Enhancing Credit Card Fraud Detection through Adversarial Learning: A Generative Adversarial Network Approach

Tushar Malankar
Dept. of AI & DS,
Thadomal Shahani EngineeringCollege, Bandra, Mumbai

Omkar Pathare
Dept. of AI & DS,
Thadomal Shahani EngineeringCollege, Bandra, Mumbai

Yash Shah
Dept. of AI & DS,
Thadomal Shahani EngineeringCollege, Bandra, Mumbai

Bhushan Jadhav, PhD
Dept. of AI & DS,
Thadomal Shahani EngineeringCollege, Bandra, Mumbai

ABSTRACT

In the fast-evolving landscape of credit card fraud detection, the imperative to swiftly and accurately identify fraudulent transactions has become paramount. Class imbalance poses a significant challenge to traditional methods, hindering their adaptability to the intricate nature of fraud. This study investigates the efficacy of Synthetic Minority Over-sampling Technique (SMOTE) and Generative Adversarial Networks (GANs) in mitigating class imbalance. Utilizing a dataset of 1.75 million transactions, both methods generate balanced datasets, evaluated through Isolation Forest models. Despite suboptimal GAN training, the GAN-generated dataset closely aligns with SMOTE in performance metrics. Precision, recall, F1-score, and an overall accuracy of 0.51 for GAN and 0.54 for SMOTE reveal competitive results, indicating the promising potential of GANs in handling class imbalance for anomaly detection in credit card transactions. The study emphasizes the significance of advanced generative techniques for improved model performance and robust handling of imbalanced data scenarios.

Keywords

Credit card fraud detection, class imbalance, anomaly detection, Synthetic Minority Over-sampling Technique (SMOTE), Generative Adversarial Networks (GANs), Isolation Forest, machine learning.

1. INTRODUCTION

In the fast-paced digital era, combatting credit card fraud has become increasingly challenging, particularly with the surge in online transactions and the evolving sophistication of fraudulent activities. Swift and accurate detection of these fraudulent transactions is crucial to safeguard financial systems and protect users from identity theft, account takeover, and other illicit financial schemes.

A significant challenge faced by credit fraud detection systems is the class imbalance problem, where the majority of transactions being legitimate makes it difficult for traditional methods to discern the subtle signs of fraud. This challenge hampers the adaptability of rule-based systems and the effectiveness of statistical models in understanding the intricate nature of fraudulent transactions.

To address the class imbalance issue, prevailing methods often resort to techniques such as the Synthetic Minority Over- sampling Technique (SMOTE). SMOTE generates synthetic instances of the minority class to balance the dataset, but it may fall short of capturing the complex patterns inherent in fraudulent behavior.

In our methodology, Synthetic Minority Over-sampling Technique (SMOTE) and Generative Adversarial Networks (GANs) are employed separately to address the class imbalance challenge in credit card fraud detection. The process involves generating balanced datasets independently using each method. Subsequently, both the SMOTE-enhanced dataset and the GAN-generated dataset are individually fitted to an Isolation Forest model.

The results from the Isolation Forest models applied to these separate datasets are then compared to evaluate the effectiveness of each approach. The intention is to analyze how well each method contributes to anomaly detection in the context of credit card transactions.

The advantage of GANs lies in their ability to learn and mimic the complex data distribution of the minority class. Unlike SMOTE, GANs are inherently generative and can create synthetic instances by understanding the underlying features of fraudulent behavior. By training a GAN on the minority class, our approach aims to produce synthetic instances that not only mitigate class imbalance but also enrich the dataset with more realistic representations. GANs introduce a level of sophistication in the generated instances, making them less susceptible to overfitting and potentially better aligned with the complex nature of credit card fraud.

