Real-time Credit Card Fraud Detection Using Machine Learning

Anuruddha Thennakoon[1], Chee Bhagyani[2], Sasitha Premadasa[3], Shalitha Mihiranga[4], Nuwan Kuruwitaarachchi[5]
Faculty of Computing
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
[1]anuruddha.thennakoon@gmail.com,[2]bhagyani.lochana@my.sliit.lk,[3]sasitha.premadasa@my.sliit.lk,
[4]shalitha.mihiranga@my.sliit.lk,[5]nuwan.ku@sliit.lk

***Abstract*—Credit card fraud events take place frequently and then result in huge financial losses [1]. The number of online transactions has grown in large quantities and online credit card transactions holds a huge share of these transactions. Therefore, banks and financial institutions offer credit card fraud detection applications much value and demand. Fraudulent transactions can occur in various ways and can be put into different categories. This paper focuses on four main fraud occasions in real-world transactions. Each fraud is addressed using a series of machine learning models and the best method is selected via an evaluation. This evaluation provides a comprehensive guide to selecting an optimal algorithm with respect to the type of the frauds and we illustrate the evaluation with an appropriate performance measure. Another major key area that we address in our project is real-time credit card fraud detection. For this, we take the use of predictive analytics done by the implemented machine learning models and an API module to decide if a particular transaction is genuine or fraudulent. We also assess a novel strategy that effectively addresses the skewed distribution of data. The data used in our experiments come from a financial institution according to a confidential disclosure agreement.**

Keywords— credit card frauds, fraud detection system, fraud detection, confidential disclosure agreement, real-time credit card fraud detection, skewed distribution.

I. INTRODUCTION

Fraud has been increasing drastically with the progression of state-of-art technology and worldwide communication. [5] Fraud can be avoided in two main ways: prevention and detection. Prevention avoids any attacks from fraudsters by acting as a layer of protection. Detection happens once the prevention has already failed. Therefore, detection helps in identifying and alerting as soon as a fraudulent transaction is being triggered. Recently, card-not-present transactions [6] in credit card operations have become popular among web payment gateways. According to the Nilson Report in October 2016, more than $31 trillion were generated worldwide by online payment systems in 2015, increasing 7.3% than 2014. Worldwide losses from credit card fraud have been rising to $21 billion in 2015, and will possibly reach $31 billion by 2020. [3] However, there has been an extreme increase in fraudulent transactions that affect the economy dramatically. Credit card fraud can be classified into several categories. The two types of frauds that can be mainly identified in a set of transactions are Card-not-present (CNP) frauds and Card-present (CP) frauds. Those two types can be described further by bankruptcy fraud, theft/counterfeit fraud, application fraud, and behavioural fraud. Our study aims at addressing four fraud natures that belong to the CNP fraud category described above and we propose a method to detect those frauds real time.

Machine learning is this generation's solution which replaces such methodologies and can work on large datasets which is not easily possible for human beings. Machine learning techniques fall into two main categories; supervised learning and unsupervised learning. Fraud detection can be done in either way and only can be decided when to use according to the dataset. Supervised learning requires prior classification to anomalies. During the last few years, several supervised algorithms have been used in detecting credit card fraud.

The data which is being used in this study is analyzed in two main ways: as categorical data and as numerical data. The dataset originally comes with categorical data. The raw data can be prepared by data cleaning and other basic preprocessing techniques. First, categorical data can be transformed into numerical data and then appropriate techniques are applied to do the evaluation. Secondly, categorical data is used in the machine learning techniques to find the optimal algorithm.

This paper consists of selecting optimal algorithms for the four fraud patterns through an extensive comparison of machine learning techniques via an effective performance measure for the detection of fraudulent credit card transactions.

The rest of this paper is presented as follows. Section 2 presents the literature review. Section 3 provides the experimental methodology including results. Finally, conclusions and discussions of the paper are presented in Section 4.

II. LITERATURE REVIEW

In earlier studies, many approaches have been proposed to bring solutions to detect fraud from supervised approaches, unsupervised approaches to hybrid ones; which makes it a must to learn the technologies associated in credit card frauds detection and to have a clear understanding of the types of credit card fraud. As time progressed fraud patterns evolved introducing new forms of fraud making it a keen area of interest for researchers. The remainder of this section describes single machine learning algorithms, machine learning models and fraud detection systems that were used in fraud detection. The problems that came across the review have analyzed for the

978-1-5386-5933-5/19/$31.00 ©2019 IEEE

later use of implementing an efficient machine learning model.

