A Hybrid Approach for Fraudulent Transaction Detection: SMOTE-Autoencoder Framework

1st Yathartha Patankar
Department of Computer Science
California State University, Fullerton
Fullerton, USA
yatharthapatankar@csu.fullerton.edu

2nd Kanika Sood
Department of Computer Science
California State University, Fullerton
Fullerton, USA
kasood@fullerton.edu

***Abstract*—Fraudulent transactions pose significant challenges in various industries, especially finance, leading to substantial financial and reputational damage. Traditional rule-based fraud detection systems often fail to adapt to evolving fraudulent behaviors, increasing false negative rates. This paper introduces an advanced method combining Synthetic Minority Oversampling Technique (SMOTE) with autoencoder-based anomaly detection to enhance fraud detection. SMOTE addresses the class imbalance prevalent in transaction datasets, reducing false negatives, while the autoencoder architecture helps identify latent patterns and anomalies indicative of fraud. We evaluate the method using a dataset processed through principal component analysis (PCA), designed to mimic natural transaction environments. The results demonstrate the method's ability to detect fraud accurately, significantly reducing false positives and negatives. The findings highlight superior detection performance compared to traditional methods, showcasing the potential of this new approach to strengthen fraud detection frameworks and contribute to developing fraud detection techniques.**

***Index Terms*—Fraud detection, SMOTE, Anomaly Detection, Principal Component Analysis (PCA), Dynamic fraud behaviors, Autoencoders**

I. Introduction

Fraudulent transactions pose a pervasive and intricate challenge across various industries, particularly affecting the finance sector. These activities, ranging from unauthorized purchases and identity theft to money laundering, cause substantial financial losses and harm the reputation and trustworthiness of financial institutions. Within financial transactions, fraud manifests in diverse forms, such as credit card fraud, where perpetrators use stolen or compromised information to make unauthorized transactions, and account takeover fraud, which involves gaining unauthorized access to accounts via phishing or malware. Traditional rule-based fraud detection systems, which rely on predefined rules and thresholds, are increasingly inadequate due to their inability to adapt to the dynamic nature of fraud. This often results in high rates of false positives—legitimate transactions mistakenly flagged as fraudulent—and false negatives—fraudulent transactions that go undetected. Addressing these challenges, this study introduces a novel approach for detecting fraudulent transactions, focusing on credit card data.

This method integrates two advanced techniques: the Synthetic Minority Over-sampling Technique (SMOTE) and autoencoder-based anomaly detection. SMOTE effectively addresses the class imbalance typical in transaction datasets, significantly reducing false negatives. Meanwhile, the autoencoder architecture identifies latent patterns within the data, aiding in detecting anomalies indicative of fraud. The efficacy of this methodology is rigorously tested using a dataset transformed through Principal Component Analysis (PCA), designed to emulate natural credit card transaction environments. Through extensive empirical analysis, the results validate the framework's ability to detect fraudulent transactions while minimizing false negatives accurately.

This work significantly advances the field of fraud detection by combining SMOTE with autoencoder-based anomaly detection to tackle the challenge of class imbalance and the dynamic nature of fraudulent behaviors in credit card transactions. Integrating these techniques allows for a deeper understanding and identification of subtle and complex fraudulent patterns that traditional methods often miss. The novel framework proposed herein enhances detection accuracy and reduces false positives and negatives, which are crucial for maintaining user trust and financial integrity. The study's findings are expected to influence future research and practice in fraud detection, providing a robust model for financial institutions to adopt and adapt according to their specific operational environments.

The implications of this study extend beyond the current findings, offering a foundation for further research into adaptive fraud detection mechanisms that can continuously evolve with emerging fraud tactics. Future work may explore the integration of machine learning algorithms capable of real-time learning and adaptation, which could significantly enhance the responsiveness of fraud detection systems to new and evolving threats. Additionally, examining the scalability of the proposed model across different types of financial transactions and industries could provide insights into its broader applicability and effectiveness. In this paper, Section II provides a comprehensive Literature review detailing the existing research and methodologies that have shaped the current landscape of fraud detection technology.