2. LITERATURE REVIEW:

The escalating threat of credit card fraud in the wake of technological advancements, including the widespread adoption of contactless payment methods, underscores the urgency for robust detection and prediction mechanisms. In this comprehensive review spanning the years 2015 to 2021, the authors examine 40 relevant articles to gain insights into cutting-edge research on credit card fraud detection and prediction. The literature is categorized based on topics such as the class imbalance problem and feature engineering, as well as the machine learning technologies employed, encompassing traditional and deep learning approaches. The analysis reveals a limited exploration of deep learning techniques, emphasizing the need for further research to effectively tackle the challenges associated with emerging technologies like big data analytics, large-scale machine learning, and cloud computing. By raising pertinent research issues and delineating future directions, this study serves as a valuable resource for both academic and industrial researchers, offering guidance in the evaluation of financial fraud detection systems and the design of resilient solutions.[1]

Al-Hashedi and Magalingam's (2021) research delves into a comprehensive analysis of financial fraud research papers spanning the years 2009 to 2019. The focus of the study revolves around the application of data mining techniques, categorizing the literature based on various factors, including publication year, publisher, methodology, and research domain (such as credit fraud, cryptocurrency fraud, insurance, and general financial fraud). The paper provides a detailed overview of works dedicated to detecting financial fraud, encompassing credit card fraud, insurance fraud, and other fraudulent activities. The utilization of data mining techniques for fraud detection is elucidated, along with the specification of datasets and validation metrics.
However, it is noteworthy that the review exclusively concentrates on 'classification' techniques, omitting a comprehensive exploration of the entire detection chain. This aspect is crucial for understanding the features employed in the detection process. Additionally, the authors did not specifically address credit card fraud and overlooked issues related to class imbalance and feature engineering. Notably, the study lacked consideration for emerging trends, such as big data. In summary, while providing valuable insights into the detection of financial fraud, the review could benefit from a broader perspective encompassing the complete detection chain and addressing contemporary challenges and trends.[2]

In their investigation, Cui et al. (2021) devised an innovative anomaly detection mechanism geared towards the identification of online banking fraud. The system was purposefully designed to address limitations observed in previous fraud detection systems, specifically concerning insufficient historical user data, skewed transaction data, and the absence of standardized methods for handling user attribute values. A novel ranking metric embedding system named ReMEMBeR was introduced to leverage multi-contextual behavior, aiming to alleviate issues related to false positives and errors. To address challenges related to historical user data, the authors innovatively introduced a pseudo recommender system, treating individuals as pseudo-users and their behaviors as pseudo items. Collaborative filtering was incorporated to harness insights from similar user behaviors. The ranking methodology (legitimate/fraudulent) involved mapping pseudo-users to pseudo items. The system's effectiveness was rigorously evaluated across four dimensions, encompassing real-world transactions, skewed data, model combination (involving machine learning algorithms such as SVM, RF, NN2L, logR), and diverse contextual behaviors.[3]

Lucas et al. (2019) proposed a credit card fraud detection method centered on quantifying covariate shift, specifically differences in user behaviors. The classification efficiency of transactions from each day against those from other days was used to identify a covariate shift when inefficiency was observed. The authors employed a distance matrix to characterize data, followed by clustering using the RF algorithm. The resulting clusters revealed distinctions corresponding to working days, school holidays, Saturdays, and Sundays. Integrating covariate shift as a new feature in the credit card fraud detection system proved beneficial in addressing variations in user purchase behaviors and fraudulent mechanisms over time. However, the study utilized a relatively dated dataset from 2015, potentially overlooking recent fraudulent strategies.[4]

Ingole et al. (2021) employed unsupervised models, including the Isolation Forest and Local Outlier Factor, to tackle credit card fraud detection as an outlier detection problem. The Isolation Forest, akin to the random forest model, facilitated outlier detection by building upon decision trees. The Local Outlier Factor calculated an anomaly score, measuring the isolation of a sample from its neighbors. This distinctive approach provided a nuanced solution to credit card fraud detection, taking into account the unique characteristics of each model.[5]]