With the analysis of various detection models, past researchers have found many problems regarding fraud detection. In [14] and [3] they have mentioned Lack of real-life data as a huge issue. Real life data are lacking because of the data sensitivity and privacy issues. Papers [3] and [7] have studied Imbalance data or skewed distribution of data. The reason behind this is having quite a less amount of frauds when compared to non-frauds in the transaction datasets. Paper [3] states that data mining techniques take time to execute when dealing with big data. Overlapping of data is another major drawback in preparation of credit card transaction data. According to paper [2] and [7] the issue occurs due to some scenarios when the legitimate transactions look exactly like fraudulent transactions. In another way, fraudulent transactions may appear as legitimate transactions. Also, they have come across the difficulty in dealing with categorical data. When considering the credit card transaction data, most of the features have categorical values. In this case, almost all the machine learning algorithms do not support the categorical values. In [3][4] they have mentioned choice of detection algorithms and feature selection as a challenge in detecting frauds since most of the machine learning algorithms take much time for training purposes than predicting. Another key issue that affects financial fraud detection is the feature selection. It aims to filter out the attributes that most describes the aspects of fraud detection and its characters. In paper [7] they have highlighted fraud detection cost and lack of adaptability as challenges in the fraud detection process. When considering a system, the cost of fraudulent behaviour and the prevention cost should be taken into consideration. Lack of adaptability occurs when the algorithm is exposed to new types of fraud patterns and normal transactions. Effectiveness can change according to the problem definition and its specifications, so having a good understanding of the performance measure is necessary [4].

There are different kinds of models implemented for credit card fraud detections. In those models, different algorithms have been used.

Adapting the fraud detection system to newly introduced frauds can be problematic whether to retrain the machine learning model due to drastic changes in the fraud patterns, also may be costly and risky. For instance, Tyler *et al.* extended a framework proposed in [12], implemented the model and the model was applied to a real-world transaction log. To address the classification problem Logistic Regression (LR) has been used. The instances of fraudulent transactions have been discretized into strategies by using Gaussian Mixture Models (GMMs). Here synthetic minority oversampling technique was used to address the class imbalance. To stand out the significance of estimates in economic value sensitivity analysis has been used. The results have proven that a practical method which uses minimal steps to retrain a model could function as same as a classifier that typically retrains every round [13].

There is another model called Risk-Based Ensemble (RBE) that can handle the data consisting of issues and give outstanding results. For handling imbalanced data, a highly efficient bagging model has been used. To handle the implicit noise in the transaction dataset they have used Naive Bayes algorithm [9]. Peter *et al.* evaluated several deep learning algorithms with respect to their efficacy. The four topologies are Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), Long Short-term Memory (LSTMs), and Artificial Neural Networks (ANNs). In their project in addition to data cleaning and other data preparation steps, they have overcome class imbalance and scalability problems by using undersampling. To discover which hyper-parameters had the highest influence on the performance of the model, the sensitivity analysis was carried out. They have discovered that the performance of the model was affected by the size of the network. They concluded that larger the network it showed better performance. [11]

Credit card data have the issue of skewed distribution which is also known as the class imbalance. According to Andrea et al., their project addresses class imbalance including other issues such as concept drift and verification latency. They have also illustrated the most relevant performance matrix that can be used in credit card fraud detection. The achievement of the research also includes a formal model and a powerful learning strategy for addressing the 'verification latency' and an 'alert and feedback' mechanism. According to experiments they have declared the precision of the alerts as the most important measure [15].

Chee et al. used twelve standard models and hybrid methods which use AdaBoost and majority voting methods to achieve better accuracy rates in credit card fraud detection [16]. They were evaluated using both benchmark and real-world data. A summary of the strengths and limitations of the methods were evaluated. The Matthews Correlation Coefficient metric (MCC) has been taken as the performance measure. To evaluate the robustness of the algorithms noise was added to the data. Also, they have proved that the majority voting method was not affected by the added noise.

