IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Credit Card Fraud Detection using Machine Learning

Bhagyashri R Hanji
Computer Science Engineering.
Dayananda Sagar Academy of Technology and Management
Bengaluru, India.

Akarshan Kumar
Computer Science Engineering.
Dayananda Sagar Academy of Technology and Management
Bengaluru, India.

Akash Roy
Computer Science Engineering.
Dayananda Sagar Academy of Technology and Management
Bengaluru, India.

Ayman Saleem
Computer Science Engineering.
Dayananda Sagar Academy of Technology and Management
Bengaluru, India.

Ayush Chandak
Computer Science Engineering.
Dayananda Sagar Academy of Technology and Management
Bengaluru, India.

***Abstract*—** In this research paper, we present a novel Credit Card Fraud Detection System that leverages machine learning-based feedback. Leveraging machine learning as a revolutionary method, the study explores the adaptation of fraud detection methods to data characteristics, specifically distinguishing between supervised and unsupervised machine learning techniques. Our research evaluates the performance of various machine learning algorithms, including random forests, distribution trees, neural networks, support vector machines, naive Bayes, logistic regression, and gradient boosting classifiers. Notably, logistic regression models excel in binary classification problems, offering accurate scores for each transaction and enabling the identification of potential fraud patterns based on historical transaction data. In addition, the study proposes a supervised classifier, addressing the challenges posed by a lack of publicly available information and the overwhelming prevalence of fraud cases compared to regular transactions. The study evaluates under-sampling and over-sampling methods such as random under-sampling and near miss sampling, meticulously assessing their effectiveness across classifiers to mitigate the negative consequences of fraudulent transactions and improve overall model performance.

Keywords*— *Credit Card Fraud Detection, Classification Method, Supervised Learning, Cyber security.

I. INTRODUCTION

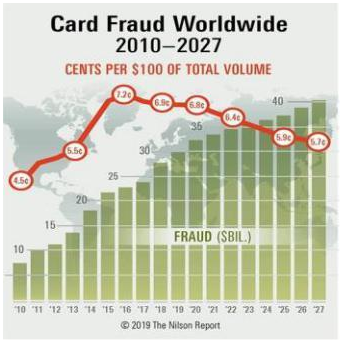
In today's world, a single touch can have big consequences. These conveniences, such as taking the bus, interacting with the virtual assistant, getting recommendations, following the recipe, ordering food to your door, are provided by restoring computing power and sharing IT infrastructure. Advances in technological capabilities have led to the creation of vast amounts of data. The simultaneous rise of artificial intelligence (AI) and machine learning (ML) is a direct result of the data explosion. What many people don't realize is that we now rely heavily on machine learning in our daily lives. A prime example of this is integrating machine learning technology into credit card fraud systems, thereby strengthening our payment system.

This research article provides a comprehensive analysis of credit card fraud in today's digital environment influenced by online platforms that store various important information in the cloud. The article highlights the growing concern about fraud and emphasizes the importance of communication technology, especially given the proliferation of online credit card transactions [1]. As a revolutionary method [2]. machine learning goes beyond traditional methods and processes big data that humans cannot directly access [3]. This article distinguishes between supervised and unsupervised machine learning techniques and demonstrates the adaptation of fraud detection methods to data characteristics. Despite progress, information gaps are still a problem and lead to fraud as most businesses are legitimate. This work demonstrates an insurance card system that uses recommendations from machine learning to improve detection and overall performance We implement the random credit data to evaluate a variety of classification techniques, including random forests, distribution trees, neural networks, support vector machines, naive Bayes, logistic regression, and gradient boosting classifiers. This article draws on a number of research papers to discuss the rise of credit card fraud in electronic payments, the importance of data science and machine learning in credit card fraud, machine learning algorithms, and artificial neural networks for classification and comparison of algorithms such as random forest and Adaboost. Research results are presented and potential avenues for further development are presented, including feedback, short-term memory improvement, and interactive modules.