The pervasive challenge of class imbalance combined with overlap in electronic fraud transaction detection provides a severe barrier, as fraudsters deliberately craft fraudulent transactions to closely resemble real ones, making them difficult to identify. While existing literature focuses on class imbalance in machine-learning-based fraud transaction detection algorithms, the issue of transaction overlap has received little attention. Using a divide-and-conquer method, this study proposes a unique hybrid technique to addressing the complexities of class imbalance with overlap. To begin, an anomaly detection model is trained on minority samples in order to eliminate outliers from the minority class as well as a substantial portion of the majority samples from the original dataset. The remaining samples, constituting an overlapping subset, exhibit a reduced imbalance ratio and diminished interference from both minority and majority classes compared to the original dataset. Subsequently, this challenging overlapping subset undergoes classification using a non-linear classifier to effectively distinguish between genuine and fraudulent transactions. The paper introduces a novel assessment criterion, Dynamic Weighted Entropy (DWE), tailored to evaluate the quality of the overlapping subset, serving as a trade-off between the exclusion of minority class outliers and the imbalance ratio of the subset. The proposed method, aided by DWE, demonstrates superior performance in terms of efficiency and accuracy, as validated through extensive experiments on Kaggle fraud detection dataset and a large real electronic transaction dataset, outperforming state-of-the-art approaches.[6]

Several investigations have delved into machine learning approaches for credit card fraud detection. Khatri et al. [7] explored diverse algorithms, pinpointing kNN as the most effective. Rajora et al. [10] concentrated on RF and kNN, noting the high accuracy achieved by RF. Trivedi et al. [8] proposed an efficient engine using GB and RF, both demonstrating notable accuracy and precision. Tanouz et al. [9] addressed class imbalance through under-sampling, emphasizing the necessity for further exploration of this challenge. Riffi et al. implemented ELM and MLP, underscoring MLP's superiority. Randhawa et al. introduced AdaBoost-SVM, showcasing impressive accuracy and MCC. These collective studies contribute to the comprehension of credit card fraud detection, with potential implications for integrating SMOTE to contend with class imbalance complexities.

SMOTE stands as a crucial method within machine learning to tackle the challenge of class imbalance in datasets. This technique involves creating synthetic instances for the minority class by interpolating between existing minority class instances. By doing so, SMOTE aims to alleviate the impact of skewed class distribution, thereby enhancing the effectiveness of classification models through the provision of a more balanced dataset.

In addressing credit card fraud detection challenges, researchers have explored supervised learning methods, spanning machine learning and deep learning. To mitigate the impact of imbalanced credit card data on classification outcomes, strategies include enhancing classifiers, selecting more effective classification modes, and addressing imbalanced data directly. Studies comparing different machine learning algorithms, such as Logistic Regression, C5.0 decision tree, and Support Vector Machine for classification, have been conducted. Artificial neural networks, exemplified by Asha RB, have been employed to predict fraud occurrences, demonstrating improved accuracy compared to unsupervised learning. Additionally, oversampling and undersampling methods, particularly Synthetic Minority Over-sampling Technique (SMOTE), have become prevalent for ensemble classification models, although challenges like information loss have been noted [10] [11]. This study adopts the

VAEGAN model as an oversampling module, leveraging synthetic minority class data to rebalance the training set while enhancing the model's expressive capabilities for more realistic and diverse data outputs

2.1 Isolation Forest:

Isolation Forest is a data anomaly detection algorithm introduced by Fei Tony Liu in 2008. It operates using binary trees, exhibiting linear time complexity and low memory requirements, making it suitable for high-volume data. The algorithm focuses on anomalies' characteristics, utilizing random splits in the data space. Unlike decision tree algorithms, Isolation Forest relies solely on path-length measures to generate anomaly scores, eliminating the need for leaf node statistics.

Isolation Forest achieves speed through random data space splits using randomly selected attributes and split points. The anomaly score inversely correlates with path-length, as anomalies require fewer splits due to their uniqueness. The algorithm capitalizes on the concept that anomalies are easily separable, constructing Isolation Trees iteratively through recursive partitioning. Each tree isolates anomalies efficiently, where anomalies correspond to shorter path lengths. The Isolation Tree's probabilistic interpretation is detailed in the original Isolation Forest paper.

2.2 Generative Adversarial Network (GAN)

GAN is a powerful machine learning model that employs two neural networks - the generator and the discriminator - in a competitive learning process. GANs adopt a zero-sum game strategy while operating unsupervised, where one network's development comes at the price of the other.