The analysis carried out on highly imbalanced data in paper [17] show that KNN shows outstanding performance for sensitivity, specificity and MCC, except for accuracy. The paper [18] discussed commonly used supervised techniques and they have provided a thorough evaluation of supervised learning techniques. Also, they have shown that all algorithms change according to the problem area.

Fraud detection system presented in paper [19] is built to handle class imbalance, the formation of labelled and unlabeled, and processing of large datasets. The proposed system was able to overcome all the challenges.

III. Experimental Methodology

A. Data description

The dataset was created combining two data sources; the fraud transactions log file and all transactions log file. The fraud transactions log file holds all the online credit card fraud occurrences while all transactions log file holds

all transactions stored by the corresponding bank within a specified time period. Due to the confidential disclosure agreement made between the bank and the authors of the paper, some of the sensitive attributes such as card number were hashed. When evaluating the combined dataset, the shape of the data was much skewed due to the imbalanced numbers of legitimate transactions and fraudulent occurrences. The file with the fraud cases had 200 records while the transaction log file had 917781 records. Attributes of the two data sources are as follows.

TABLE I. ATTRIBUTES OF THE GENUINE TRANSACTIONS LOG

Field Name	Description
CARD_NO	Credit cardholder's hashed card number.
DATE	Date of the transaction.
TIME	Time of the transaction.
TRANSACTION_AMOUNT	Transaction Amount.
MERCHANT_NAME	Merchant name relevant to the transaction
MERCHANT_CITY	Registered city of the Merchant.
MERCHANT_COUNTRY	Registered country of the Merchant.
RESPONSE_CODE	ISO Response code related to the transaction.
MERCHANT_CATEGORY_CODE	Category code of the Merchant.
APPROVED/NOT APPROVED	Status of the transaction.

TABLE II. ATTRIBUTES OF THE FRAUDULENT TRANSACTIONS LOG

Field Name	Description
CARD_NO	Credit cardholder's hashed card number.
DATE	Date of the transaction.
TIME	Time of the transaction.
SEQ	A unique sequence number which was given for frauds.
NATURE OF THE FRAUD	Nature of the frauds as card present or not present.
MCC	Category code of the Merchant.
AMOUNT OF FRAUD	Transaction Amount.
REVERSAL	Bank-related field.

B. Data preparation

Collected raw data were first divided into 4 data sets according to its fraud pattern. This process was done with the information gained by the bank. The four datasets are,

1. Transactions with Risky Merchant Category Code (MCC).
2. Transactions larger than $100.
3. Transactions with risky ISO Response code.
4. Transactions with unknown web addresses.

Those 4 datasets were used in two different ways.

1. By transforming raw data into a numerical form. (Type A)
2. By preparing raw data categorically without making any transformation. (Type B)

Type A was applied to datasets 1, 2, 3 and type B was applied to data set number 4. In data preparation the data are cleaned, transformed, integrated and reduced. First 3 sets of data were subjected to all above-mentioned steps to prepare them numerically. To prepare the data categorically all the steps except for data transformation were applied. The basic steps which were involved in the type A are described below.

- Data Cleaning - Filling in missing values is an important task in the data cleaning process. There are many ways to overcome this issue, such as ignoring the whole tuple, but most of them are likely to bias the data. Since the source file which contained genuine transactions did not contain records with missing values, filling them was no more an issue. Tuples with meaningless value were removed from the files as they do not contribute to producing important data as well as they do not bias the data. Additionally, following changes such as removing unnecessary columns, separating the date time column into two.

- Data Integration - Before the data were subjected to further change the two data sources were integrated together since fraudulent and genuine record files were in two separate files. Figure 1 shows how the mapping process was done.

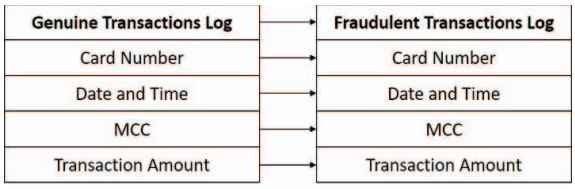

Fig. 1. Data Mapping

- Data Transformation - Here, all the categorical data were consolidated into an understandable numerical format. The transactional dataset contains several data types with several ranges. Therefore, data transformation comprises of data normalization. Data

normalization scales the attribute data to fall in a small numeric range.