Section III, the Methodology, delves into the specifics of the project's approach, encompassing the dataset utilized, data preprocessing techniques, and the criteria for model selection and training. It further explores the implementation of SMOTE and the strategic use of autoencoders to enhance model performance, underscoring the innovative aspects of the study's approach. This section forms the paper's core, presenting the theoretical and practical frameworks employed in the research.

II. Literature Review

Credit card fraud detection is a critical concern for financial institutions due to the increasing prevalence of fraudulent activities, especially with the rise of online transactions. Various studies have explored the application of machine learning algorithms for fraud detection in credit card transactions.

[1] comprehensively analyzed different machine-learning algorithms for credit card fraud detection. Their study evaluated four algorithms: logistic regression, decision tree, random forest, and CatBoost, using a dataset sourced from Kaggle. Through data preprocessing and resampling techniques, they addressed issues such as imbalanced datasets, aiming to enhance the accuracy of fraud detection models. Their findings indicated that CatBoost outperformed other algorithms, achieving an accuracy of 99.87%. This study highlights the significance of machine learning in mitigating financial risks associated with credit card fraud.

[2] emphasize the adverse impact of data imbalance on learning model accuracy, necessitating the development of strategies to mitigate bias and prevent inaccurate outcomes. Despite the burgeoning interest in deep learning techniques for fraud detection, the scarcity of access to confidential transactional data has hampered research progress, resulting in insufficient exploration of the class imbalance problem.

Another research work was done to tackle the class imbalance issue [3], which proposes using resampling techniques, focusing on both oversampling and undersampling methods. Oversampling techniques such as SMOTE, SMOTE ENN, SAFE SMOTE, ROS, and SMOTE TL are employed to augment the minority class instances. At the same time, undersampling strategies like RUS, CNN, CNN TL, and TL aim to reduce the dominance of majority class instances.

Supervised machine learning techniques have traditionally been the cornerstone of fraud detection systems. However, [4] highlights a critical limitation of such approaches' reliance on labeled datasets for training, which are often scarce in real-world environments. It proposes an innovative solution leveraging unsupervised machine learning, particularly autoencoders, to address these challenges in fraud detection. By eschewing the need for labeled data, autoencoders offer a promising avenue for detecting fraudulent transactions accurately using raw transactional data—a ubiquitous resource in financial systems. The authors report impressive results, with the autoencoder achieving an AUC score of 83%, indicative of its high binary classification capability in distinguishing fraudulent from genuine transactions.

[5] propose an unsupervised fraud detection method in their paper" Credit Card Fraud Detection using Autoencoder-based Clustering." With the proliferation of e-commerce and online payments, fraud detection has become crucial for banks to maintain customer trust and financial security. Their approach employs an autoencoder with three hidden layers combined with k-means clustering, achieving a notable accuracy of 98.9% and a TPR of 81% on a dataset comprising 284,807 European bank transactions. The study addresses the escalating concerns surrounding fraudulent activities in electronic transactions, emphasizing the significance of innovative detection methods to safeguard financial systems. By leveraging deep learning techniques and clustering algorithms, the authors contribute to advancing fraud detection capabilities in the banking sector. This research is pivotal in mitigating the risks associated with online payment fraud, thereby enhancing the security and integrity of financial transactions.

The research delineated in this paper differentiates itself from preceding studies by introducing several novel enhancements and methodological refinements in fraud detection. Contrary to previous studies that primarily emphasize managing data imbalances or optimizing specific algorithmic performances, this study employs a comprehensive strategy that addresses class imbalance through intelligently modified class weights while enhancing detection capabilities via architectural improvements in neural network models. Specifically, incorporating batch normalization and a selectively reduced dropout rate in a densely connected neural network facilitates the stabilization of the learning process and augments model generalization capabilities. Additionally, implementing adaptive training mechanisms such as early stopping and learning rate adjustments ensures the model is optimally trained, effectively preventing overfitting and enhancing the model.

