An efficient implementation of credit card fraud detection using CatBoost algorithm

Vadhri Suryanarayana[1], Kuruva Maddileti[2], Dune Satyanarayana[3], R Leela Jyothi[4], Kavuri Sreekanth[5], Praveen Mande[6], Raghava Naidu Miriyala[7], Oggi Sudhakar[8]
[1]Department of Artificial Intelligence and Data Science, Lakireddy Bali Reddy College of Engineering, Mylavaram, India
[2]Department of Mathematics, K.V. Subba Reddy Engineering College, Kurnool, India
[3]Department of Computer Science and Engineering, Aditya College of Engineering and Technology, Surampalem, India
[4]Department of Computer Science and Engineering, SRKR Engineering College, Bhimavaram, India
[5]Department of Computer Science and Engineering, Koneru Lakshmaiah University, Guntur, India
[6]Department of Electrical Electronics and Communication Engineering, GITAM School of Technology,
GITAM Deemed to be University, Visakhapatnam, India
[7]Department of Computer Science and Engineering, Krishna University College of Engineering and Technology, Machilipatnam, India
[8]Department Electronics and Communication Engineering, GIET Engineering College, Rajahmundry, India

Article Info

Article history:

Received Jun 1, 2024
Revised Dec 9, 2024
Accepted Feb 27, 2025

Keywords:

CatBoost
Credit card
Financial services
Fraud transaction
Fraudsters
Machine learning

ABSTRACT

Transaction fraud has grown to be an important issue in worldwide, banking and commerce security is easier access to trade information. Every day, there are more and more incidents of transaction fraud, which causes large financial losses for both consumers and financial professionals. The ability to identify transaction fraud is getting closer to reality due to improvements in computer science's machine learning (ML) and data mining areas. So, one of them that is becoming dangerous is credit card fraud (CCF). Millions of people are experiencing financial loss and identity theft as a result of these malicious operations. The CCF of many illegal activities that fraudsters are always using new methods to carry out. One major problem facing financial services sector is CCF. To overcome this, categorical boosting (CatBoost) algorithm is explained as a solution to these problems. Fraud or fraudulent transactions are identified using this effective CatBoost algorithm implementation for identification of CCF. Thus, in terms of accuracy, precision, and detection rate this method gives better performance.

Corresponding Author:

Vadhri Suryanarayana
Department of Artificial Intelligence and Data Science, Lakireddy Bali Reddy College off Engineering
Mylavaram, Krishna District, Andhra Pradesh, India
Email: vadhri.abhiram@gmail.com

1. INTRODUCTION

In every sector of people’s lives, including banking and business, internet use is increasing. As a result, more data is being integrated and virtualized. The amount and frequency of online transactions have increased significantly in the recent times, leading to a growing number of people using the internet as a platform for conducting business [1]. An increase in activity is seen in the transactions conducted online. Fraudsters frequently use a number of techniques to rapidly move large amounts of money and get user information, causing both individuals and institutions to face serious financial losses [2].

In every sector of people’s lives, including banking and business, internet use is increasing. As a result, more data is being integrated and virtualized. The amount and frequency of online transactions have increased significantly in the recent times, leading to a growing number of people using the internet as a platform for conducting business [1]. An increase in activity is seen in the transactions conducted online.

Fraudsters frequently use a number of techniques to rapidly move large amounts of money and get user information, causing both individuals and institutions to face serious financial losses [2].

Fraud in the financial sector is an extensive problem with serious consequences that affects both businesses and the government. Credit cards can be connected to two different types of fraudulent activity: internal and external card fraud. The people who are using credit card to make transactions with another person's approval, that is called as internal card fraud; incase credit card is lost, that is known as external card fraud. When cardholders and financial institutions crash to carry out fraudulent operations by adopting a false identity, inside card fraud occurs [3]. The majority of credit card fraud (CCF) incidents are caused by external card fraud, which has been the focus of much research on identifying and preventing. The traditional methods utilized in past to identify fraud transactions have shown to be more time-consuming and ineffective [4]. As a result, using of big data has led to growing increasing of use of manual processes. On the other hand, financial institutions have shifted their attention to current computational methods that deals with issue of CCF.

Fraud is defined as improper or criminal fraud committed with the intention of obtaining money or personal benefit. Fraud prevention and detection are two methods which are used to prevent fraud losses [5]. The proactive strategy of fraud prevention works to stop fraud before it starts. Therefore, detection of frauds become necessary, in case of fraudsters before completing fraud transaction. This CCF mentions illegal using of credit card details for financial transactions. Hence, credit card transactions can be made digitally or physically [6]. The credit card is used for in-person transactions. Digital transactions can take place online or over the phone. The cardholders frequently mail or phone in their card number, expiration date, and card verification number.