GANs use a continuous feedback loop between the generator and discriminator, which are both neural networks. The generator creates synthetic data that attempts to mimic genuine data, whereas the discriminator distinguishes between authentic and produced data. This adversarial process evolves throughout training iterations, refining the generator's ability and the discriminator's discernment.

2.3 Working of GAN:

The generator and discriminator in a GAN's complicated relationship always attempt to surpass one other. The generator generates synthetic data from random noise, developing its output to seem like real data. Simultaneously, the discriminator improves its ability to differentiate between actual and fake data. This dynamic interaction, which is guided by a loss function, leads to a convergence in which the generator generates very convincing synthetic data. The GAN design strives for seamless merging, blurring the lines between the imaginary and the real.

2.4 Architecture of GAN:

The two essential components of GAN's are the creative generator and the discriminating discriminator. The generator initiates the creative process by creating synthetic data out of randomness. The discriminator assesses both genuine and synthetic data at the same time, honing its capacity to discriminate between the two. This constant feedback loop enables the generator to adapt in response to the discriminator's insights, resulting in a fusion in which created data blurs the lines between the real and the imaginary.

3. METHODOLOGY:

The methodology for the proposed work consists of various phases including Input selection, Initial data analysis, Data Preprocessing, Class imbalance mitigation etc. which are explained in following subsections. The different stages in proposed framework are shown in Fig.1.

3.1 Input Selection:

The dataset comprises 1.75 million transactions generated from simulated users across diverse terminals spanning from January 2023 to June 2023. Notably, the dataset exhibits a substantial imbalance, with a mere 0.1345% of transactions falling within the fraudulent classification. This disparity in class distribution necessitates careful consideration and specialized techniques for handling imbalanced data in the context of fraud detection research.

3.2 Initial data analysis:

The dataset contains 1,754,155 entries and 10 columns, including TRANSACTION_ID, TX_DATETIME, CUSTOMER_ID, TERMINAL_ID, TX_AMOUNT, TX_TIME_SECONDS, TX_TIME_DAYS, TX_FRAUD, and TX_FRAUD_SCENARIO. The Unnamed: 0 column appears to be an index. The TX_DATETIME column is currently stored as an object and might require conversion for temporal analysis. Additionally, the TX_FRAUD column indicates the presence of fraud in transactions, with 1,518,186 non- fraudulent instances and 235,969 instances flagged as fraudulent

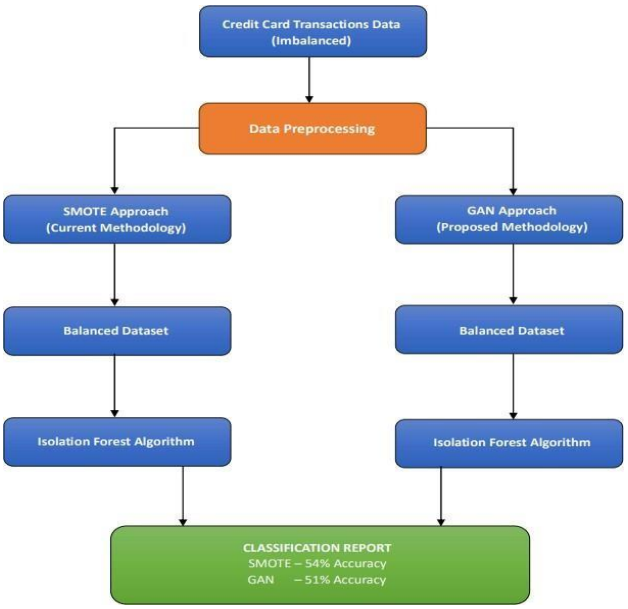

Fig 1: Proposed framework

3.3 Data Preprocessing:

In the preprocessing stage, several columns, including TERMINAL_ID, CUSTOMER_ID, TX_DATETIME, and TX_FRAUD_SCENARIO, were dropped from the dataset as they were deemed either redundant or not directly contributing to the model's predictive capabilities. TERMINAL_ID and CUSTOMER_ID might have been excluded due to their potential limited impact on fraud detection, while TX_DATETIME and TX_FRAUD_SCENARIO may have been removed for simplicity or after extracting relevant features from them.

