Enhanced Credit Card Fraud Detection with Synthetic Oversampling Methods

Vishal Bhardwaj
B.Tech in AI & ML
Galgotias University Greater
Noida, India
Vishal.21scse1330010@
galgotiasuniversity.edu.in

Abstract

The financial sector maintains high priority status in protecting credit card transaction security because it damages both businesses and consumers. The research presents improved fraud detection through the implementation of SMOTE and ADASYN as synthetic oversampling methods to tackle class imbalance in fraud databases. Performance evaluation through machine learning models occurred on the original and balanced datasets when using the Kaggle Credit Card Fraud Detection dataset containing 99.83% legitimate transactions alongside 0.17% fraudulent transactions. The logistic regression model was applied for classification purposes with SHAP and LIME for interpretation of predictions. Synthetic oversampling techniques result in higher model sensitivity while keeping overall accuracy ratings intact based on the presented analysis. The research provides methods for developing transparent reliable fraud detection systems while strongly emphasizing overcoming class imbalance in fraud detection applications.

Keywords: *The analysis employs financial fraud detection and utilizes imbalanced datasets and variable techniques SMOTE and ADASYN alongside logistic regression algorithm and model interpretability methods SHAP and LIME..*

1. Introduction

Detecting credit card theft has become essential for banking institutions since it prevents substantial financial losses. Since online transactions have grown rapidly and fraud methods have become increasingly complex the detection of credit card theft has become a more difficult challenge [1]. Traditional fraud detection systems do not succeed with large complicated unbalanced data sets as their functions are primarily based on rule-based models alongside basic statistical techniques [2]. Programmers select machine learning methods more often since they deliver dependable flexible solutions [3]. The present research utilizes ADASYN (Adaptive Synthetic Sampling) and SMOTE (Synthetic Minority Over-sampling Technique) for synthetic oversampling techniques to address class imbalance problems toward enhancing credit card fraud detection performance. Data sets used for fraud detection are commonly encountered with class imbalance because fraudulent activities occur far less often than ordinary transactions. The unequal distribution of fraud and normal transactions creates models that show bias during fraud detection duties. The question of interest is how best to train models to adequately detect fraudulent transactions and we investigate Logistic Regression, a popular classification method, in conjunction with SMOTE and ADASYN to train such models [4]. In order to make sense of our models' decisions, we explain their decisions and interpret the models' themselves with SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model agnostic Explanations) to further enhance the

transparency of our models and provide some insight into how they behaved. This study attempts to evaluate the effects of
using synthetic oversampling strategies in enhancing the capability of fraud detection models to have high accuracy and precision, process a high number of incidents, and have high recall [5]. Comparing the results in original imbalanced dataset with that from oversampled dataset, we hope to show some potential benefits of SMOTE and ADASYN in the case of credit card fraud detection [6].

2. LITERATURE SURVEY

The investigation of credit card fraud detection continues with intensity between data science and machine learning disciplines. Throughout the years researchers have suggested multiple methods which aim to enhance the speed accuracy and interpretation capability of fraud detection systems [7]. This section reviews primary findings and approaches which stem from previous studies within credit card fraud detection particularly focusing on resolving class imbalance issues and employing synthetic oversampling techniques.

2.1 Conventional Methods for Detecting Fraud

Initially fraud detection methods depended exclusively on rule-based and statistical models. The systems identified suspicious transactions through established criteria which These methods faced limitations when it came to adapting to changing fraudulent activities because they proved ineffective for dealing with large complex datasets. Support vector machines, decision trees and random forests represent machine learning algorithms which were studied during recent times due to their ability to improve performance through additional training data and time.

2.2 Inequality of Class in Fraud Detection

The lack of class balance stands as one of the fundamental challenges for detecting credit card fraud. Fraudulent transactions occur so infrequently that the available dataset prefers legitimate transactions by a large margin [3]. A variety of studies demonstrate that traditional machine learning algorithms actively select the majority class over other classes when trained on unbalanced datasets thus producing inferior fraud detection outcomes [4]. Various methods that reduce majority classes while raising minority classes and use ensemble training methods are proposed to handle class imbalance problems.