Fraud with credit cards is defined as the unauthorized use of funds in any transaction [7]. These credit card scams rose quickly with digital transactions, particularly during pandemics. CCF occurs in two types: the first category is card-not-present fraud when the card number, card verification code (CVC), expiration date are collected without cardholder having to physically present the card to the vendor. Online transactions are where this type of scam is most frequently observed. Card-present fraud is the theft of credit card details through a point-of-sale system during physical transactions [8]. It is now possible to prevent card-present fraud using chip-based cards. In 2020, there were 4,59,297 reported cases of CCF based on the Annual Report of the Federal Trade Commission. According to the research, identity theft is frequent type of CCF, with a 44.6% increase in cases in 2020 over the statistics for 2019. While credit cards offer significant financial convenience, they also present a serious threat by creating opportunities for criminals to commit fraud through methods such as forgery and theft of credit card information. Such criminal acts not only lead to huge economic losses for financial institutions, but also cause a far-reaching threat to the socio-economic system. Traditional fraud detection methods depend on statistical and tree models to identify anomalies in transaction data [9].

This makes it clear that effective fraud detection techniques are required to prevent monitoring losses. Credit fraud can take many different forms. These include card fraud related to non-mail receipts, account takeovers, credit card imprints prepared manually or electronically, counterfeit cards, identity theft, form jacking redirection, intercepting mail, application fraud, credit card theft, identity theft, spoofing phone numbers and locations, fraudulent merchant websites, lost and stolen cards, and merchant cooperation are various types of fraud. The target class is mostly predicted by the classification methods.

In order to prevent financial loss due to card fraud, a fraud-detection system is developed with a great deal of effort and finance. Numerous machine learning (ML) algorithms are used to evaluate huge amounts of data, integrating both traditional techniques like decision trees (DT), logistic regression (LR), support vector machine (SVM), and hidden Markov models with advanced techniques such as deep learning (DL) as well as gradient boosting tree [10]. The most promising of them are gradient boosting tree and DL, particularly deep neural network (DNN) and categorical boosting (CatBoost) because of their excellent fraud detection performance. When it comes to managing both new users and users with historical transaction histories, we take CatBoost for granted because time-based DNN designs are unable to include the user's transaction history, but CatBoost-based techniques have this advantage.

This spiral oversampling balancing technique (SOBT) combined with a fraud feature-boosting technique is used by a CCF detection model. Specifically, we describe an elimination method of compound classifications to improve the quality of the credit card transaction dataset [11]. This strategy aims to remove features that are excessively redundant and connected.

With the help of ensemble synthesized minority oversampling techniques-generative adversarial network (ESMOTE-GAN), the suggested CCF detection model is designed around ensemble learning and generative adversarial network (GAN). To stop the GAN from modeling the noise, multiple subsets were retrieved using under-sampling, and SMOTE was then applied to create less skewed sets. The synthesized subsets were produced by training several sets of GAN models on these subsets. The suggested ESMOTE-GAN method was then used to train a collection of random forest (RF) classifiers [12]. For decision-making, a weighted voting mechanism was used to aggregate probabilistic results of trained classifiers. According to the results, suggested model improved overall performance by 1.9% and the detection rate by 3.2%, respectively, while having a false alarm rate of 0%. The research gap observed in this research is precision.

The ML techniques are utilized to identify the CCF. Initially, traditional methods are used. Then, hybrid techniques are used, emerging majority voting with adaptive boosting (AdaBoost). To evaluate the model's efficiency, a publicly accessible credit card data set is utilized. Subsequently, the study looks at real credit card data set that was acquired from financial institution. Experimental data indicate that the majority voting method, when combined with noise [13], has great accuracy rates in identifying instances of CCF.

A time-aware gate that stores behavioral changes brought on by the user's subsequent transactions is combined with lengthy short-term memory in an enhanced unique model. By establishing connections between previous and current transactional actions, in order to integrate behavioral periodicity into the model, a current-historical attention module is created. Complete and logical behavioral representations are supposed to be learned through an interaction module [14].

A ML system that uses skewed real-world datasets created from European credit card holders is utilized to identify CCF [15]. Our method of sampling dataset with the SMOTE addressed the problem of class imbalance. The following ML models were used to calculate this system: SVM DT, extra tree (ET), extreme gradient boosting (XGBoost), RF, and LR. To improve ML models categorization quality, they were combined with the AdaBoost method. AUC (area under the curve) and the Matthews correlation coefficient (MCC) were utilized to calculate the accuracy, recall, and precision of this model. Additionally, to further validate results of this study, the suggested method was applied to highly skewed synthetic dataset of CCF. The research gap observed in this research is detection rate.