II. BACKGROUND AND LITERATURE *A. Credit Card Frauds*

Our dependence on mobile phones and the internet has increased tremendously over the last few years, leading to a rise in online payments. Unfortunately, the growth of the digital economy has also led to an increase in credit card fraud, resulting in significant revenues for financial institutions worldwide and additional stress for credit cardusers. This project is essential due to the escalating occurrence of credit

card fraud in digital environments, necessitating advanced fraud detection methods using machine learning. The pressing need to address the growing concern about fraudulent activities and to fortify defenses against sophisticated fraud strategies underscores



the significance of this research. As of 2017, there are approximately 20.48 billion cards worldwide, including credit cards, debit cards, and all prepaid cards .

Fig. 1. Card Fraud Worldwide

Fig. 1 Provides an overview of global card fraud totals and $100 losses as of 2027 globally. It can be seen that credit card fraud reached nearly $30 billion in 2019 and is expected to increase each year, but the $100 rate is expected to decrease. The Australian Payments Network's annual report shows credit card use is on the rise in Australia, with fraud reaching $574 million. Credit card fraud increased 10% overall, with card not present (CNP) fraud accounting for 84.9% of all fraud. To address this issue, the Australian government has introduced a number of anti-fraud measures, including the CNP Fraud Reduction Framework, which sets standards for manufacturers and traders. Additionally, partnerships such as the Australian Payments Council's partnership with the Cyber Security Center also aim to collect data effectively. The government is also regulating EMV chips, blocking chargebacks to other cards and encouraging financial institutions to implement fraud detection systems.

B. Fraud Detection using Machine Learning

In recent years, the important role of machine learning in data processing, especially in detecting card fraud, has gained importance. Existing studies, such as those presented in the Genetic Algorithm and Logistic Regression literature [1], propose a variety of methods to detect fraud and understand its various types, including eight combinations of supervised, unsupervised, and proprietary. Artificial Neural Network (ANN), Genetic Algorithm (GA), Support Vector Machine (SVM), Lightweight Object Mining (FISM), Decision Making (DT), Migratory Bird Optimization Algorithm (MBO) and Naive Bayes Procedure etc. Technology NB has been proposed and tested in the field.

The goal of card fraud continues to be to identify the behavior of the card at the time of purchase using techniques that include a lot of transportation, unknown Bayesian analysis, Bayesian and nervous system output evaluation. Decision trees, machine learning, and logistic regression are an important part of fraud detection, as reviewed in the Genetic Algorithms and Logistic Regression paper. This paper explores new learning techniques such as support vector machines, random forests, and logistic regression, and aims to improve fraud detection using neural networks and logistic regression as described in random forests and Adaboost files. Detecting credit card fraud faces challenges due to complex fraud patterns, inconsistent and limited data, optimal sample selection, and efforts to test fraud strategies. The effectiveness of credit card fraud detection, influenced by sampling approaches, parameter choices, and identification techniques, is addressed in the Supervised learning algorithms paper. Various types of credit card fraud, including physical card theft and stealing confidential information, pose threats to financial transactions. Businesses leverage machine learning techniques, as explored in the Optimized Light Gradient Boosting Machine, to identify suspicious transactions, especially with the increasing use of credit cards in online and regular transactions. Traditional manual methods for fraud detection become impractical with the advent of big data, prompting financial companies to adopt intelligent methods based on computing intelligence (CI), including supervised and unsupervised techniques. In this data analysis paper, experts primarily examine a hybrid data model with three levels of functionality choice and heuristic classification, drawing insights from genetic algorithms, data gain ratio, and evaluation of recovery characteristics, as detailed in the Genetic algorithm and logistic regression.

Logistic regression plays an important role in detecting credit card fraud and provides a powerful analytical tool for measuring the likelihood of business fraud. In this case, logistic regression models perform well on binary classification problems, making them ideal for distinguishing between legitimate and fraudulent. One of their strengths is their ability to provide accurate and well- defined scores for each transaction, thus simplifying the decision- making process to prevent fraud. Logistic regression uses historical transaction data to learn patterns that indicate fraud, including variables such as currency, location, and time. This model evaluates the impact of each predictor on the probability of fraud, providing insight into the relative importance of different features. In addition, logistic regression is powerful enough to handle unequal data; This is a challenge in detecting credit card fraud, where the number of legitimate transactions outweighs fraud. The simplicity and interpretation of logistic regression have made it an important part of machine learning techniques used by financial institutions, providing a good way to improve the accuracy and efficiency of fraud systems in the face of ever- changing fraud strategies. The imbalanced nature of credit card data, with more legitimate transactions than fraud, poses challenges to prediction accuracy, a concern addressed in the Random Forest and Naive Bayes paper. Class allocation strategies, involving oversampling of minority classes, help mitigate this issue. Supervised optimization techniques may fall short in detecting fraud cases, leading to exploration of deep autoencoder and restrained Boltzmann machine designs, integrated into a hybrid technique with AdaBoost and Majority Voting, as discussed in the Long Short term memory paper [8].