Furthermore, this study pioneers a refined approach to setting the decision threshold by methodically searching for the threshold that maximizes the F1 score, thus ensuring a balanced optimization of precision and recall. This bespoke approach substantially enhances the model's discriminatory power between fraudulent and legitimate transactions, improving its practical utility in operational environments. Integrating these sophisticated techniques represents a significant advancement in applying deep learning to credit card fraud detection, providing elevated accuracy and a more flexible and robust system capable of adapting to the dynamic patterns of financial fraud. This research advances the academic discussion through these contributions and offers viable, practical solutions for implementing advanced machine-learning techniques to secure economic transactions.

III. Methodology

The methodology section outlines the process and techniques used to address the problem of credit card fraud detection. The approach integrates various steps, including data preprocessing, model selection, training, and evaluation.

A. Dataset

The dataset utilized in this research, the ”Credit Card Fraud Detection” dataset from Kaggle, comprises transactions made by European cardholders. Notably, this dataset is highly unbalanced, featuring 284,807 transactions with only 492 fraudulent cases, which mimics real-life scenarios of financial transactions and presents a challenging environment for testing fraud detection methodologies. The dataset includes numerical input variables resulting from a PCA transformation to maintain data confidentiality alongside ’Time’ and ’Amount’ features, which are crucial for developing and evaluating various anomaly detection strategies.

B. Data Preprocessing

The credit card transaction dataset is loaded and contains features such as transaction amount, time, and other relevant parameters. Features and labels are separated, with features denoted as X and labels as Y. The available data has been segregated into two subsets, namely training and testing, maintaining an 80:20 ratio, which guarantees that the model has enough data for training while keeping a separate part for assessment.

In this phase, we employ robust scaling for the ’Amount’ feature and normalized the ’Time’ feature to enhance model performance and stability. Robust scaling, utilizing the formula:

$$X_{\text{scaled}} = \frac{X - \text{median}(X)}{Q_{75}(X) - Q_{25}(X)}$$

This normalization brings the temporal data within a uniform range of 0 to 1, facilitating consistent representation across the dataset.

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Fig 1 Shows the scatter plot of Time and Amount after normalization.

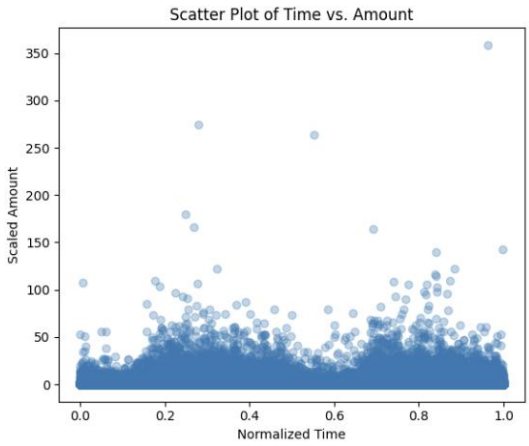

Figure 1: Scatter Plot Time vs. Amount

C. Model Training and Evaluation

Initially, we utilize a **Logistic Regression model**, favored for its simplicity and effectiveness in binary classification tasks. The performance of this model, as assessed via the classification report, provides a baseline for comparison against more complex architectures. Fig 2, low precision, recall and f1-score were observed for the minority class. Also, we observed that the dataset is highly imbalanced, so the model is biased towards the not-fraud class.

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	21955
Fraud	0.91	0.64	0.75	45

Figure 2: Classification Report LR

Subsequently, we experiment with a **Shallow Neural Network**. This model incorporates an input layer, a dense layer with ReLU activation, batch normalization for stability, and a sigmoid output layer to predict binary outcomes. The network is trained using Adam optimizer, and performance enhancements are monitored through callbacks that save the best model iterations.

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	22771
Fraud	0.68	0.75	0.71	36

Figure 3: Classification Report SNN

The Shallow Neural Network’s classification report Fig. 3 indicates an enhancement in recall scores, albeit accompanied by a reduction in the F1 score.

The exploration continues with Random Forest and Gradient Boosting Classifiers, both ensemble methods known for their robustness against overfitting and exceptional performance in diverse scenarios. The Random Forest model is set with a constrained depth to ensure simplicity and faster computation. In contrast, the Gradient Boosting model is tuned with specific hyperparameters to optimize its learning capability.