Following the feature selection, the dataset was divided into training and testing sets using a standard train-test split. The TX_FRAUD column, denoting the occurrence of fraud in transactions, was designated as the target variable, while the remaining columns were treated as input features.

3.4 Class imbalance mitigation:

In our methodology, we address the challenge of class imbalance through separate implementations of two distinct approaches: Generative Adversarial Networks (GANs) and the traditional Synthetic Minority Over-sampling Technique (SMOTE). Firstly, we employ GANs, utilizing a generator-discriminator architecture to synthetically generate minority class instances, enriching the dataset with diverse representations. Additionally, we independently implement the SMOTE algorithm to oversample the minority class instances through synthetic data interpolation. By maintaining a clear separation between these approaches, we aim to systematically assess and compare the individual impacts of GANs and SMOTE in mitigating class imbalance.

3.4.1 Generative Adversarial Network Model (GAN):

The implemented Generative Adversarial Network (GAN) comprises three key components: the generator, discriminator, and the combined GAN model. The generator is constructed with two dense layers, having 128 neurons each and employing the rectified linear unit (ReLU) activation function for non-linearity. It is designed to take random noise as input (with a dimension specified by latent_dim) and output synthetic data with a dimensionality matching that of the original dataset features. On the other hand, the discriminator, responsible for distinguishing between real and synthetic data, consists of two dense layers with 128 neurons each and utilizes the sigmoid activation function. The GAN model shown in Fig. 2 is formed by combining the generator and discriminator, with the discriminator's trainable parameter set to False to facilitate the training of the generator. The models are then compiled with the binary cross entropy loss function and the Adam optimizer.

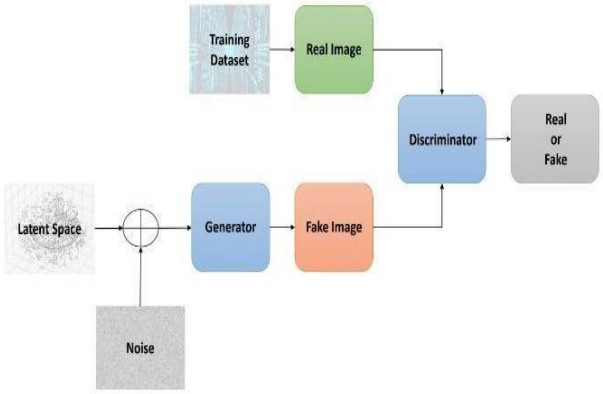

Fig 2: GAN Architecture

3.4.1.1 MODEL TRAINING

The training process of the GAN unfolds over a specified number of epochs, where at each iteration, a random batch of real data from the minority class is selected. This real data batch is used to train the discriminator, which aims to correctly classify real instances as genuine (label: 1) and synthetic instances as generated (label: 0). Simultaneously, the generator produces synthetic data by inputting random noise, attempting to fool the discriminator into misclassifying it as real. The adversarial training involves optimizing the discriminator and generator iteratively, updating their weights to enhance the ability of the generator to generate realistic data. The training progress is printed at regular intervals, displaying the discriminator and generator losses. This iterative process continues until the specified number of epochs is reached.

3.4.1.2 SYNTHENTIC DATA GENERATION

Once the GAN training completes, the generator is used to produce synthetic data. The number of synthetic samples equals the number of cases in the original dataset that are not fraudulent. The model generates synthetic data instances by feeding random noise into the generator. This synthetic dataset is thus ready for comparison and assessment, giving a fair representation of the minority class for training machine learning models, which is notably useful for correcting class imbalance in fraud detection. The created synthetic data is used as an augmentation approach to help the model detect fraudulent transactions with greater efficiency.

3.4.2 SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) comes with the imbalanced-learn library. The original feature set 'x' and accompanying labels 'y' are resampled using SMOTE, resulting in an enhanced dataset labelled as X_data and Y_data. SMOTE balances the class distribution by producing synthetic instances of the minority class. This guarantees that both the dominant and minority classes are sufficiently represented, reducing the influence of class imbalance in future model training.