Credit card fraud has become prevalent in the digital environment, particularly with the rise of e-commerce. Attackers use a variety of deception techniques and methods that must be constantly investigated. Past research such as neural network modeling, Bayesian networks, intelligent decision engines, optimization algorithms, meta-learning agents, artificial intelligence, image processing, rule-based systems, logistic regression, support vector machines, decision trees, k-nearest neighbors, transformational learning and many more are making changes to the environment. Unsupervised techniques such as self- configuring neural network models are widely used in instant payment applications to provide solutions to social problems and achieve results. Random forests have become a powerful tool in detecting credit card fraud, unraveling the complexities of the ever- changing system with remarkable results. Random forests contain a set of decision trees and are good at handling the complexity of rogue scenarios by

combining the outputs of many individual trees. Each decision tree measures characteristics such as transaction amount, frequency, and location, allowing the model to capture negative patterns that indicate legitimate and fraudulent behavior. The combination of different forests increases power, reduces competition, and increases access to new, unprecedented information. Additionally, the ability to perform fundamental analysis helps identify the most useful features that help detect fraud. This is especially true in the dynamic credit card fraud environment where fraudsters are constantly changing their strategies. Random forests are effective in handling contradictory data common in fraud detection, where there are more cases of legitimate transactions than cases of fraud. The integrated model has the ability to quickly analyze large amounts of data, which is important for immediate fraud prevention. The versatility and adaptability of Random Forests make them the core of an arsenal of machine learning techniques that allow financial institutions to strengthen their defenses against the complex strategies used by fraudsters in the ever-expanding world of digital marketing [4].

The application of Adaboost (i.e. transaction support) to credit card fraud has been successful and increased efficiency. This learning algorithm creates a robust and accurate model by combining multiple weak classifiers. In terms of fraud detection, Adaboost plays an important role in solving problems caused by inconsistent information in the credit card market. This algorithm assigns different weights to individual data points and assigns the highest priority to error cases in each iteration. This weight change ensures that subsequent iterations are incorrect before previous events are correctly classified. Adaboost often incorporates decision trees or other classifications into its categories when detecting credit card fraud. The combination of these techniques allows the algorithm to learn and adapt to the complexity of fraud patterns; This results in a good model for distinguishing legitimate transactions from deception. By leveraging its focus on complex examples, Adaboost has been instrumental in improving the accuracy and reliability of credit card fraud systems, thereby increasing financial security in the energy sector.

C. Literature Survey

This extensive literature survey provides an in-depth exploration of diverse methodologies and cutting-edge technologies within the dynamic field of mental health solutions. The research conducted by Ileberi, E., Sun, Y. & Wang, Z. investigates the critical need for effective credit card fraud detection methods, given the surge in online transactions and the heightened activities of fraudsters targeting credit card transactions. The study employs supervised machine learning algorithms such as Decision Tree, Random Forest, Artificial Neural Network, Naive Bayes, and Logistic Regression, using a credit card fraud dataset generated from European credit cardholders. This work is essential due to the imperative to develop models capable of accurately detecting credit card fraud, thereby safeguarding users from financial losses.

The document by Asha RB, Suresh Kumar KR presents an overview of credit card fraud detection, identifying various classifications of credit card fraud and emphasizing the rising need for robust fraud detection models. The proposed system aims to leverage Artificial Neural Networks (ANN) supplemented by classification algorithms like Support Vector Machine and k-Nearest Neighbor for credit card fraud detection. The research compares the performance of these three algorithms, determining that ANN outperforms the others. The dataset used in the experiment consists of 31 attributes, with the last attribute indicating the transaction outcome.