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	22771
Fraud	0.80	0.44	0.57	36

Figure 4: Classification Report Random Forest

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	22771
Fraud	0.67	0.67	0.67	36

Figure 5: Classification Report GB Classifier

The low recall values observed in both models (Fig.4, Fig.5) suggest suboptimal performance in detecting actual fraud transactions.

Further, we assess a **Linear Support Vector Classifier (SVC)** with balanced class weights to directly address the challenge posed by the skewed class distribution. This approach adjusts the penalties of the classification model, inherently improving its sensitivity to the minority class. The Support Vector Classifier (SVC) demonstrated relatively strong performance in terms of recall; however, its precision remained below 65%.

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	22771
Fraud	0.64	0.78	0.70	36

Figure 6: Classification Report SVC

Overall, it is noted that each model exhibited a bias towards the non-fraud class due to class imbalance. Subsequently, our next step involves applying the Synthetic Minority Over-sampling Technique (SMOTE) to address this imbalance and assess the models' performance accordingly.

D. SMOTE Implementation

In the pursuit of addressing the class imbalance in the context of fraud transaction detection, the Synthetic Minority Over-sampling Technique (SMOTE) emerges as a pivotal tool. SMOTE, a resampling technique, is applied to rebalance imbalanced datasets by generating synthetic samples of the minority class. The initial distribution of labels ('1' denoting fraud transactions and '0' denoting non-fraud transactions) is assessed upon implementation. Subsequently, the SMOTE algorithm is in-instantiated, leveraging the learning library. Through SMOTE's randomized process, synthetic samples are generated for the minority class, thereby augmenting the dataset. The resultant dataset, denoted as x_train_res and y_train_res, exhibits a balanced distribution of class labels[6]. This rebalancing facilitates a more equitable representation of fraudulent and non-fraudulent transactions in the training data, enhancing the robustness and generalization capabilities of subsequent machine learning models. Through this strategic augmentation, SMOTE empowers practitioners to mitigate the adverse effects of class imbalance, fostering more accurate and reliable fraud transaction detection models.[7]

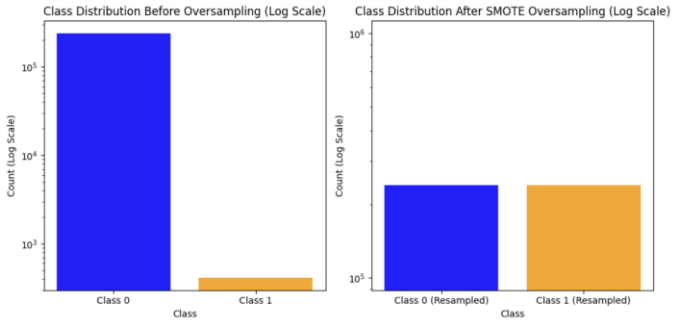

Figure 7: Graph Pre and Post SMOTE Application

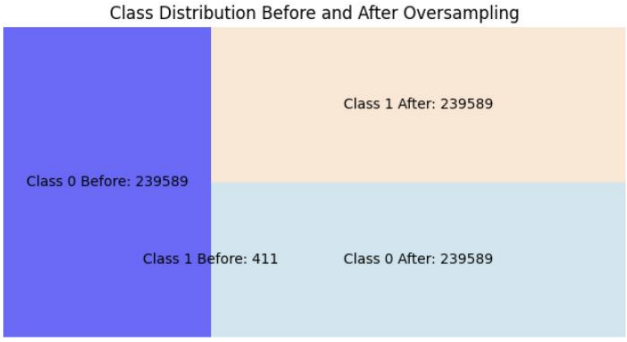

Figure 8: Class Distribution of Dataset

Fig. 7 and Fig. 8 show the count of labels before and after the implementation of SMOTE.