In this powerful DL method, the meta-learner is a multi-layer perceptron (MLP), a stacking ensemble architecture with base learners that are gated recurrent units (GRU) and long short-term memory (LSTM) neural networks. The dataset's class distribution is balanced using a hybrid approach called SMOTE and edited nearest neighbor (SMOTE-ENN). A sensitivity of 1.000 and specificity of 0.997, were shown by the experiment's results, might be achieved by combining the SMOTE-ENN method with the recommended DL ensemble [16]. The research gap observed in this research is detection rate.

The type of self-paced learning used in this work is adaptive hybrid weighted self-paced learning, this improves system's goal function by changing the fundamental learner selection approach of the Adaboost algorithm. It is possible to decreases effect of human experience in model training by using self-adaptive threshold finding approach used here. Using a double-fault measurement, degree of diversity among base categories is also computed from the perspective of generalization error, this is included in this paper's weight computation for weak learners. To determine the optimal range of effect coefficients, experiments are carried out [17].

A real-world fraud-detection system (FDS) frequently utilizes the assumptions used in fraud detection. There are two primary dimensions to this realism: 1) the timing and manner in which supervised information is propagated, and 2) the measurements by which fraud-detection effectiveness is evaluated [18]. There are three main contributions in this paper. First, they offer formalization the identification of fraud issue with support of industry partner which represents the real-world operational characteristics of FDSs handling numerous credit card transactions on every day. Additionally, they provide examples of the best measurements of performance to apply to fraud detection. Secondly, we develop and evaluate a unique learning approach that successfully addresses verification delay, concept drift, and class imbalance. Third, in our research, they use a real-world data stream over three years and more than seventy-five million authorized transactions to denotes the effects of idea drift and class imbalance.

An innovative method that simulates fraudulent behavior as a process that is dependent on time. Modeling temporal link among frauds is basis for design and evaluation of an oversampling method known as "Adversary-based oversampling" (ADVO), which is primary contribution. Two learning approaches are used to implement the strategy: the first is a novel regression-based oversampling model that makes predictions about future fraudulent activity based on characteristics of past fraud. Secondly, the time GAN oversampling technique has been updated and adapted to the CCF detection context. For creating artificial frauds time series, for this change, it is necessary to treat a collection of card frauds as a time series [19].

A neural network (NN) ensemble classifier combined with hybrid data resampling technique that is an effective way to prevent CCF. By utilizing an LSTM–NN as base learner in AdaBoost technique, ensemble classifier is obtained [20]. This SMOTE-ENN approach is used to achieve hybrid resampling in the meantime. Real-world credit card transaction datasets that are publicly available are used to demonstrate the efficiency of the suggested approach. The following models are used to benchmark the performance of the suggested method: DT, MLP, SVM, conventional AdaBoost, and LSTM. With a sensitivity of 0.996 and specificity of 0.998, recommended LSTM ensemble performs better than the earlier algorithms. The results of the experiment show that the classifiers performed better after being trained on resampled data. The research gap observed in this research is detection rate.

From their sequential historical transactions over time, behavioral information is extracted, a module has been developed and utilized in their previous transactional activities, this makes it possible for the proposed model to represent the periodicity and behavioral purpose. To acquire more thorough and logical representations, an interaction module is proposed. Two large real-world transaction datasets were used in tests to demonstrate the effectiveness of the learnt transactional behavioral representations, one public and the other large-scale [21]. The results demonstrate that the learnt representation is capable of accurately differentiating between fraudulent and genuine activities, and that suggested solution performs advanced techniques in terms of CCF detection when measured against a number of evaluation criteria.

Convincing and diversified minority class data are produced using a novel oversampling technique. To get the ensemble learning classification model trained, minority class fraud data is generated and added to the training set [22]. The enhanced VAEGAN oversampling approach works better than the other oversampling methods, according to experiments on an open credit card dataset, including GAN, variational autoencoder (VAE), and SMOTE, with respect to F1_score, precision, and other measurements. The imbalanced data categorization issue is successfully resolved by the oversampling technique, which is based on the improved VAEGAN.

An optimized-light gradient boosting machine (O-LightGBM) is an intelligent method for detecting fraud in credit card transactions. The suggested method modifies light gradient boosting machine's parameters by intelligently integrating hyperparameter optimization algorithm based on Bayesian theory [23]. Two public credit card transaction data sets from the real world, one with genuine transactions and the other with fraudulent transactions, were used in the tests to demonstrate that effectively our proposed OLightGBM detects CCF.

Four steps made up a new fraud detection system that improves a cardholder's behavioral patterns, they first categorize all of the cardholders into different categories based on comparable transaction behaviors by using the cardholders past transaction data. In order to aggregate the transactions in each group, we thus suggest a window-sliding method. Based on both the cardholder's past transactions and the combined transactions, we then extract a set of unique behavioral patterns for every cardholder [24]. After that, they train a set of classifiers using all of the behavioral patterns for every group.