- Data Reduction - The strategy used for this is Dimension reduction. We must prevent the risk of learning wrong patterns of data and the selected features should eliminate the irrelevant aspects and qualities of the fraud domain [10]. The principal component analysis which is well-known that PCA is a popular transform method. Applying this method resolves the feature selection issue from the perspective of numerical analysis. PCA performed feature selection successfully by finding the suitable number of principal components.

In type B, data cleaning and data integration were involved as same as in type A. Then those data were taken to the next step of the process.

C. *Resampling Techniques.*

The two data sources were characterized by a highly imbalanced distribution of examples among the classes. Fraudulent transactions contained a much smaller number of examples than the genuine transactions. To overcome this, we conducted under-sampling and over-sampling by reducing the majority occurrences and by raising the minority occurrences respectively. For over-sampling, Synthetic Minority Oversampling Techniques (SMOTE) and for under-sampling, condensed nearest neighbour (CNN) and random under-sampling (RUS) were used. The minority class is over-sampled by producing "synthetic" examples in SMOTE method [20]. Out of RUS and CNN, RUS is a non-heuristic method which balances the class distribution by using a method to eliminate random majority class examples [21] [22].

Additionally, we have used 10-fold cross-validation. Then the data to which the cross-validation was applied, were resampled by the above-mentioned resampling techniques.

D. *Modelling and testing*

Our study analyses four different fraud patterns. For analyzing each pattern, we have reflected the following process as described in figure 2. Quite a few numbers of techniques were used in the data analysis. Four machine learning algorithms were prioritized in our analysis with the help of the literature. They are Support Vector Machine, Naive Bayes, K-Nearest Neighbor and Logistic Regression. We applied those selected supervised learning classifiers to our resampled data. When selecting machine learning models which can capture each fraud, the accuracy and performance of each model were taken into consideration. Optimal models were selected by filtering them out comparatively against an appropriate performance matrix (Table 3).

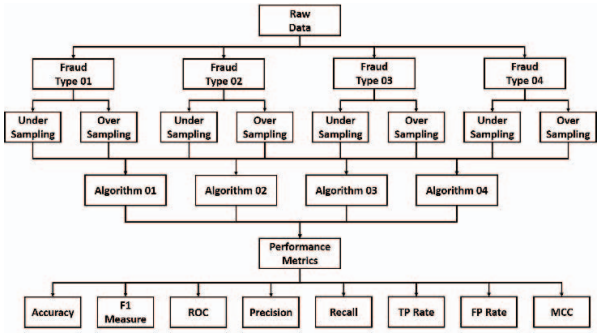

Fig. 2. Model Selection

TABLE III. PERFORMANCE METRICS

Measure	Formula
Accuracy	TN + TP/TP + FP + FN + TN
Precision	TP/TP + FP
Recall	TP / (TP + FP)
True positive rate	TP/TP + FN
False positive rate	FP/FP+TN
F1-measure	2 × (Precision ×Recall)/ (Precision +Recall)
ROC	The TP rate against the FP rate
MCC	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$

TP = True Positive, TN =True Negative, FP = False Positive, FN = False Negative

The following graphs show the accuracy rates from the 4 types of fraud when the ML classifiers are applied to preprocessed and resampled data.