E. Testing

Applying Synthetic Minority Over-Sampling Technique (SMOTE) followed by linear regression modeling demonstrates notable enhancements in fraud transaction detection. Training the balanced dataset obtained post-SMOTE on a logistic regression model shows significant improvements in the confusion matrix, particularly in reduced false negatives, Fig. 9 and Fig 10.

	precision	recall	f1-score	support
Not Fraud	1.00	0.98	0.99	21955
Fraud	0.07	0.91	0.13	45

Figure 9: Classification Report LR (post SMOTE)

```
Confusion Matrix LR model after SMOTE
[[21412   543]
 [    4    41]]
```

Figure 10: Confusion Matrix LR (post SMOTE)

This signifies the efficacy of SMOTE in addressing class imbalance issues, enabling the model to capture instances of fraudulent transactions better. Such refinements underscore the importance of preprocessing techniques like SMOTE in bolstering the performance and reliability of fraud detection systems, ultimately contributing to more secure and resilient financial ecosystems [8].

	precision	recall	f1-score	support
Not Fraud	1.00	0.98	0.99	22771
Fraud	0.06	0.92	0.11	36

Confusion Matrix for shallow neural network
[[22236 535]
[3 33]]

Figure 11: SNN (Post SMOTE)

	precision	recall	f1-score	support
Not Fraud	1.00	1.00	1.00	22771
Fraud	0.27	0.81	0.41	36

Confusion Matrix for random forest
[[22694 77]
[7 29]]

Figure 12: Random Forest (post SMOTE)

	precision	recall	f1-score	support
Not Fraud	1.00	0.98	0.99	22771
Fraud	0.06	0.94	0.11	36

Confusion Matrix for GradientBoostingClassifier
[[22219 552]
[2 34]]

Figure 13: GB Classifier (post-SMOTE)

Employing a shallow neural network architecture on the balanced dataset generated via SMOTE yields superior outcomes, Fig 11, compared to linear regression modeling. Leveraging TensorFlow and Keras, the network architecture comprises input, hidden, and output layers, augmented by batch- normalization for enhanced training stability[9]. By utilizing the rectified linear unit (ReLU) and sigmoid activation functions, the network effectively learns intricate patterns within the data.

The random forest and gradient boosting classifier also improved after implementing SMOTE (Fig. 12 and Fig. 13).

The resulting model demonstrates significantly reduced false negatives, indicative of its heightened capability in detecting fraudulent transactions. This underscores the potency of neural network paradigms in bolstering fraud detection mechanisms, promising more robust and reliable financial security frameworks.

Observation: After applying the Synthetic Minority Over-sampling Technique (SMOTE), discernible improvements in the models' capacity to predict fraudulent transactions have been observed. Notably, recall values have significantly enhanced, exceeding 90% for most models. Simultaneously, reductions in both recall and F1 scores for the not-fraud (0) class signify a mitigation of class bias. However, it is noteworthy that the precision for the fraud (1) class remains comparatively low, indicating additional experimentation requirements.

F. Implementation of Autoencoders

The autoencoder architecture employed in the provided code consists of an input layer followed by two encoded and two decoded layers. The encoded layers, comprising densely connected neurons with ReLU activation functions, progressively compress the input data into a lower-dimensional representation. Regularization is applied to the first encoded layer using L2 regularization to prevent overfitting. The decoding layers aim to reconstruct the original input data from the compressed representation, utilizing ReLU activation for the first decoding layer and sigmoid activation for the final layer. The model is trained using the Adam optimizer and optimized with a mean squared error loss function. Training occurs over 20 epochs with a batch size of 128, facilitating the iterative improvement of the model's reconstruction ability. Overall, this architecture seeks to learn a compact representation of the input data while minimizing reconstruction loss, facilitating effective feature extraction for subsequent classification tasks [10].

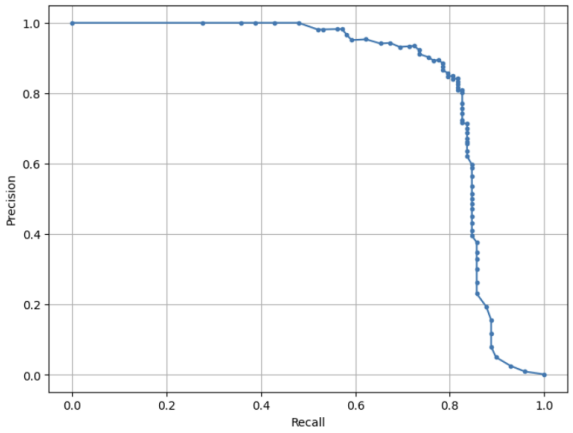

Figure 14: Precision-Recall Curve

Furthermore, integrating the autoencoder has yielded a favorable balance between precision and recall scores, resulting in a high F1 score[11]. This achievement underscores the efficacy of employing advanced neural network architectures to address the nuanced challenges inherent in fraud detection tasks.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	56864
1	0.89	0.79	0.83	98