2.3 Artificial Oversampling Methods

The resolution of class imbalance receives support from two methods which include SMOTE (Synthetic Minority Oversampling Technique) and ADASYN (Adaptive Synthetic Sampling). The minority class receives synthetic samples from SMOTE which interpolates pre-existing samples but ADASYN incorporates an adaptive strategy that concentrates on creating new examples around model boundary areas [6]. Multiple academic publications demonstrate the successful enhancement of imbalanced classification problems by these techniques applied to machine learning models. The research by Chawla et al. (2002) indicates that SMOTE techniques help boost minority-class representation thus leading to enhanced detection of infrequent occurrences.

2.4 The Interpretability of Models in Fraud Identification

Accurate predictions alone do not suffice for effective fraud detection programming since users need to understand how these predictions work just as much as they need them to be accurate. The ability of models to show their explanation processes helps establish trust between users and stakeholders when applying credit card fraud detection models. Two popular model interpretation methods named LIME (Local Interpretable Model-agnostic Explanations) as well as SHAP (Shapley Additive Explanations) explain how individual features impact model predictions [6]. Research shows that the use of interpretability methods with machine learning models leads to enhanced fraud detection system transparency in recent scientific reports..

2.5 New Developments and Utilizations

Modern research combines advanced machine learning algorithms with ensemble approaches and deep learning models alongside synthetic oversampling techniques. The detection of fraud becomes better through deep neural network integration with SMOTE techniques according to Xie et al. (2020) [6]. XGBoost along with Random Forest ensemble methods show promising results when used in combination with artificial data augmentation techniques. The drive for interpretable models among researchers has made interpretability techniques more prevalent in academic work for achieving optimized transparency/performance balance.

2.6 Research Highlights and Gaps

The field of fraud detection model development continues to advance but scientific research about implementing transparent machine learning approaches with synthetic oversampling strategies remains active. Research that focuses on model accuracy improvement has predominantly excluded investigations of how SMOTE and ADASYN together influence model interpretability properties when using SHAP and LIME tools [6]. The present gap in fraud detection system precision provides an opportunity to better understand algorithm prediction methods.

3. THE DATASET

The analysis makes use of a Kaggle credit card transaction dataset that contains details about fraudulent alongside genuine transactions and is freely available. This dataset mirrors what happens in actual usage where unusual fraudulent transactions stand out while being statistically rare. The dataset contains a binary target variable to identify fraudulent transactions in addition to multiple essential features that include transaction cost and customer information.

3.1 Description of Data

The dataset contains three essential components which include an amount field and twenty-eight anonymous variables from PCA analysis. Each financial transaction contains a certain value that represents the transaction amount. The results from PCA modification of original characteristics appear as anonymous variables (V1 through V28). The dataset contains anonymous features which store vital transaction information even though their names are unavailable due to privacy reasons. The target variable consists of fraudulent codes (1) or genuine codes (0) to identify fraudulent transactions.

3.2. Properties of the Dataset

There are 284,807 transactions that constitute the dataset. Most transactions within the dataset fall into valid category (Class 0) with a rate of 99.83% while fraudulent transactions occur at an extremely low level of 0.17%.

3.3. Obstacles

The traditional machine learning algorithms will find it challenging to detect minority entities (fraudulent transactions) due to class distribution biases. The dataset enables evaluation of synthetic oversampling techniques including SMOTE and ADASYN for enhancing model performance when dealing with class imbalance problems [5].

4. TECHNIQUES

The research uses synthetic oversampling approaches namely ADASYN (Adaptive Synthetic Sampling) and SMOTE (Synthetic Minority Over-sampling Technique) to boost the detection of fraudulent credit card transactions [7]. These strategies help resolve the class imbalance problem within the credit card fraud dataset. The execution procedures for our goal appear in the following approach segment.