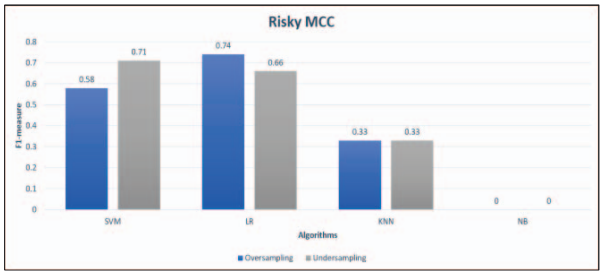

Fig. 3. Risky MCC Results

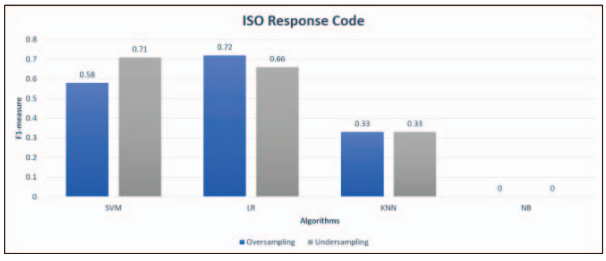

Fig. 4. ISO-Response Code Results

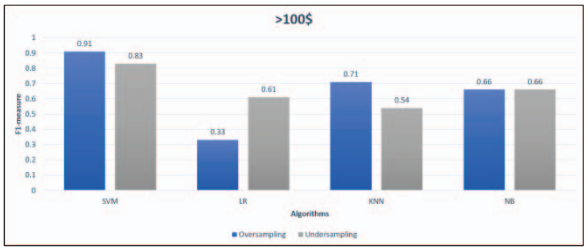

Fig. 5. >100$ Results

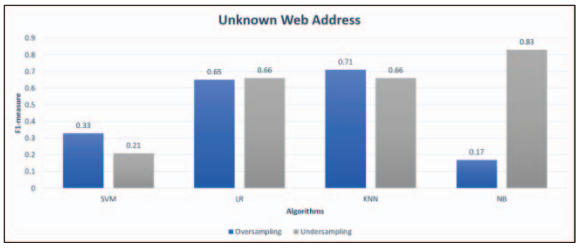

Fig. 6. Unknown web address results

E. *Real time Fraud Detection.*

In the past, fraud detection has been done by taking already happened transactions in bulk and applying machine learning models on them. Since the results can be seen after weeks or months, tracking down of detected frauds was found extremely difficult, and there have been many cases where the fraudsters were able to commit many more fraudulent purchases before being exposed. Real-time fraud detection is the execution of fraud detection models the second an online purchase is taken place. That way our system is capable of detecting frauds real-time. It gives an alert to the bank indicating its fraud pattern and accuracy rate, making it easy for fraud monitoring teams to move into their next action without having to waste their time and money.

F. *Fraud Detection System.*

Real-time detection of credit card fraud can be stated as one of the main contributions of this project. The real-time fraud detection system consists of three main units; API MODULE, FRAUD DETECTION MODELS and DATA WAREHOUSE. All the components are involved in fraud detection simultaneously. Fraudulent transactions are being classified into four fraud types (Frauds occur due to Risky MCC, ISO-Response Code, Unknown web address, Transaction above 100$) using three supervised learning classifiers. API module is responsible for transferring real time transactions between the Fraud detection model, GUI, and Data warehouse. A Data Warehouse has been used for storing live transactions, the predicted results and other important data of the machine learning models. The user can interact with the fraud detection system with GUIs where it shows the real time transactions, alerts regarding frauds and historical data regarding frauds in a graphical representation. When a transaction is recognized as fraudulent by the fraud detection model, a message will be sent to the API module. Then the API module will notify the end user by sending a notification and the feedback given by the end user will be stored. Figure 7 describes the overall system flow of the fraud detection system.

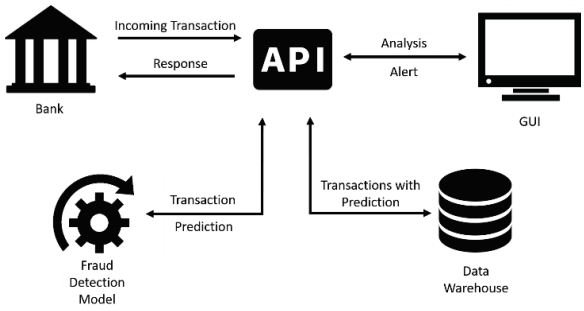

Fig. 7. System Diagram

IV. Conclusions

Credit card fraud detection has been a keen area of research for the researchers for years and will be an intriguing area of research in the coming future. This happens majorly due to continuous change of patterns in frauds. In this paper, we propose a novel credit-card fraud detection system by detecting four different patterns of fraudulent transactions using best suiting algorithms and by addressing the related problems identified by past researchers in credit card fraud detection. By addressing real time credit-card fraud detection by using predictive analytics and an API module the end user is notified over the GUI the second a fraudulent transaction is taken place. This part of our system can allow the fraud investigation team to make their decision to move to the next step as soon as a suspicious transaction is detected. Optimal algorithms that address four main types of frauds were selected through literature, experimenting and parameter tuning as shown in the methodology. We also assess sampling methods that effectively address the skewed distribution of data. Therefore, we can conclude that there is a major impact of using resampling techniques for obtaining a comparatively higher performance from the classifier. The machine learning models that captured the four fraud patterns (Risky MCC, Unknown web address, ISO-Response Code, Transaction above 100$) with the highest accuracy rates are LR, NB, LR and SVM. Further the models indicated 74%, 83%, 72% and 91% accuracy rates respectively. As the developed machine learning models present an average level of accuracy, we hope to focus on improving the prediction levels to acquire a better prediction. Also, the future extensions aim to focus on location-based frauds.

