REPUBLIC OF TURKEY

ALTINBAS UNIVERSITY

Institute of Graduate Studies

Information Technologies

EMPLOYING MACHINE LEARNING TECHNIQUES AND FUZZY MEMBERSHIP FOR DETECTING FRAUD TRANSACTIONS IN CREDIT CARD

Ahmed ABDULGHANI

Master of Science

Supervisor

Prof. Osman Nuri UCAN

Istanbul, 2022

EMPLOYING MACHINE LEARNING TECHNIQUES AND FUZZY MEMBERSHIP FOR DETECTING FRAUD TRANSACTIONS IN CREDIT CARD

Ahmed ABDULGHANI

Information Technologies

Master of Science

ALTINBAŞ UNIVERSITY

2022

The thesis titled " EMPLOYING MACHINE LEARNING TECHNIQUES AND FUZZY MEMBERSHIP FOR DETECTING FRAUD TRANSACTIONS IN CREDIT CARD " prepared by AHMED ABDULGHANI and submitted on 10/2/2022 has been **accepted unanimously of votes** for the degree of Master of Science in Information Technologies

Prof. Osman Nuri UCAN

Supervisor

Thesis Defense Jury Members:

Prof. Dr. Osman Nuri UÇAN	Faculty of Engineering and Architecture, Altinbas University	_____
Asst.. Prof. Selin ÖZBEK ŞİMŞEK	Faculty of Engineering and Architecture, Altinbas University	_____
Assoc. Prof. Adil DENİZ DURU	Faculty of Sports Science Marmara University	_____

I hereby declare that this thesis meets all format and submission requirements of a master thesis

Submission date of the thesis to the Graduate Education Institute: ___/___/___

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Ahmed ABDULGHANI

Signature

DEDICATION

To the absent who will not return is my beloved mother and her pure soul that accompanies me. It is impossible to thank you enough for everything you have done. I hope to meet you in heaven To my father, for his endless love, encouragement, and support. He taught me noble values, a love of benevolence, and helping people To the soul of my brother Omar To my sisters and my brother To my wife and my two children To my friends and everyone who loves me.

ACKNOWLEDGEMENTS

First of all, thanks and praise be to God for granting me success in completing this path. Thanks to my supervisor Prof. Osman Nuri UCAN, I am proud that this thesis by his supervision. Thanks to my co-supervisor, Dr. KHATTAB M.ALI. You have helped me overcome any obstacle that I consider impossible. You helped me identify, accept and improve my weaknesses. You are a true teacher, and many students' lives will be changed forever because of your kindness and passion. So much thanks must be given to my mother, my mother's pure soul, my father, my father's tender heart, my beloved wife Marwa, my children Yamin and Taleen, my brother Dr.Mohammed, my sisters Rania and Saja, and my colleague & my sister Zainab, I am forever indebted to you all. So thanks to my undergraduate classmate and close family friend Hanaa Khudhur for giving me the help and advice. I wish her a good and happy life. Thanks to my friend Abo obaida for supporting me. Finally, thanks to my manager Sameer Mahdi and my co-workers in my job. Thank you to my friends and comrades and everyone who helped me and extended a helping hand.

ABSTRACT

EMPLOYING MACHINE LEARNING TECHNIQUES AND FUZZY MEMBERSHIP FOR DETECTING FRAUD TRANSACTIONS IN CREDIT CARD

Abdulghani, Ahmed,

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Prof. Dr. OSMAN Nuri UCAN

Date: 01/2022

Pages: 76

Credit card fraud is becoming an increasingly prevalent problem in today's financial sector. An alarming rise in the number of fraud-related activities has been observed in the past few years, which has resulted in significant financial losses for numerous organizations, businesses, and government bodies. Because the numbers are projected to grow in the future, many researchers in this discipline have centered their efforts around identifying fraudulent behavior early on using advanced machine learning algorithms, which are becoming increasingly popular. However, the identification of credit card fraud is not easy for many reasons, including the fact that the fraudulent behaviors vary from one attempt to the next and the available datasets in this field are very few. In this thesis, a system is proposed to detect fraudulent transactions based on some methods of machine learning such like Logistic regression (LR), NaiveBayes (NB), as well as the Linear Discriminant Analysis(LDA), in addition to XGBoost algorithm. the techniques used to generate the models of the suggested system, they were trained and evaluated on the basis of two different datasets. Where the first dataset was the European Cardholders, and the second dataset was the Turkish dataset provided by the Yapi Kredi company. These datasets suffer from a big imbalance problem. Therefore, we used SMOTE to fix this issue. The system models' effectiveness was