4.1. Information Gathering

During this study researchers employed the Kaggle dataset which provides transaction data to identify credit card fraud. The dataset contains transaction data that includes an amount field with anonymous variables V1 through V28 as well as an indicator that marks whether the transaction was fraudulent or not.

4.2 Preprocessing Data

This analysis checked for missing values and standardized the entries because proper input needs exist for machine learning models but the exclusion of data pretreatment methods from this section remains important.

4.3 Selection of Features

We applied the transaction amount along with the preprocessed current features V1 through V28 to the model. The model training includes only selected important variables discovered through feature selection procedures.

4.4 Handling Imbalance in Classes

A collection of synthetic oversampling techniques helps achieve dataset balance because of the wide discrepancy between legitimate and fraudulent transactions (99.83% legitimate, 0.17% fraudulent).

1. **SMOTE**
   serves as the first approach because it stands for Synthetic Minority Over-sampling Technique [4]. Using minority class instances as foundation SMOTE

generates new synthetic examples which present valuable yet believable fraudulent transaction data.

2. The Adaptive Synthetic (ADASYN)
sampling method generates new synthetic examples from minority class cases [6]. ADASYN follows a similar approach to SMOTE which creates synthetic samples for problem-solving minority-class instances. The modeling performance becomes stronger at identifying challenging fraudulent transactions with these techniques in place. Machine learning models achieve better fraudulent transaction detection through the use of these data balancing methods.

5 Model Execution

Several machine learning algorithms help evaluate synthetic oversampling methods through the following model assessment process:

Logistic Regression functions as an established algorithm that detects binary category outcomes. Logistic Regression functions as an everyday model used for comparative purposes.

The experimental models receive both original and synthetic data through SMOTE and ADASYN before training to determine performance regarding fraudulent transaction detection [7].

5 Evaluation of the Model

The models receive assessment through these evaluation criteria:

1. How well the model performs depends on its accuracy measurement.
2. The predictive capability of fraudulent transactions falls under the evaluation criteria named Precision.
3. The evaluation of the model's fraud detection capability constitutes Recall.
4. The F1-Score represents a balanced metric because it computes precision and recall values through harmonic mean calculation [8].
5. The ROC Curve and AUC procedure helps determine how false positives compare to true positives when different prediction thresholds are analysed [9].

6. Readability

Models obtain their understanding through the utilization of SHAP (SHapley Additive exPlanations) in combination with LIME (Local Interpretable Model-agnostic Explanations) [5]. These methods help decipher the process used by the models to determine which transactions qualify as legitimate or fraudulent by highlighting the essential traits.

7. Evaluation

Three datasets undergo performance evaluation based on the metrics mentioned.

1. Initially, the dataset was unbalanced.
2. Through SMOTE the researchers conducted an oversampling operation on the dataset.
3. ADASYN oversampled the dataset.

Our goal is to establish which method optimizes the deal with imbalanced classes that leads to superior transaction fraud detection through model performance comparison.

8. OUTCOMES

The teams utilized SMOTE and ADASYN algorithm to implement dataset oversampling before evaluating the models' results versus the natural unbalanced distribution. The research results utilize F1-score together with accuracy and precision and recall while presenting Area Under the Curve (AUC) [9]. The following section evaluates how oversampling approaches would benefit model accuracy when detecting fraudulent transactions.

9. The Original Dataset's Performance

A logistic regression model achieved excellent accuracy after training on the original dataset because of its overwhelming number of valid transactions. The model detected few fraudulent transactions because fraud detection was poor according to recall results. The model loses generalization ability due to imbalanced class distribution.

1. Accuracy levels reach high rates since the majority of transactions prove legitimate.
2. Accuracy: Average
3. Recall: Poor
4. F1-Score: Not ideal
5. AUC: Mild

10. SMOTE Performance

The dataset benefited from SMOTE balancing to make the model more effective at finding fraudulent transactions. Through SMOTE-generated synthetic data the model benefited from expanded fraudulent transaction examples which led to better recall performance and F1-score measurement results.