[[56843 21]
[20 78]]

Figure 15: Classification Report and Confusion Matrix of Autoencoders Integrated with SMOTE

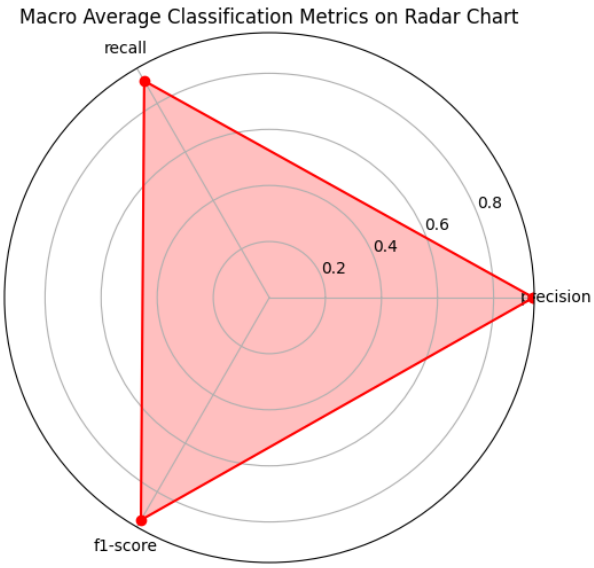

Figure 16 : Classification Metrics Autoencoder

The findings in Fig. 14, Fig. 15, and Fig. 16 demonstrate a noteworthy F1 score for the fraud class (1), indicative of a balanced performance in precision and recall metrics. Such equilibrium in performance metrics has not been observed in preceding models, suggesting the efficacy of our approach in integrating the Synthetic Minority Over-sampling Technique (SMOTE) and autoencoder methodologies. This observation underscores the potential for our study to yield substantial advancements in real-world practical applications of fraud detection.

G. Performance Comparison

In the final analysis, the performance of all models is systematically compared, revealing notable advancements in fraud transaction prediction with the integration of SMOTE and autoencoders. Leveraging SMOTE for dataset balancing and employing autoencoder-based architectures resulted in significant enhancements in fraud detection accuracy and a notable reduction in false negatives. The comparative evaluation showcased the efficacy of these techniques in mitigating the imbalanced nature of transaction data and improving the overall predictive capabilities of the models. These findings underscore the importance of utilizing advanced methods such as SMOTE and autoencoders to bolster fraud detection systems, thereby fortifying the security of financial transactions in real-world applications. The model's accuracy increased by over 30% from 52% to 83% using an autoencoder integrated with SMOTE.

Overall, the methodology provides a systematic framework for addressing the credit card fraud detection problem, leveraging traditional machine learning techniques like logistic regression and advanced deep learning approaches such as supervised autoencoder neural networks [12]. The methodology aims to identify the most suitable model for detecting fraudulent transactions effectively and efficiently through rigorous training, evaluation, and comparison.