4. RESULTS AND DISCUSSION

The evaluation metrics from the Isolation Forest model applied to datasets generated by both GAN and SMOTE reveal interesting insights into the performance of these approaches in handling class imbalance. Despite the GAN not being trained to its optimal capacity, its results closely approximate those of the SMOTE-enhanced dataset.

Table 1: Actual Result of GAN & Isolated Forest

	PRECISION	RECALL	F1-SCORE	SUPPORT
FRAUD	0.72	0.13	0.22	350949
NON-FRAUD	0.48	0.94	0.64	303520
ACCURACY	**0.51**			

Table 2: Result of smote oversampling & isolated forest

	PRECISION	RECALL	F1-SCORE	SUPPORT
FRAUD	0.69	0.14	0.23	303609
NON-FRAUD	0.52	0.94	0.67	303666
ACCURACY	**0.54**			

As results from Table 1, In the case of the GAN-generated dataset, Class 0 (non- fraudulent instances) shows a precision of 0.48, recall of 0.94, and an F1-score of 0.64. Class 1 (fraudulent instances) exhibits a precision of 0.72, recall of 0.13, and an F1-score of 0.22. The overall accuracy is reported at 0.51. Comparatively, as results from Table 2, The SMOTE-generated dataset shows similar precision, recall, and F1-score for both classes, with an overall accuracy of 0.54. These metrics suggest that, even with suboptimal training, the GAN approach performs competitively and approaches the results achieved by SMOTE.

This convergence in performance indicates the promising potential of GANs in addressing class imbalance, showcasing their ability to generate synthetic instances that closely approximate the impact of traditional oversampling techniques like SMOTE. The similarity in performance, even without full optimization, suggests that with further fine-tuning and training, GANs have the potential to outperform or at least match the efficacy of established techniques like SMOTE in handling imbalanced datasets for anomaly detection. The results underscore the significance of exploring advanced generative techniques in the pursuit of improved model performance and more robust handling of imbalanced data scenarios.

5. CONCLUSION:

In conclusion, this study highlights the enormous potential of Generative Adversarial Networks (GANs) as a compelling alternative to existing oversampling strategies, such as the Synthetic Minority Over-sampling Technique (SMOTE), for resolving class imbalance in credit card fraud detection. Despite inadequate training, GANs exhibit competitive performance, closely resembling SMOTE outcomes. The convergence of precision, recall, and F1-score measures, together with competitive accuracy, demonstrates GANs' resilience in producing synthetic examples that successfully alleviate class imbalance.
This implies that, with additional optimization and training, GANs have the potential to outperform or at least match known methodologies, thus opening the door to improved model performance in anomaly identification for data sets with imbalances. This article suggests for more research into sophisticated generative approaches to strengthen the robustness of fraud detection models in dynamic and difficult circumstances.

6. FUTURE WORK:

Future work for this proposed article could explore several avenues to enhance the performance and capabilities of the Generative Adversarial Network (GAN) in the context of credit card transaction fraud detection:

1. Architectural Refinement:

Examine more complicated GAN designs, such as deep convolutional GANs (DCGANs) or Wasserstein GANs (WGANs), to see how they affect the quality of synthetic transactions created. Experimenting with different generator and discriminator structures may yield better results.

2. Hyperparameter Tuning:

Conduct a thorough hyperparameter tuning procedure to optimize the learning rates, batch sizes, and other generator and discriminator parameters. Fine-tuning these parameters may help to make the training process more reliable and effective.
3. Extensive Loss Functions:

Investigate alternate loss functions that are adapted to the unique issues of fraud detection. This might entail creating unique loss functions or implementing extra goals, such as reducing false negatives in fraud scenarios.

4. Adversarial Training Methods:

Investigate further adversarial training strategies, such as employing reinforcement learning or more complex optimization algorithms, to further stabilize and improve the GAN training process.

5. Ensemble Techniques:

Investigate the advantages of integrating the GAN-based strategy with additional fraud detection models or ensemble approaches. Integrating the capabilities of many models may result in a more robust and accurate overall system.

6. Data Augmentation Techniques:

Create and test a variety of data augmentation options for the synthetic data production process. This might include adding extra characteristics, investigating temporal aspects of transactions, or constructing transaction sequences rather than individual instances.