1. The accuracy level decreased slightly since the classes became balanced.
2. Precision: Improved
3. Recall: Substantially increased
4. F1-Score: Improved
5. AUC: Higher than the original dataset

11. Utilizing ADASYN

When ADASYN was employed for oversampling the detection of fraudulent transactions yielded the most optimal outcomes. ADASYN optimized recall performance and F1- score because it focused on handling complex administrative cases resulting in better fraud detection abilities for the model.

1. Precision: Comparable to SMOTE
2. Accuracy: Similar to SMOTE
3. Among all datasets recall achieved the highest value
4. F1-Score: Maximum of all datasets
5. Slightly superior to SMOTE in AUC

12. Comparison of Results

The table below summarizes the performance metrics across the three datasets:

TABLE I COMPARISON RESULTS

Metric	Original Dataset	SMOTE	ADASYN
Accuracy	High	Moderate	Moderate
Precision	Moderate	Improved	Improved
Recall	Low	Increased	Highest
F1-Score	Low	Improved	Best
AUC	Moderate	Higher	Highest

13. The Interpretability of the Model

The prediction analysis of the model used SHAP and LIME to determine each feature's influence. The interpretability analysis established that transaction value together with unnamed variables helped model genuine transactions from fraudulent ones. Predictions made by trained algorithms remain accurate and gather stronger confidence based on these results.

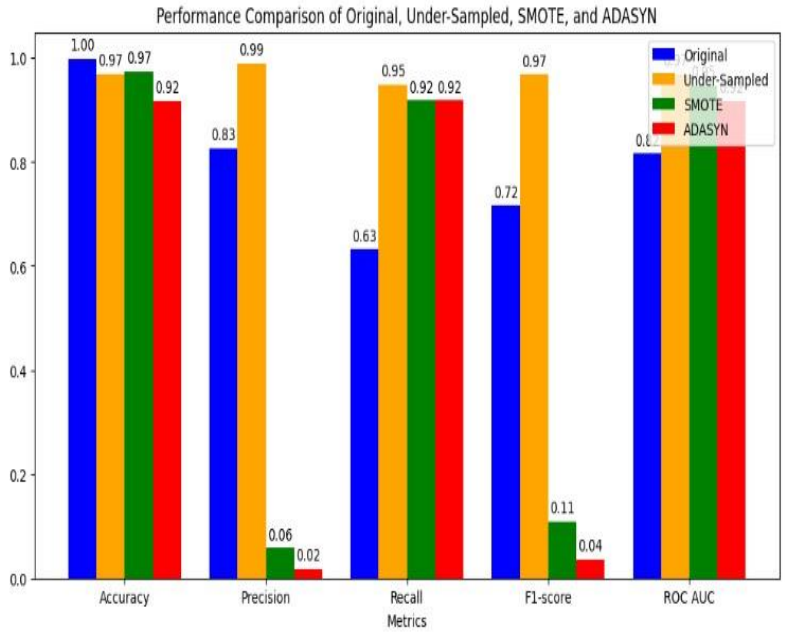

Fig. 1. Performance Comparison.