Using transactions made by European cardholders over a period of two days in September 2013, an open CCF dataset is used to test approach [25]. Based on experimental results, the VAE method outperforms both traditional DNN methods and synthetic minority oversampling methods. Furthermore, it performs better than the new GAN model-based oversampling techniques. After the expanded dataset was submitted to the training baseline, the VAE model test exhibits positive results on indicators like as precision, F1-measure, accuracy, and specificity. The results of this experiment imply that imbalanced classification issues can be successfully solved using the VAE-based oversampling approach. Following is the arrangement of the remaining paper: in section 1, the literature survey is included, in section 2 the framework of an efficient implementation of CCF detection using CatBoost algorithm is presented; section 3 explains result analysis; section 4 presents the conclusion in the paper.

2. FRAMEWORK OF AN EFFICIENT IMPLEMENTATION OF CREDIT CARD FRAUD DETECTION USING CATBOOST ALGORITHM

In this section, framework of an efficient implementation of CCF detection using CatBoost algorithm is observed in Figure 1. Every time a credit card is used for a financial transaction, information is recorded electronically and added to a credit card database. This comprises the transaction's date and time, the merchant's details, the transaction sum, and the credit card holder's data, this may contain the credit card number, name, and address of the cardholder. The process of transforming raw data into a form that can be understood is known as data preparation since raw data processing is impossible, so it is also an essential stage in data mining. Data mining and ML techniques should not be used without first verifying the quality of the data. Credit card database is taken as input, then the data is pre-processed. After removing the unwanted data, the remaining data is analyzed and cleaned. So, that features are selected from that data, from that features selection related features are extracted. Finally, CatBoost algorithm is applied to test database and trained database. So, that fraud transaction will be detected by this system.

Applying statistical and/or logical methods methodically to explain and show, illustrate and evaluate, and describe data is the process of data analysis. Various structural problems in data sets are corrected through data cleansing. This includes typographical problems such as misspellings, incorrect numeric entries, syntax issues, and missing values, including empty or null fields that should have data. Feature selection eliminates noise in the data and uses only the relevant information to lower the input variable in your model. The technique involves selecting the relevant features for your ML model automatically, taking consideration of the type of problem you are attempting to address. ML algorithms can utilize feature extraction to transform unprocessed data image files are a typical data into numerical features.

For ML applications, data scientists can create new features by extracting an object's form or an image's redness value. The database is divided into training and testing databases at this step. The testing database's objective is to test (evaluate) the generated classifier, whereas the training database's objective is to build the classifier (model). Separate from the training dataset, an additional subset of the original data is called the test dataset. Once the model training is finished, it is used as a benchmark for model evaluation because it includes some attributes and a class probability distribution that are similar. Then the database is trained. The data used to train a ML algorithm or model is known as training data in ML. In order to evaluate or prepare training data for ML applications, human intervention is necessary.

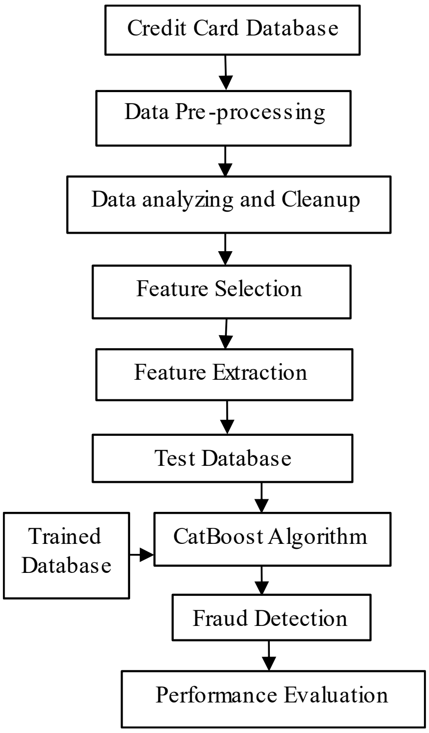

Figure 1. Framework of an efficient implementation of CCF detection using CatBoost algorithm

Next, make use of DT are utilized in CatBoost, a supervised ML technique, to perform regression and classification tasks in the train using AutoML tool. The two primary characteristics of CatBoost are that it utilizes gradient boosting (the Boost) and operates on categorical data (the Cat). When working with data sets that contain categorical features, the ML gradient-boosting technique CatBoost performs exceptionally. CCF investigations generally involve banks analyzing transaction patterns and details for signs of unauthorized activity. They may collaborate with law enforcement, merchants, and cybersecurity experts if the situation requires more extensive scrutiny. After the application of CatBoost algorithm fraud is detected. Finally, the out is evaluated. To determine how well a model is doing at the particular task it was created, ML models are monitored and evaluated. Utilizing metrics like regression and classification, there are numerous approaches to model evaluation in model monitoring.