7. Evaluation of Real-World Datasets:

Extend the assessment to incorporate a larger range of real-world credit card transaction datasets to ensure the model's generalizability and performance across various domains and circumstances.

8. Interpretability and Explainability:

Examine approaches for increasing the interpretability of the GAN model's judgements, revealing how and why specific transactions are classed as real or synthetic. This is critical for establishing trust in the model and facilitating its implementation in real-world applications.

9. Scalability and effectiveness:

Address scaling issues and investigate strategies for making the GAN model more computationally efficient, allowing it to be used in real-time or near-real-time fraud detection systems.

10. Integration of User Feedback:

Integrate user feedback mechanisms into the training process, allowing domain experts to offer feedback on the quality of produced transactions and iteratively refining the model based on practical concerns.

By following these routes, the proposed research work will be able to improve the state-of-the-art in GAN-based fraud detection and contribute to the creation of more effective and dependable financial transaction security solutions.

7. REFERENCES

[1] Cherif, A., Badhib, A., Ammar, H., Alshehri, S., Kalkatawi, M., & Imine, A. (2023). Credit card fraud detection in the era of disruptive technologies: A systematic review. Journal of King Saud University - Computer and Information Sciences, 35(1), 145-174. https://doi.org/10.1016/j.jksuci.2022.11.008.

[2] Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to

2019. *Computer Science Review, 40*, 100402. https://doi.org/10.1016/j.cosrev.2021.100402.

[3] Li, Z., Huang, M., Liu, G., & Jiang, C. (2021). A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection. Expert Systems with Applications, 175, 114750. https://doi.org/10.1016/j.eswa.2021.114750.

[4] Cui, J., Yan, C., Wang, C., 2021. ReMEMBeR: Ranking Metric Embedding-Based Multicontextual Behavior Profiling for Online Banking Fraud Detection. IEEE Trans. Comput. Social Syst. (2021) 1–12. Conference Name: IEEE Transactions on Computational Social Systems. https://doi.org/10.1109/TCSS.2021.3052950

[5] Y. Lucas, P.-E. Portier, L. Laporte, S. Calabretto, L. He-Guelton, F. Oblé, M. Granitzer Dataset Shift Quantification for Credit Card Fraud Detection 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE) (2019), pp. 97-100, 10.1109/AIKE.2019.00024

[6] S. Ingole, A. Kumar, D. Prusti, S.K. Rath Service-Based Credit Card Fraud Detection Using Oracle SOA Suite SN Comput. Sci., 2 (3) (2021), p. 161, 10.1007/s42979-021-00539-2

[7] S. Khatri, A. Arora, and A. P. Agrawal, ''Supervised machine learning algorithms for credit card fraud detection: A comparison,'' in Proc. 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence), Jan. 2020, pp. 680–683.

[8] K. Trivedi, S. Simaiya, U. K. Lilhore, and S. K. Sharma, ''An efficient credit card fraud detection model based on machine learning methods,'' Int. J. Adv. Sci. Technol., vol. 29, no. 5, pp. 3414–3424, 2020.

[9] R. Sailusha, V. Gnaneswar, R. Ramesh, and G. R. Rao, ''Credit card fraud detection using machine learning,'' in Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS), May 2020, pp. 967–972.

[10] A. Singh, R. K. Ranjan, and A. Tiwari, ''Credit card fraud detection under extreme imbalanced data: A comparative study of data-level algorithms,'' J. Experim. Theor. Artif. Intell., vol. 34, no. 4, pp. 571–598, Jul. 2022.

[11] N. S. Alfaiz and S. M. Fati, ''Enhanced credit card fraud detection model using machine learning,'' Electronics, vol. 11, no. 4, p. 662, Feb. 2022

[12] F. Z. El Hlouli, J. Riffi, M. A. Mahraz, A. El Yahyaouy, and H. Tairi, ''Credit card fraud detection based on multilayer perceptron and extreme learning machine architectures,'' in Proc. Int. Conf. Intell. Syst. Comput. Vis. (ISCV), Jun. 2020, pp. 1–5.