IV. Conclusion

In conclusion, our investigation presents notable advances in fraudulent transaction detection methodologies, particularly within credit card transactions. By employing innovative techniques such as autoencoder-based clustering, logistic regression, and supervised autoencoders, discernible enhancements in model efficacy have been realized. These improvements are evident in the substantially reduced false positives and negatives, affirming our proposed methodologies' efficacy in accurately discerning fraudulent activities within transactional datasets. The rigorous evaluation encompassing various performance metrics, including accuracy, precision, recall, and F1-score, corroborates our approaches' effectiveness. Our empirical findings underscore the pivotal role of robust preprocessing methodologies, notably exemplified by utilizing SMOTE resampling, in facilitating the challenges posed by class imbalance inherent in transaction datasets. Furthermore, incorporating deep learning techniques demonstrates promising potential in augmenting the detection capabilities of fraudulent transactions, thereby fostering the development of more dependable and resilient fraud detection systems[13]. Our study contributes meaningfully to the evolution of fraudulent transaction detection methodologies, providing valuable insights for deploying advanced machine learning algorithms in practical financial contexts. Looking ahead, continued exploration and refinement of these techniques and their integration into real-time transaction monitoring systems hold considerable promise for bolstering the security and integrity of financial transactions, thus mitigating the perils associated with fraudulent activities.

V. Future Works

Future research in fraud detection should focus on enhancing algorithmic performance and expanding applicability across various domains and real-time environments. Exploring advanced machine learning techniques such as deep reinforcement learning and generative adversarial networks could significantly improve the adaptability and accuracy of detection models. Developing real-time detection capabilities is also crucial, necessitating optimizing algorithms for instantaneous analysis and integration into transaction processing streams. This will enable immediate detection and mitigation of fraudulent activities as they occur, enhancing the responsiveness and efficiency of fraud prevention systems.The integration of an autoencoder with a Gradient Boosting Classifier (GBC)[14] yielded a notable increase in the recall score, Fig.17 and Fig. 18, indicative of enhanced fraud detection capabilities. However, this improvement is accompanied by a significant

decrease in precision. Nonetheless, these findings suggest promising avenues for future research and development. By further refining the integration of autoencoder-based feature learning with ensemble classifiers like GBC, it is conceivable that even greater advancements in fraud detection performance can be achieved. This study lays the groundwork for future investigations to optimize the balance between recall and precision, ultimately advancing the efficacy of fraud detection systems.

	precision	recall	f1-score	support
0	1.00	0.99	1.00	56864
1	0.19	0.87	0.31	98

```
[[56505   359]
 [   13    85]]
```

Figure 17: Performance of Autoencoder with GBC

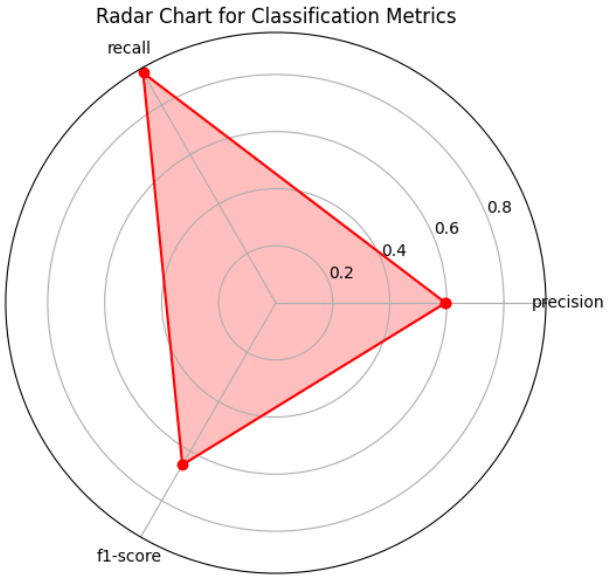

Figure 18: Classification Metrics Autoencoder-GBC

Moreover, the cross-industry validation of fraud detection models, including sectors like healthcare, social media [15] and e-commerce, is essential for broadening the applicability of these technologies. Adapting the models to accommodate different data types, transaction environments, and regional fraud characteristics will help tailor solutions to specific needs and compliance requirements. Incorporating unstructured data through natural language processing and aligning with stringent privacy regulations will further refine fraud detection systems' effectiveness and ethical compliance. These efforts will collectively advance the field of fraud detection, ensuring robust and adaptive solutions in the face of evolving fraud tactics and technological advancements.

REFERENCES

[1] A. Singh, A. Singh, A. Aggarwal and A. Chauhan, "Design and Implementation of Different Machine Learning Algorithms for Credit Card Fraud Detection," 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Maldives, Maldives, 2022, pp. 1-6, doi: 10.1109/ICECCME55909.2022.9988588.