CatBoost stands for both boosting and category. CatBoost is used to solve business challenges and in numerous frameworks, including TensorFlow and core ML. It is simple to use, has great performance, and generates results without requiring a lot of data training. It also supports more descriptive data formats out of the box. It reduces the need for intensive hyper-parameter adjustment and lowers overfitting, is applicable to both regression and classification and doesn't require the transformation of the dataset. "CatBoost classifier" is used for classification. This approach differs from other boosting algorithms primarily because it uses symmetric trees. For each data point, CatBoost trains log (number of data points) models to determine the residuals. It requires less prediction time because it simply takes past information points into account when calculating. In traditional system, detection of CCF is not accurate. Even it takes a lot of time to find fraud transaction. Therefore, by using this method fraud is detected accurately. Even fraud transactions are detected in less time also. The credit card fraudulent transaction detection, is important to find out the fraud transactions on products as well as general payments. So that customers can able to avoid payment for the purchase of products that they didn't buy as well as normal transactional payments.

3. RESULT ANALYSIS

In this section, a performance analysis of the framework for an efficient implementation of CCF detection using the CatBoost algorithm is performed. In Table 1 the comparison between CCF detection using CatBoost algorithm with RF and DT is observed.

In Figure 2, x-axis demonstrated CCF detection and y-axis demonstrated accuracy. The accuracy comparison graph is observed between RF, DT, and CatBoost for an efficient implementation of CCF detection using CatBoost algorithm. CatBoost algorithm shows high accuracy to detect the fraud transactions when compared with the other algorithms.

A comparison graph of detection rates between RF, DT, and CatBoost is shown in Figure 3. In this Figure 3, x-axis demonstrated CCF detection and y-axis demonstrated detection rate. This demonstrates that the CatBoost algorithm is effectively used to detect CCF. Therefore, RF, DT shows low detection rate and CatBoost achieves high detection rate.

Precision is high for CatBoost algorithm for implementation of CCF detection when compared with RF and DT as shown in Figure 4. Therefore, high precision is achieved for CatBoost, whereas RF and DT has low precision compared with CatBoost. In this Figure 4, x-axis demonstrated CCF detection and y-axis demonstrated precision.

Table 1. Performance analysis

Parameters	RF	DT	CatBoost
Accuracy	91.2	89.3	95.4
Detection rate	88.7	92.1	94.5
Precision	94.7	92.7	96.8