The research provided by Mr. Kapil Dev Tripathi, Mr.Vikash Singh Rajput addresses the escalating issue of fraudulent credit card transactions (CCTs) by employing a Randomized Search CV with XGB Classifier to accurately predict fraudulent activities. It underlines the growing financial costs incurred by companies and consumers due to fraudulent transactions and the necessity for effective solutions in fraud identification. The study utilizes a large volume of credit card fraud (CCF) data from the UCI repository, conducting simulations through the use of Python in a Jupyter notebook. The proposed model's significance is evaluated based on various performance parameters, including accuracy, precision, recall, F1 score, MCC, and ROC, ultimately achieving an 83.02% accuracy rate.

The work by S P, Maniraj & Saini, Aditya & Ahmed, Shadab & Sarkar, Swarna stands out, offering valuable insights into detecting anomalous activities, particularly focusing on fraudulent transactions. It employs the Local Outlier Factor and Isolation Forest Algorithm for outlier detection. The dataset used for the experiment is obtained from Kaggle, containing 31 columns, with sensitive data masked, and includes details such as time, amount, and transaction class. The analysis includes various graphs to visualize inconsistencies, transaction times, and transaction amounts. The dataset is processed and standardized, with the removal of the "Class" column to ensure fairness of evaluation. The study utilizes the Local Outlier Factor and Isolation Forest Algorithm from the sklearn ensemble module for outlier detection. The paper provides a detailed explanation of the modules, including pseudocodes for their algorithms and output graphs. The research employs Jupyter Notebook in Python to demonstrate the proposed approach, and the program can also be executed on the cloud using the Google Colab platform. The Local Outlier Factor and Isolation Forest Algorithm are thoroughly described, along with their pseudocodes and associated output graphs.

The research provided by Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou and M. Li, addresses the challenge of detecting credit card fraud by developing a novel model that creates improved representations of transaction records. The proposed approach enhances long short-term memory with a time-aware gate to capture changes in user behavior across consecutive transactions. Additionally, it introduces a current-historical attention module to establish connections between present and past transactional behaviors, enabling the model to recognize behavioral periodicity. An interaction module is also included to gain comprehensive and sensible behavioral representations. The research validates the effectiveness of the learned behavioral representations through experiments conducted on a large real-world transaction dataset from a financial company in China and a public dataset. The results demonstrate that the proposed method effectively distinguishes between legitimate and fraudulent behaviors, outperforming existing state-of-the-art methods in fraud detection.

III. PROPOSED SUPERVISED CLASSIFIER

Researching machine learning (ML) techniques in credit card fraud faces a major challenge: a lack of publicly available information due to the urgency of understanding financial information and protecting consumer privacy. This limitation forces research in this area to rely on a single data set for analysis, raising concerns about the

generalizability of the findings. The effectiveness of machine learning models, the fact that they are based on a large amount of data with different features and different products, and offer many differences, make the research process important to analyze these changes in our data. This study focuses on the business scenario of diversifying the subtle behavior of the model. Adding another layer of difficulty is credit card fraud, where the number of fraud cases exceeds that of regular businesses. It is quite difficult to solve this dilemma and this study aims to see the impact of various models on the behavioral pattern in this context. ML algorithms such as Support Vector Machine (SVM), Decision Making (DT) and Logistic Regression (LR) have traditionally been applied in credit card analysis.

However, their scalability to large data sets raises concerns. To address these questions, this work explores the uncharted territory of deep learning and specifically examines the connection between convolutional neural networks (CNN) and time, short-term memory networks (LSTM). These methods have been praised for their effectiveness in image classification and natural language processing (NLP) and have been reviewed for their applicability in bad credit classification. In addition, this study examines significant preliminary information. - Work on ML pipeline. It is important to examine how distribution efficiency affects the credit card fraud context. Logistic regression plays an important role in this research and this research expands its method by incorporating the use of fuzzy logic. This addition aims to improve model accuracy and interpretation, providing a comprehensive study of the many challenges and opportunities of machine learning, particularly fuzzy logistic regression for credit card fraud.