measured employing a variety of metrics, specifically the confusion matrix, accuracy, F1score, recall, precision as well as AUC. Also, the fuzzy membership function was adopted to the dataset in order to raise the system's efficiency. The final results showed the high efficiency of XGBoost.

Keywords: Credit Card Fraud Detection, Supervised Machine Learning, Imbalanced Dataset, SMOTE, Fuzzy Membership.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

CCF	:	Credit Card Fraud
ML	:	Machine Learning
LR	:	Logistic regression
LDA	:	Linear Discriminant Analysis
NB	:	Naïve Bayes
SMOTE	:	Synthetic Minority Oversampling Technique
AUC	:	Area Under the ROC curve
TP	:	True Positive
TN	:	True Negative
FN	:	False Negative
FP	:	False Positive

1. INTRODUCTION

credit card is frequently utilized for expediting corporate transactions within today's global economy. Credit cards have been the subject of cyberattacks and fraud because of their widespread use. In order to efficiently deal with violations and unauthorized users, more security is needed. Crooks and fraudsters frequently threaten credit card networks who use unethical tactics to conduct unauthorized or illegal transactions [1]. As a result, fraud detection methods must be extremely quick to detect and able to keep up with rapid changes in fraud methods and their ever-changing strategies in order to be effective [2].

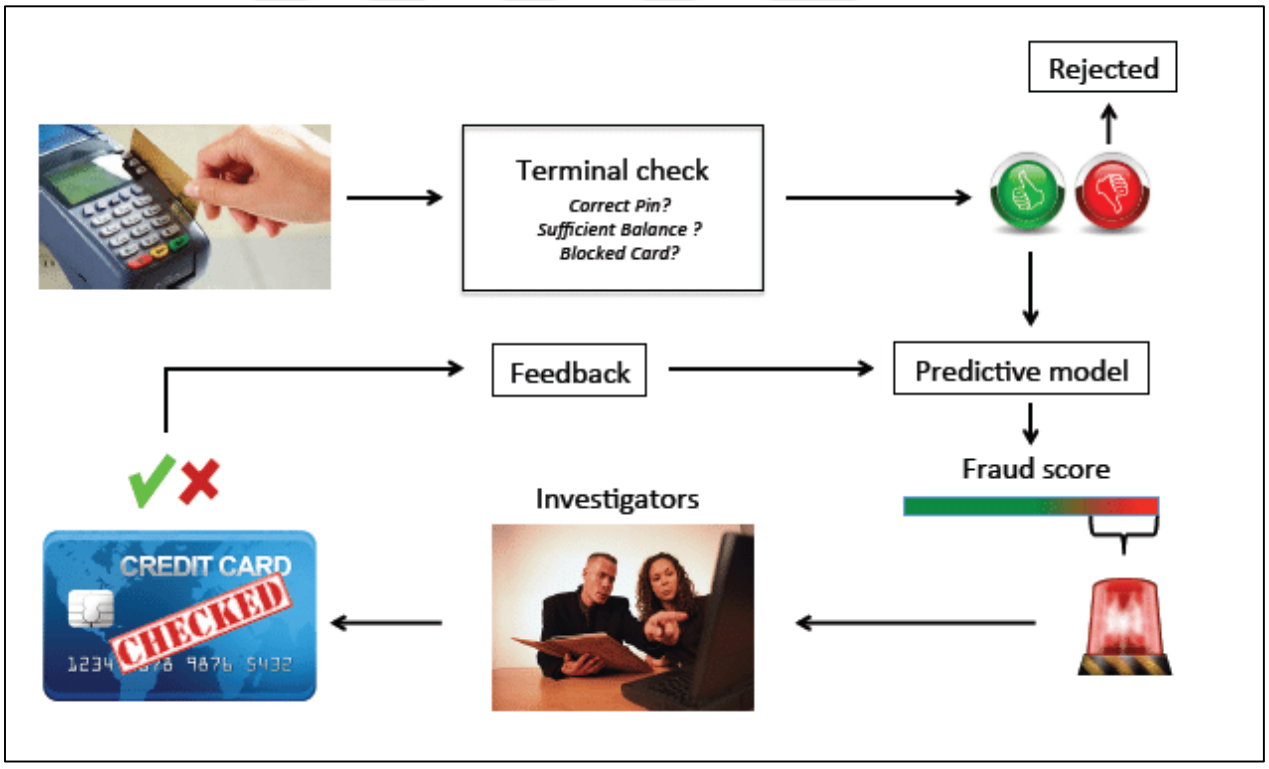

Figure 1.1: CCF Detection Operation [3]

The detection systems that are now available and in use rely on a variety of approaches and tools, including artificial intelligence techniques, statistics, and data mining techniques. The detection method follows a logical sequence, as seen in Figure 1.1; initially, a terminal check is performed on the card. Then, if the transaction is not refused, it is passed on to a predictive algorithm that