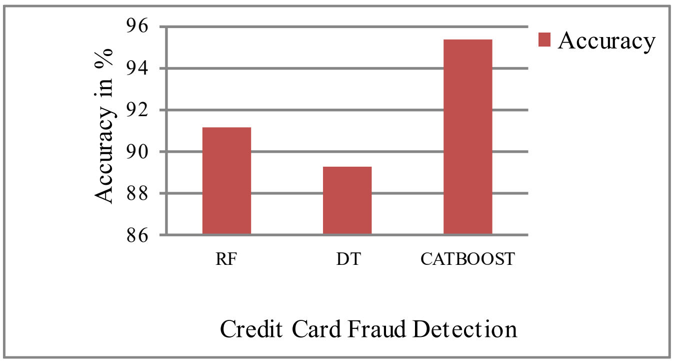

Figure 2. Accuracy comparison graph

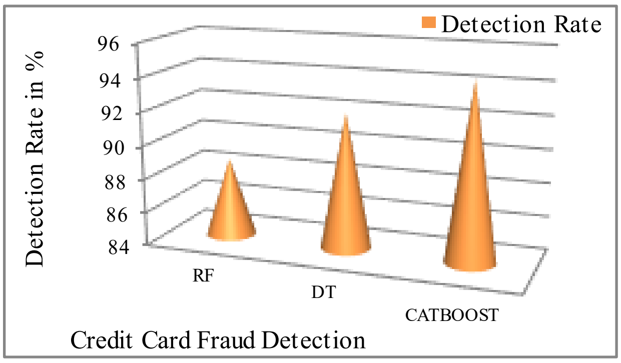

Figure 3. Detection rate comparison graph

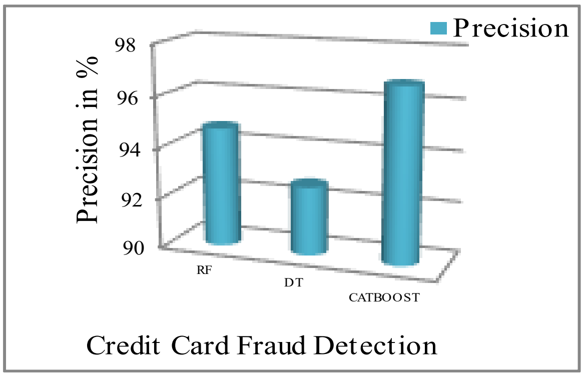

Figure 4. Precision comparison graph

4. CONCLUSION

Hence, an efficient implementation of CCF detection using CatBoost algorithm is concluded in this section. The demand for digital affairs has expanded as the cashless economy has developed. One of the most common types of digital transactions is the usage of credit cards for online transactions. Thus, the main concern is to identify those fake transactions. In traditional methods, detection of CCF is not accurate. Even it takes a lot of time to detect or identify fraud transactions in credit card. Therefore, by using this method fraud is detected or identified accurately. Even fraud transactions are detected in less time. This detection is important to find out the fraud transactions. So that customers don't get charged for the purchase of products that they didn't buy as well as avoid fraud transactions also and it reduces organizations reputational damage with increases security level. Therefore, to detect or identify these fake transactions a novel algorithm is used in this analysis. So, by using this CatBoost algorithm fraud or fake transactions are detected or identified accurately. The parameters like accuracy, detection rate, and precision are compared between CatBoost algorithm with RF and DT methods. Therefore, CatBoost algorithm shows better results when compared with other methods. Thus, in terms of detection rate, accuracy, and precision, this method better performance. In future, this method achieves better results even for large datasets and also sends alert about fraud transaction if it initiated.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Vadhri Suryanarayana	✓		✓		✓	✓		✓		✓			✓	
Kuruva Maddileti		✓		✓		✓		✓	✓		✓	✓		
Dune Satyanarayana			✓			✓			✓	✓			✓	
R Leela Jyothi				✓			✓				✓			
Kavuri Sreekanth	✓		✓		✓			✓		✓			✓	
Praveen Mande		✓		✓					✓			✓		
Raghava Naidu Miriyala	✓		✓		✓		✓				✓		✓	
Oggi Sudhakar					✓		✓			✓		✓		

C	: **C**onceptualization	I	: **I**nvestigation	Vi	: **Vi**sualization
M	: **M**ethodology	R	: **R**esources	Su	: **Su**pervision
So	: **So**ftware	D	: **D**ata Curation	P	: **P**roject administration
Va	: **Va**lidation	O	: Writing - **O**riginal Draft	Fu	: **Fu**nding acquisition
Fo	: **Fo**rmal analysis	E	: Writing - Review & **E**diting		

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

- The data that support the findings of this study are openly available in [(I. Sadgali, N. Sael, and F. Benabbou), (A. Priyadarshini, S. Mishra, D. P. Mishra, S. R. Salkuti, and R. Mohanty)] at http://doi.org/[(10.11591/ijai.v10.i3.pp698-706), (10.11591/ijece.v14i1.pp759-771), (10.11591/ijeecs. v23.i3.pp1634-1642)], reference number [8], [10].
- The data that support the findings of this study will be available in [(A. Mniai, M. Tarik, and K. Jebari), (W. Ning, S. Chen, S. Lei, and X. Liao), (Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang)] [(10.1109/ACCESS.2023.3323842), (10.1109/ACCESS.2023.3290957), (10.1109/ACCESS.2023. 3302339)] following a [6 month] embargo from the date of publication to allow for the commercialization of research findings.
- The data that support the findings of this study are available on request from the corresponding author, [Vadhri Suryanarayana]. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

REFERENCES

[1] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, “Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms,” *IEEE Access*, vol. 10, pp. 39700–39715, 2022, doi: 10.1109/ACCESS.2022.3166891.

[2] S. N. Kalid, K.-C. Khor, K.-H. Ng, and G.-K. Tong, “Detecting frauds and payment defaults on credit card data inherited with imbalanced class distribution and overlapping class problems: a systematic review,” *IEEE Access*, vol. 12, pp. 23636–23652, 2024, doi: 10.1109/ACCESS.2024.3362831.

[3] H. Zhu, M. Zhou, G. Liu, Y. Xie, S. Liu, and C. Guo, “NUS: noisy-sample-removed undersampling scheme for imbalanced classification and application to credit card fraud detection,” *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 1793–1804, Apr. 2024, doi: 10.1109/TCSS.2023.3243925.

[4] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M.-S. Hacid, and H. Zeineddine, “An experimental study with imbalanced classification approaches for credit card fraud detection,” *IEEE Access*, vol. 7, pp. 93010–93022, 2019, doi: 10.1109/ACCESS.2019.2927266.

[5] A. Mniai, M. Tarik, and K. Jebari, “A novel framework for credit card fraud detection,” *IEEE Access*, vol. 11, pp. 112776–112786, 2023, doi: 10.1109/ACCESS.2023.3323842.

[6] H. Zhu, M. Zhou, Y. Xie, and A. Albeshri, “A self-adapting and efficient dandelion algorithm and its application to feature selection for credit card fraud detection,” *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 2, pp. 377–390, Feb. 2024, doi: 10.1109/JAS.2023.124008.