A. Logistic Regression

Analyzing credit card fraud using logistic regression involves collecting historical transaction data, including attributes such as payment, location, and time. Data processing included imputation of missing values, statistical modeling, and coding of categorical variables. The dataset is divided into training and testing, which ensures that the model performs well from the analysis of invisible data. This approach provides a good basis for identifying business risks associated with fraud in the credit card industry.

The first important step in using logistic regression to detect credit card fraud involves data preparation. This includes credit card history; Each entry is classified as LEGAL or fraudulent based on historical data. The dataset is divided into two parts: one for legitimate trading and the other for fraudulent trading. Logistic regression uses key features such as exchange rate, location, and time as dependent variables. The LEGIT array contains instances of legitimate transactions that constitute the good class, while the FRAUD array contains instances of casts that represent the bad class. Run a logistic regression model on this data to identify potential patterns and relationships between input devices and proximity. During training, the model optimizes its ability to make correct decisions. Logistic functions (also known as sigmoid functions) help identify potential credit card fraud with higher accuracy by transforming the output into a binary function between 0 and 1.

After training, the model is validated on a test model to evaluate its performance. Metrics such as precision, recall, and F1 score measure the model's performance in detecting fraud and reducing false positives. Factor analysis is used to explain the impact of different characteristics on fraud detection. Logistic regression models are praised for their simplicity, interpretability, and data processing abilities, making them invaluable in combating credit card fraud. Continuous monitoring and integration of new profile models is essential to ensure continued impact and effectiveness in reducing the risk of credit card fraud. This application demonstrates the potential of logistic regression as a tracking learning algorithm in credit card fraud, enabling financial institutions to identify legitimate transactions, legitimate operations, and fraud as inferred from historical data structures.

To improve the performance of logistic regression in fraud, feature engineering, including creating new information changes or modifying existing information, has proven useful. Variables such as frequency shift, scale factor, and correlation by negative scores help improve the model's ability to identify negative behavioral patterns. Physical integration, such as time of day and day of the week, can improve the model's ability to detect fraud patterns based on trading hours. Considering the high rate of fraud, consistent maintenance and order is essential. To increase the flexibility of the model, update the model regularly to include new variables and information. Logistic regression allows joint or hybrid models to be combined with other machine learning methods to leverage different algorithms to provide strong protection against credit card changes. Using logistic regression for credit card fraud requires a strategic approach that includes specialized engineering skills, continuous model development, and regulatory compliance. This holistic approach provides a strong and flexible defense against credit card fraud.

IV. SYSTEM ARCHITECTURE

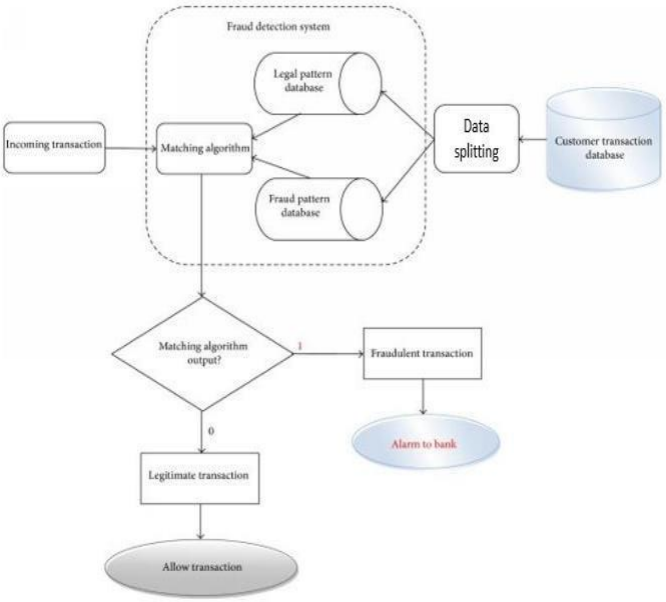


Fig: System architecture of proposed system.