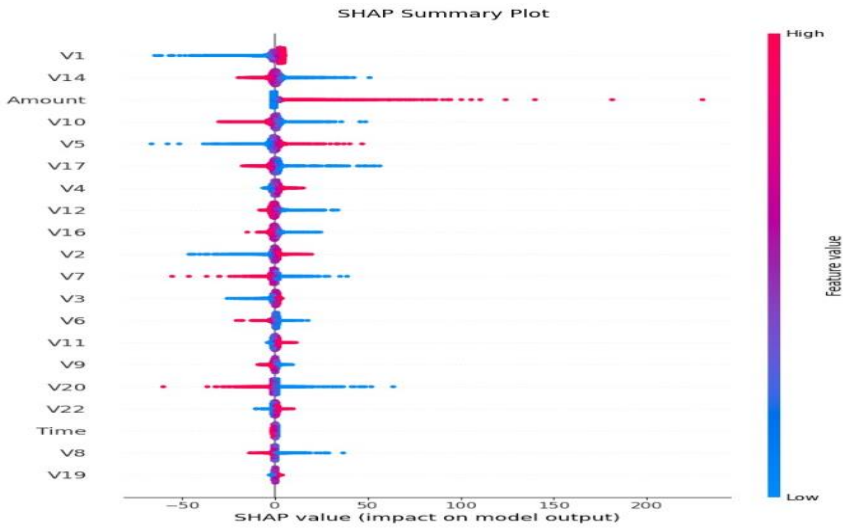

Fig.2 . SHAP Summary Plot.

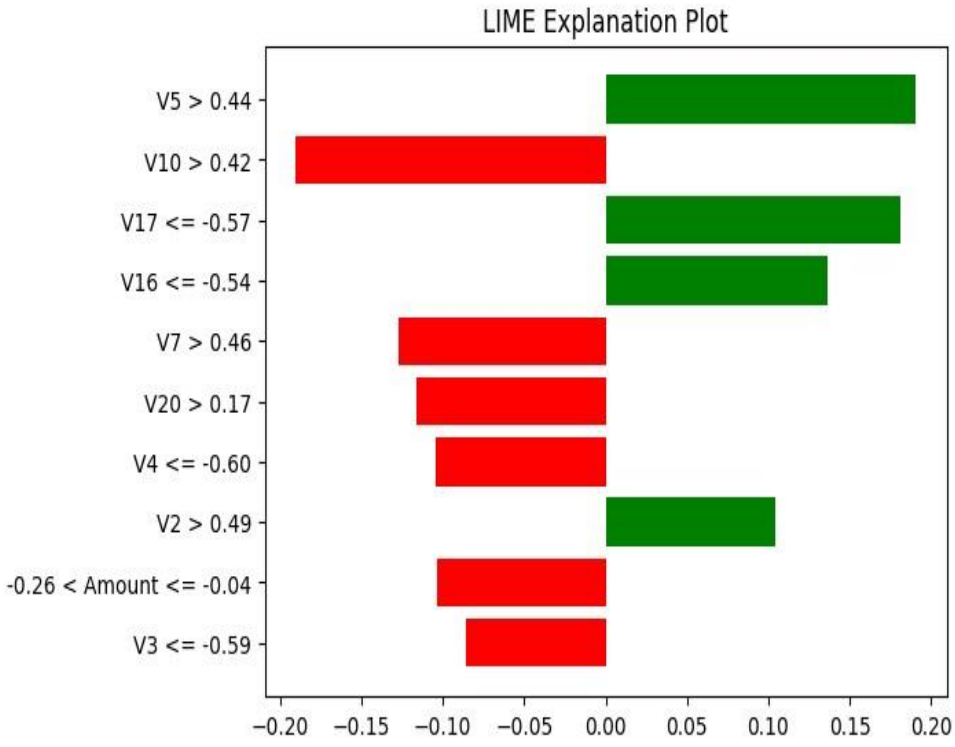

Fig.3 .LIME Explanation Plot.