References

[1] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and G. N. Surname, “Random Forest for credit card fraud,” 15th Int. Conf. Networking, Sens. Control, 2018.

[2] M. Zareapoor, S. K. . Seeja.K.R, and M. Afshar Alam, “Analysis on Credit Card Fraud Detection Techniques: Based on Certain Design Criteria,” Int. J. Comput. Appl., vol. 52, no. 3, pp. 35–42, 2012.

[3] David Robertson, “Investments & Acquisitions — September 2016 Top Card Issuers in Asia–Pacific Card Fraud Losses Reach $21.84 Billion,” Nilson Rep., no. 1096, 1090.

[4] J. West and M. Bhattacharya, "An Investigation on Experimental Issues in Financial Fraud Mining," *Procedia Comput. Sci.*, vol. 80, pp. 1734–1744, 2016.

[5] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance Evaluation of Class Balancing Techniques for Credit Card Fraud Detection," IEEE Int. Conf. Power, Control. Signals Instrum. Eng., pp. 2747–2752, 2017.

[6] G. Liu, W. Luan, Z. Li, and Y. Zhang, "A new FDS for credit card fraud detection based on behavior certificate," 2018.

[7] Z. Zojaji, R. E. Atani, and A. H. Monadjemi, "A Survey of Credit Card Fraud Detection Techniques : Data and Technique Oriented Perspective," pp. 1–26, 2016.

[8] Suman and Nutan, "Review Paper on Credit Card Fraud Detection," Int. J. Comput. Trends Technol., vol. 4, no. 7, 2013.

[9] S. Akila and U. S. Reddy, "Risk based Bagged Ensemble (RBE) for Credit Card Fraud Detection," no. Icici, pp. 670–674, 2017.

[10] D. P. Methods, "Data Preprocessing Techniques for Data Mining," Science (80-.)., p. 6, 2011.

[11] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, and P. Beling, "Deep Learning Detecting Fraud in Credit Card Transactions," pp. 129–134, 2018.

[12] M. F. Zeager, A. Sridhar, N. Fogal, S. Adams, D. E. Brown, and P. A. Beling, "Adversarial learning in credit card fraud detection," *2017 Syst. Inf. Eng. Des. Symp.*, pp. 112–116, 2017.

[13] T. Cody, S. Adams, and P. A. Beling, "A Utilitarian Approach to Adversarial Learning in Credit Card Fraud Detection," pp. 237–242, 2018.

[14] M. Rafało, "Real-time fraud detection in credit card transactions," *Data Science Warsaw*. 2017.

[15] A. Dal Pozzolo, G. Boracchi, O. Caelen, and C. Alippi, "Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy," *Ieee Trans. Neural Networks Learn. Syst.*, pp. 1–14, 2018.

[16] 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. XX, pp. 1–1, 2018.

[17] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," *2017 Int. Conf. Comput. Netw. Informatics*, pp. 1–9, 2017.

[18] R. Choudhary and H. K. Gianey, "Comprehensive Review On Supervised Machine Learning Algorithms," *2017 Int. Conf. Mach. Learn. Data Sci.*, pp. 37–43, 2017.

[19] G. E. Melo-Acosta, F. Duitama-Muñoz, and J. D. Arias-Londoño, "Fraud detection in big data using supervised and semi-supervised learning techniques," *Commun. Comput. (COLCOM), 2017 IEEE Colomb. Conf.*, pp. 1–6, 2017.

[20] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer,

"SMOTE: Synthetic minority over-sampling technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002.

[21] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance Evaluation of Class Balancing Techniques for Credit Card Fraud Detection," *IEEE Int. Conf. Power, Control. Signals Instrum. Eng.*, pp. 2747–2752, 2017..

[22] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance Evaluation of Class Balancing Techniques for Credit Card Fraud Detection," IEEE Int. Conf. Power, Control. Signals Instrum. Eng., pp. 2747–2752, 2017.