A pivotal figure emerges as the custodian of a sophisticated customer transaction database fortified with a cutting-edge fraud detection system. This guardian meticulously oversees the vast sea of transactions, ensuring the integrity of the financial ecosystem. Upon encountering a transaction, the data undergoes a meticulous process of data splitting. The system then deftly applies logistic regression, a statistical method specifically crafted for fraud detection. Logistic regression is a type of machine learning

algorithm that is well-suited for binary classification problems, such as determining whether a transaction is fraudulent (1) or legitimate (0). The algorithm has been trained on historical data, learning patterns that distinguish between genuine and potentially fraudulent transactions. It evaluates the input data and calculates the probability that a given transaction belongs to the fraudulent class.

In this intricate dance between data and logistic regression, the custodian stands as a stalwart defender, safeguarding the financial well-being of countless individuals and institutions alike. The logistic regression model, with its ability to analyze and interpret complex data patterns, plays a crucial role in fortifying the fraud detection system and ensuring the security of the financial transactions it scrutinizes.

V. DATASETS

The main purpose of this study is to evaluate the effectiveness of products in credit card fraud by using data with different models and features. European Card Data (ECD) is especially used for this purpose. Similar to many credit card fraud reports, ECD exhibits high levels of volatility with fewer instances of fraud compared to traditional transactions. In the sample code of the data set, "0" is shown in case of no fraud, and "1" is shown in case of fraud. This study aims to investigate the nuances of the classifier's performance, with a particular focus on handling inconsistent data, as demonstrated in Eurocard data.

A. European Card Data

This data, generously provided by the Machine Learning Group of the Université Libre de Bruxelles and obtained by Kaggle, captured two-day European credit cards in September 2013. In this study, the data includes 284,807 samples with 31 factor characteristics. Importantly, only 492 samples were identified as fake, accounting for only 0.172% of the data. To protect customer privacy and flexibility, everything in the configuration file except "time" and "price" is changed by the analytics manager (PCA). The "time" property represents the number of seconds since the first instance, and the "quantity" property represents the total amount of work. This dataset serves as a valuable resource for evaluating credit card fraud detection methods in the study.
https://www.kaggle.com/mlg-ulb/creditcardfraud

VI. EVALUATION METRICS

Given the lack of class weighting in the literature, accuracy proved unsuitable as a benchmark in this study. The measurement method chosen depends on the solution; The main goal for detecting credit card fraud is to detect all fraud while minimizing the downside (business risk misclassified as fraud). To facilitate this evaluation, use the conflict matrix by assigning illegal examples (according to the rules) to bad examples and bad examples. A negative result represents the fact that the truth is not a fraud, a positive result represents the fact that the truth is a fraud, a negative result represents the fact that a fraud is not considered a fraud, and a lie is just as bad a representation as a representation of a lie. - Lie. To gain a deeper understanding of the evaluation process, the study considers the equation of accuracy, precision, recall, and F1 score and validates performance metrics that are important for credit card fraud procedures.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

As shown in Equation 1, accuracy is related to good predictive value, and reducing the number of false positives can improve accuracy. Precision is a critical metric when the error rate is high. The equation 3 values returned relate to true positive results and reduce the negative for the return. Situations where the value of the negative is high are often important for reaching higher consciousness. The balance between good and bad is important when investigating credit card fraud. Predicting the entire model based on false positives produces higher returns but lower accuracy, precision, and F1 scores. In contrast, estimating all unfalsified models resulted in maximum but zero regression with unidentified true and F1 scores. This study uses all four metrics (accuracy, accuracy, recall, and F1 score) to provide a comparison of credit card fraud.

VII. HANDLING CLASS IMBALANCE

Class imbalance poses a challenge in labeled datasets where instances are unevenly distributed into majority and minority classes. In credit card transactional data, fraudulent transactions, though a rare minority among millions, exert significant impacts on users, merchants, and issuers. The challenge in imbalanced datasets is the potential for biased performance favoring the majority class and misclassifying the minority class as noise. Various methods, including data sampling, cost-sensitive learning, one-class learning, and ensemble learning, aim to address this issue. This study delves into the nuanced realm of under-sampling and over-sampling methods, exploring their effectiveness across classifiers and meticulously evaluating performance on sample data for each dataset. This strategy aims to reduce the negative consequences of fraudulent transactions and improve the overall performance of the model, making it effective in real situations where fraud cases are rare but disturbing. The sampling methods used are as follows:

A. Random Under Sampling

Random under-sampling (RUS) involves reducing the number of events in a class by selecting from the dataset. This approach minimizes class conflict by focusing on class in general. Implementing random under-sampling in Python makes the use of the RandomUnderSampler class in the learning library unstable. The "sampling_strategy" parameter plays an important role and determines the majority of classes for minority classes. In this way, the sample selection process supports the equivalent class by reducing the number of cases in most categories and eliminates inconsistencies in the data set related to credit card analysis.

