.org/10.1155/2021/5558918

Authors

Irfan Syamsuddin — Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, sc 25927526500, https://orcid.org/0000-0002-6017-7364, irfans@poliupg.ac.id

Sirajuddin Omsa — Associate Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, https://orcid.org/0000-0003-2776-1531, sirajud_om@poliupg.ac.id
Andi Rustam — Associate Professor, Universitas Muhammadiyah Makassar, Makassar, 90221, Indonesia, sc 25927526500, https://orcid.org/0009-0000-2082-4339, a.rustam@unismuh.ac.id
Dahsan Hasan — Associate Professor, Politeknik Negeri Ujung Pandang, Makassar, 90245, Indonesia, https://orcid.org/0009-0002-6612-0104, dahsan.hasan@poliupg.ac.id

Авторы

Сямсуддин Ирфан — PhD, профессор, Государственный политехнический институт Уджунг Панданга, Макассар, 90245, Индонезия, sc 25927526500, https://orcid.org/0000-0002-6017-7364, irfans@poliupg.ac.id
Омса Сираджуддин — PhD, доцент, Государственный политехнический институт Уджунг Панданга, Макассар, 90245, Индонезия, https://orcid.org/0000-0003-2776-1531, sirajud_om@poliupg.ac.id
Рустам Андри — доцент, Университет Мухаммадии в Макассаре, Макассар, 90221, Индонезия, sc 25927526500, https://orcid.org/0000-0002-6017-7364, a.rustam@unismuh.ac.id
Хасан Дахсан — PhD, доцент, Государственный политехнический институт Уджунг Панданга, Макассар, 90245, Индонезия, https://orcid.org/0009-0002-6612-0104, dahsan.hasan@poliupg.ac.id

Received 12.01.2025
Approved after reviewing 02.05.2025
Accepted 24.05.2025

Статья поступила в редакцию 12.01.2025
Одобрена после рецензирования 02.05.2025
Принята к печати 24.05.2025