generates alerts for questionable transactions. The investigators enhance the predictive model's accuracy by offering feedback on alerts [2].

Card fraud (CCF) occurs via two main situations, offline or online. The robbery of Credit cards is a feasible strategy for offline fraud. The Internet is used to commit online fraud through hacking operations on computers, phones, and other networked devices and the use of credit card information for a variety of purposes.

When it comes to the proliferation of electronic commerce and its rapid growth in several last years, the use of a credit card is required for the services accessing. Thanks to its services, many economic operations are made easier, which have grown in number over time. Despite this, numerous fraud and counterfeiting methods lead to large financial losses for individuals, institutions, and banks. Therefore, its use has negative and serious downsides. Figure 1.2 depicts the financial losses incurred by the CCF over the previous few years and the predicted losses for the foreseeable future, according to the study in [4],

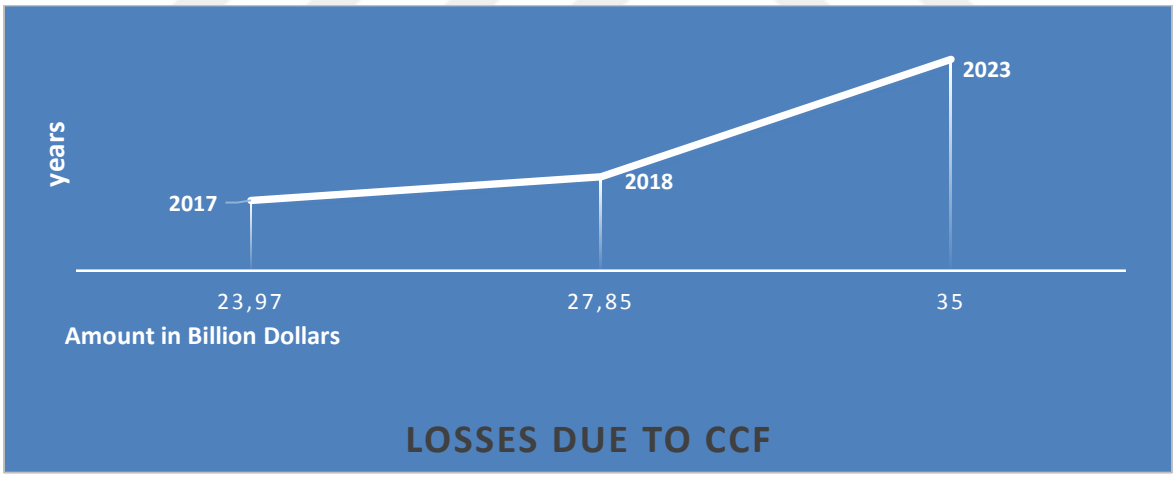

Figure 1.2: Economic Losses Linked to CCF [4]

This report demonstrated that fraud resulted in enormous economic losses worldwide, as the ratio of 27.85 billion dollars in 2018 was more than the proportion of 23.97 billion dollars in 2017, indicating an increase in the percentage year over year. Furthermore, it is expected that the share of financial losses caused by CCF will continue to rise, reaching 35 billion dollars in 2023 and possibly more [4].