[7] B. Lebichot, T. Verhelst, Y.-A. Le Borgne, L. He-Guelton, F. Oble, and G. Bontempi, “Transfer learning strategies for credit card fraud detection,” *IEEE Access*, vol. 9, pp. 114754–114766, 2021, doi: 10.1109/ACCESS.2021.3104472.

[8] I. Sadgali, N. Sael, and F. Benabbou, “Human behavior scoring in credit card fraud detection,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 3, pp. 698–706, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp698-706.

[9] Y. Tang and Z. Liu, "A credit card fraud detection algorithm based on SDT and federated learning," *IEEE Access*, vol. 12, pp. 182547-182560, 2024, doi: 10.1109/ACCESS.2024.3491175.

[10] A. Priyadarshini, S. Mishra, D. P. Mishra, S. R. Salkuti, and R. Mohanty, “Fraudulent credit card transaction detection using soft computing techniques,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 23, no. 3, pp. 1634–1642, Sep. 2021, doi: 10.11591/ijeecs.v23.i3.pp1634-1642.

[11] L. Ni, J. Li, H. Xu, X. Wang, and J. Zhang, “Fraud feature boosting mechanism and spiral oversampling balancing technique for credit card fraud detection,” *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 1615–1630, Apr. 2024, doi: 10.1109/TCSS.2023.3242149.

[12] F. A. Ghaleb, F. Saeed, M. Al-Sarem, S. N. Qasem, and T. Al-Hadhrami, “Ensemble synthesized minority oversampling-based generative adversarial networks and random forest algorithm for credit card fraud detection,” *IEEE Access*, vol. 11, pp. 89694–89710, 2023, doi: 10.1109/ACCESS.2023.3306621.

[13] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, “Credit card fraud detection using adaboost and majority voting,” *IEEE Access*, vol. 6, pp. 14277–14284, 2018, doi: 10.1109/ACCESS.2018.2806420.

[14] Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, and M. Li, “Learning transactional behavioral representations for credit card fraud detection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 4, pp. 5735–5748, Apr. 2024, doi: 10.1109/TNNLS.2022.3208967.

[15] E. Ileberi, Y. Sun, and Z. Wang, “Performance evaluation of machine learning methods for credit card fraud detection using SMOTE and AdaBoost,” *IEEE Access*, vol. 9, pp. 165286–165294, 2021, doi: 10.1109/ACCESS.2021.3134330.

[16] I. D. Mienye and Y. Sun, “A deep learning ensemble with data resampling for credit card fraud detection,” *IEEE Access*, vol. 11, pp. 30628–30638, 2023, doi: 10.1109/ACCESS.2023.3262020.

[17] W. Ning, S. Chen, S. Lei, and X. Liao, “AMWSPLAdaboost credit card fraud detection method based on enhanced base classifier diversity,” *IEEE Access*, vol. 11, pp. 66488–66496, 2023, doi: 10.1109/ACCESS.2023.3290957.

[18] A. D. Pozzolo, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, “Credit card fraud detection: a realistic modeling and a novel learning strategy,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3784–3797, Aug. 2018, doi: 10.1109/TNNLS.2017.2736643.

[19] D. Lunghi, G. M. Paldino, O. Caelen, and G. Bontempi, “An adversary model of fraudsters’ behavior to improve oversampling in credit card fraud detection,” *IEEE Access*, vol. 11, pp. 136666–136679, 2023, doi: 10.1109/ACCESS.2023.3337635.

[20] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, “A neural network ensemble with feature engineering for improved credit card fraud detection,” *IEEE Access*, vol. 10, pp. 16400–16407, 2022, doi: 10.1109/ACCESS.2022.3148298.

[21] Y. Xie, G. Liu, C. Yan, C. Jiang, and M. Zhou, “Time-aware attention-based gated network for credit card fraud detection by extracting transactional behaviors,” *IEEE Transactions on Computational Social Systems*, vol. 10, no. 3, pp. 1004–1016, Jun. 2023, doi: 10.1109/TCSS.2022.3158318.

[22] Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang, "Credit card fraud detection based on improved variational autoencoder generative adversarial network," *IEEE Access*, vol. 11, pp. 83680–83691, 2023, doi: 10.1109/ACCESS.2023.3302339.

[23] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine," *IEEE Access*, vol. 8, pp. 25579–25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[24] C. Jiang, J. Song, G. Liu, L. Zheng, and W. Luan, "Credit card fraud detection: a novel approach using aggregation strategy and feedback mechanism," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3637–3647, Oct. 2018, doi: 10.1109/JIOT.2018.2816007.

[25] H. Tingfei, C. Guangquan, and H. Kuihua, "Using variational auto encoding in credit card fraud detection," *IEEE Access*, vol. 8, pp. 149841–149853, 2020, doi: 10.1109/ACCESS.2020.3015600.