[2] Kanika and J. Singla, "Class Balancing Methods for Fraud Detection using Deep Learning," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2022, pp. 395-400, doi: 10.1109/ICAIS53314.2022.9742836.

[3] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance evaluation of class balancing techniques for credit card fraud detection," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 2017, pp. 2747-2752, doi: 10.1109/ICPCSI.2017.8392219.

[4] L. S. Chandradeva, I. Jayasooriya and A. C. Aponso, "Fraud Detection Solution for Monetary Transactions with Autoencoders," 2019 National Information Technology Conference (NITC), Colombo, Sri Lanka, 2019, pp. 31-34, doi: 10.1109/NITC48475.2019.9114519.

[5] M. Zamini and G. Montazer, "Credit Card Fraud Detection using auto encoder based clustering," 2018 9th International Symposium on Telecommunications (IST), Tehran, Iran, 2018, pp. 486-491, doi: 10.1109/IS- TEL.2018.8661129.

[6] H. Insan, S. S. Prasetiyowati and Y. Sibaroni, "SMOTE-LOF and Borderline-SMOTE Performance to Overcome Imbalanced Data and Outliers on Classification," 2023 3rd International Conference on Intelligent Cybernetics Technology and Applications (ICICyTA), Denpasar, Bali, Indonesia, 2023, pp. 136-141, doi: 10.1109/ICICyTA60173.2023.10428902.

[7] U. Jabeen, K. Singh and S. Vats, "Credit Card Fraud Detection Scheme Using Machine Learning and Synthetic Minority Oversampling Technique (SMOTE)," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 122-127, doi: 10.1109/ICIRCA57980.2023.10220646.

[8] N. S. S. Pranavi, T. K. S. S. Sruthi, B. J. Naga Sirisha, M. S. Nayak and V. S. Gupta Thadikemalla, "Credit Card Fraud Detection Using Minority Oversampling and Random Forest Technique," 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, 2022, pp. 1-6, doi: 10.1109/INCET54531.2022.9824146.

[9] A. K. Uttam and G. Sharma, "A Comparison of Data Balancing Techniques for Credit Card Fraud Detection using Neural Network," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1136-1140, doi: 10.1109/I-SMAC52330.2021.9640911.

[10] C. -H. Chang, "Managing Credit Card Fraud Risk by Autoencoders : (ICPAI2020)," 2020 International Conference on Pervasive Artificial Intelligence (ICPAI), Taipei, Taiwan, 2020, pp. 118-122, doi: 10.1109/IC- PAI51961.2020.00029.

[11] J. Chen, Y. Shen, and R. Ali, "Credit Card Fraud Detection Using Sparse Autoencoder and Generative Adversarial Network," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2018, pp. 1054-1059, doi: 10.1109/IEMCON.2018.8614815.

[12] N. T. N. Anh, T. Q. Khanh, N. Q. Dat, E. Amouroux and V. K. Solanki, "Fraud detection via deep neural variational autoencoder oblique random forest," *2020 IEEE-HYDCON*, Hyderabad, India, 2020, pp. 1-6, doi: 10.1109/HYDCON48903.2020.9242753.

[13] A. G. Putrada and N. G. Ramadhan, "MDIASE-Autoencoder: A Novel Anomaly Detection Method for Increasing The Performance of Credit Card Fraud Detection Models," *2023 29th International Conference on Telecommunications (ICT)*, Toba, Indonesia, 2023, pp. 1-6, doi: 10.1109/ICT60153.2023.10374051.

[14] M. T. Vo, T. Nguyen and T. Le, "Robust Head Pose Estimation Using Extreme Gradient Boosting Machine on Stacked Autoencoders Neural Network," in *IEEE Access*, vol. 8, pp. 3687-3694, 2020, doi: 10.1109/ACCESS.2019.2962974.

[15] K. L. Kurien and A. A. Chikkamannur, "Benford's Law and Deep Learning Autoencoders: An approach for Fraud Detection of Credit card Transactions in Social Media," *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, Bangalore, India, 2019, pp. 1030-1035, doi: 10.1109/RTEICT46194.2019.9016804.