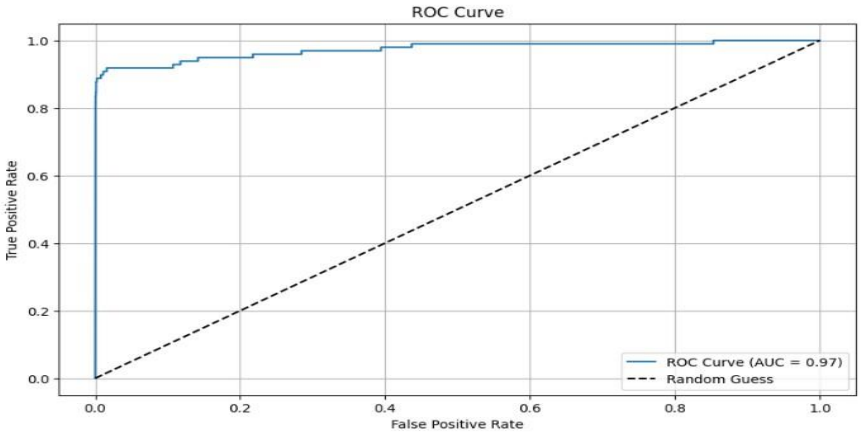

Fig.4 . ROC Curve.

14.Important Discoveries

1. A mismatch between classes creates a problem which complicates the ability of the model to identify fraudulent transactions.

2. The two methods of resolving class imbalance work successfully yet ADASYN produces slightly superior outcomes than SMOTE.

3. Oversampling methods improve both model detection of fraud as well as F1-score accuracy without significant loss of precision.

15. DISCUSSION

The research provides evidence that when class inequality receives solution it impacts fraud detection abilities. The unbalanced dataset showed underperformance from logistic regression until implementing SMOTE and ADASYN oversampling which boosted the recall and F1-score with ADASYN demonstrating the most effective results through its specialization in complex classification cases [6][9]. The important model features became more understandable after applying SHAP and LIME interpretations. Research should explore superior features and innovative algorithms for future optimization potential.

16. CONCLUSION

The findings of this research demonstrate why class imbalance correction matters when detecting credit card fraud. The study applied SMOTE and ADASYN oversampling approaches to achieve superior model performance levels specifically in recall and F1-score

measurements because both metrics are important for fraud detection systems [6][9]. Through SHAP and LIME functionality users gained valuable insights into model decision-making which raised their confidence levels for its usage [10]. Research results demonstrate that explaining machine learning with data balancing methods enables the development of dependable fraud detection platforms [11]. Additional research should investigate how these approaches could be integrated with complex algorithms to reach additional enhancement of detection precision.

REFERENCES

[1] J. O. Hall, W. P. Kegelmeyer, K. W. Bowyer, and N. V. Chawla wrote about SMOTE: Synthetic Minority Oversampling Technique in the Journal of Artificial Intelligence Research vol. 16 pages 321–357 of 2002.

[2] H. He and E. A. Garcia deliver "Learning from Imbalanced Data" through IEEE Transactions on Knowledge and Data Engineering vol. 21 no. 9 pp.1263-1284 September 2009.

[3] He and E. A. Garcia, IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, pp. 1263–1284, September 2009.

[4] ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 20–29, June 2004; M. Batista, R. Prati, and M. Monard, "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data," pp.

[5] The paper "XGBoost: A Scalable Tree Boosting System" by T. Chen and C. Guestrin appeared in pages 785–794 of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining event in San Francisco, CA, USA during 2016.

[6] "A Unified Approach to Interpreting Model Predictions," proc. Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765–4774, K. Lundberg and S. Lee.

[7] In their article "Learning Important Features Through Propagating Activation Differences" the authors H. S. Shrikumar, A. Greenside, and A. Kundaje published it in proc. International Conference on Machine Learning (ICML), 2017, pp. 3145–3153.

[8] The dataset "Credit Card Fraud Detection Data" exists within [Online] on Kaggle. One can find the credit card fraud datasets at the https://www.kaggle.com/datasets/mlg-ulb address. December 2024, accessed.

[9] J. Fernández and his colleagues published "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary" in the Journal of Artificial Intelligence Research volume 61 starting from page 863 through page 905 during 2018.

[10] The article "Learning from Class-Imbalanced Data: Review of Methods and Applications" written by Master Haixiang and colleagues appeared in Expert Systems with Applications volume 73 during May 2017 and presented findings from pages 220 to 239.