BIOGRAPHIES OF AUTHORS

Dr. Vadhri Suryanarayana iD 8 SC ◐ was received Ph.D. in CSE from Acharya Nagarjuna University, post graduated in M.Tech. (CSE) from JNTU. He is having 27 years of experience in teaching in Engineering Colleges. He was held various positions from lecturer to vice-principal. Now, he is working as a professor in AI and DS and dean of industrial relations in Lakireddy Bali Reddy College of Engineering, Mylavaram. He was published 45 papers in national and international journals and also published 13 patents. He was authored for 5 technical books. He is also having 2 scholars under his supervision for Ph.D. He is a life member of various societies. He was honored with 2 awards for his best performance in research and guidance. He can be contacted at email: vadhri.abhiram@gmail.com.

Dr. Kuruva Maddileti iD 8 SC ◐ B.Sc. degree in mathematical science from the Sri Krishna Devaraya University, the M.Sc. degree in mathematics from The Sri Venkateswara University and the Ph.D. degree in mathematics from the EIILMU University. He used to hold several administrative posts with mathematics in Shanthinikethan College of Education, Rayala Seema University from 2015 to 2023. He completed B.Ed. from Sri Krishna Devaraya Univarsity, Ananthapur and M.Ed. from Rayalaseema University, Kurnool. He is currently an associate professor with the Department of Humanities and Basic Sciences Mathematics and Applied Mathematics in Dr. K.V. Subba Reddy Institute of Technology. He has authored six journals. He has research interest in mathematics. He can be contacted at email: kmaddiletibed@gmail.com.

Dune Satyanarayana iD 8 SC ◐ received the B.Sc. degree in computer science from the Andhra University, Visakhapatnam, Andhra Pradesh, India. The M.Tech. degree in computer science and engineering from Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. He used to hold several teaching posts with the school of computing. He worked as a lecturer in V.S. Lakshmi Educational Institutes at Kakinada from 2010 to 2014. He worked as an associated processor in Chaitanya institute of Science and Technology at Kakinada from 2015 to 2019. He was also director of the CARE (Computer Applications and Research Education) at Kakinada Centre from 2019 to 2020. He is currently working as an assistant professor in the Department of Computer Science and Engineering at Aditya College of Engineering and Technology, Surampalem, Kakinada District, Andhra Pradesh, India. Her research interests include OOPS, DBMS, and cloud computing. He can be contacted at email: satyaduneprof@gmail.com.

R Leela Jyothi iD 8 SC ◐ she currently serves as an assistant professor in the Department of CSE at SRKR Engineering College (A), located in Bhimavaram, Andhra Pradesh. Has 8 years of teaching experience. Coming to my areas of interests includes artificial intelligence, machine learning, and cloud computing. She can be contacted at email: rudraraju.leela92@gmail.com.

Kavuri Sreekanth is presently working as assistant professor in the Department of Computer Science and Engineering at Koneru Lakshmaiah Deemed to be University. Previously he worked as an employee in AP state government university and taught many subjects in Computer Science and Engineering Department and Information Technology Department. He completed his post graduation from JNTU Hyderabad. Pursuing research in the area of web mining. He can be contacted at email: kavurikanth@gmail.com.

Praveen Mande is currently working as an assistant professor in Department of Electrical Electronics and Communication Engineering, GITAM School of Technology at Gandhi Institute of Technology and Management, Visakhapatnam. His research interests include power system operation and control, smart grids and micro grids, electrical vehicles, power electronics, and power quality improvement using FACTS devices and its applications. He can be contacted at email: pmande@gmail.com.

Dr. Raghava Naidu Miriyala received his Ph.D. degree in computer science and engineering from Acharya Nagarjuna University, Guntur District, Andhra Pradesh, India. From 1999 to 2004 he was an assistant professor, from 2005 to 2012 he was associate professor from 2012 to 2018 he was a professor and director (i/c) in the Department of Master of Computer Applications. Since December 2018 he has been an assistant professor in computer science and engineering at Krishna University. He served as the Coordinator of the Computer Science and Engineering Department from 2019 to 2023. Where he taught in the Departments of Computer Science and Engineering and Electronics and Communication Engineering. He can be contacted at email: mrnaidus@gmail.com.

Oggi Sudhakar is assistant professor Department of Electronics and Communication Engineering in GIET Engineering College. He completed B.Tech in the year 2005 from Lenora College of Engineering, M.Tech (wireless and mobile communication) in the year 2009, from Vardhman College of Engineering. He has a diverse mix of teaching and research experience spanning about 15 years. Keen to actively seek and impart knowledge, he currently teaches optical communication, his area in research is low power VLSI. He can be contacted at email: sudhakar@giet.ac.in.