B. Near Miss Sampling

In random under-sampling (RUS), selecting events at random can result in important information being omitted from the data set. To solve this problem, near index (NM) sampling uses distance criteria to guide sample selection. There are three versions of the closure: Version 1, Version 2, and Version 3. After evaluating the

effectiveness of all three versions, Version 1 was chosen as the option for study 1. Version 1 selects the most common examples with the smallest mean relative to the three closest classes. The implementation in Python uses the NearMiss class from the non-parallel learning library to be more efficient when storing important data.

VIII. CONCLUSIONS AND FUTURE WORK

Credit card fraud poses a threat to financial institutions and calls for fraud reform. The machine learning method is the main factor in reducing the accuracy and negativities of credit card fraud, and its results are affected by many factors such as the size of the business and getting along. Deep learning methods, including convolutional neural networks (CNN) and short-term neural networks (LSTM) designed for image and natural language processing, have shown better performance compared to traditional methods. In particular, LSTM with 50 blocks achieved the highest F1 score of 84.85%. This research shows the uncertainty in solving problems for different classes.

Although there are many ways to improve the current model, no data corruption was found and increasing levels did not improve performance. Future research focuses on advances in deep learning. As a traditional algorithm, logistic regression is always better than other algorithms with an efficiency of up to 98% in credit card fraud detection. Its simplicity, translation and computational efficiency make it an attractive choice. Logistic regression's ability to determine the importance of fraud by providing insight into more important factors makes it a powerful tool, especially clear and critical explanation. As part of future work, fuzzy logic is combined with logistic regression to create a hybrid method designed to increase the flexibility and efficiency of the model in the process. Understanding the intricacies of credit card fraud.

IX. REFERENCES

[1] Ileberi, E., Sun, Y. & Wang, Z. “A machine learning based credit card fraud detection using the GA algorithm for feature selection”. *J Big Data* **9**, 24 (2022). https://doi.org/10.1186/s40537-022-00573-8.

[2] S P, Maniraj & Saini, Aditya & Ahmed, Shadab & Sarkar, Swarna. (2019). “Credit Card Fraud Detection using Machine Learning and Data Science”. International Journal of Engineering Research and 08. 10.17577/IJERTV8IS090031.

[3] Asha RB, Suresh Kumar KR, “Credit card fraud detection using artificial neural network”, Global Transitions Proceedings, Volume 2, Issue 1, 2021, Pages 35-41, ISSN 2666-285X, https://doi.org/10.1016/j.gltp.2021.01.006.

[4] R. Sailusha, V. Gnaneswar, R. Ramesh and G. R. Rao, "Credit Card Fraud Detection Using Machine Learning," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 1264-1270, doi: 10.1109/ICICCS48265.2020.9121114.

[5] S. Khatri, A. Arora and A. P. Agrawal, "Supervised Machine Learning Algorithms for Credit Card Fraud Detection: A Comparison," 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2020, pp. 680-683, doi: 10.1109/Confluence47617.2020.9057851.

[6] A. A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine," in IEEE Access, vol. 8, pp. 25579-25587, 2020, doi: 10.1109/ACCESS.2020.2971354.

[7] Trivedi, Naresh & Simaiya, Sarita & Kumar Lilhore, Dr & Sharma, Sanjeev. (2020). “An Efficient Credit Card Fraud Detection Model Based on Machine Learning Methods”. MATTER: International Journal of Science and Technology.

[8] Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou and M. Li, "Learning Transactional Behavioral Representations for Credit Card Fraud Detection," in IEEE Transactions on Neural Networks and Learning Systems, 2022, doi: 10.1109/TNNLS.2022.3208967.