1.1 MOTIVATIONS

greater emphasis must be placed on applications for monitoring and monitoring frauds and attempting to prevent and avoid them to lower the amount of financial losses caused by CCF. Fraud detection will help minimize consumer issues and complaints, thereby earning their trust. Currently, CCF detections are heavily reliant on Machine Learning (ML) approaches[5]. Numerous methods in ML can be utilized to solve CCF problems[6][7], it encompasses both supervised [8][9] and unsupervised [10] additionally the semi-supervised learning[5][10].

CCF has garnered considerable interest among banks. Data mining was applied to facilitate the detection of fraud and the identification of the true motive for each transaction. The most effective option is to employ statistical algorithms in an effort to search for any signals of fraud based on the data that is currently accessible[11]. The underlying concept of CCF detection is that every action taken by the card is analyzed and interpreted. There are a variety of approaches available for detecting and characterizing CCF, like the support vector machine(SVM)[12], the genetic algorithm[13], the decision tree[14], the artificial neural network[15], as well as the naïve Bayes (NB)[16].

1.2 RELATED WORKS

This section will provide a quick overview of numerous previous research that have dealt with the subject of CCF detection in the past. These are the studies that were conducted:

C.Sudha and D. Akila (2021) proposed Fraud Detection System for Credit Cards that could be developed using SVM and RF classifiers. Users' operational features are initially extracted using this system, which then uses an RF classification to determine whether or not a characteristic is legitimate or suspect. Extracting transactional features from user logs is the second phase, followed by classifying the features into lawful and suspect by the SVM classifier developed by M. The suggested system's performance has been assessed via the standard measures of accuracy,

precision, recall, as well as F-1 score. The results showed that both RF and SVM classifiers were able to attain a high detection rate with good accuracy[17].

Dileep et al. **(**2021) investigated CCF. Through the use of algorithms that incorporate machine learning approaches. Two algorithms were used: RF and Tree Decision. The model's efficiency is determined using certain publicly available data. The examination is then conducted on a sample of data gleaned from a financial establishment. Additionally, the data samples are augmented with a clatter to ensure the system's robustness. The approaches utilized in the paper have the following meaning: in a first way, a tree is formed against the user's activities, and tree scams are suspected. The second method creates a forest-based on user activity and makes attempts to identify the suspect through its use. The research findings demonstrate unequivocally that the standard, elective approach achieves acceptable degrees of accuracy [18].

Tran and Dang (2021) proposed that an imbalanced dataset be balanced using two different sampling ways: SMOTE as well as Adaptive Synthetic (ADASYN). On balanced data, the following techniques are used: RF, KNN, LR, and Decision Tree. Comprehensive classification metrics, involving fundamental, combined, also visual indicators, were employed to appraise the performance. After the process of dataset resampling, the authors remark that the ML algorithms are referred to produce positive results for detecting fraudulent activities [19].

Janvitha et al. (2021) worked by analyzing previous transactions and the amount spent on those transactions to establish the transaction's status, utilizing HMM (Hidden Markov Models) and different clustering techniques. This model is trained to forecast using the cardholder's behavior pattern. Acceptance of an incoming transaction is contingent upon the likelihood of being satisfied. To begin, historical transactions must be evaluated, and the cardholder's profile's concealed

pattern. Second, these patterns must be sorted into clusters with comparable properties in order to identify fraud and genuine instances. Based on previous transactions with Hmm, this model helps determine whether an individual's new transaction is valid or fake. When combined with Hmm, the study identifies which clustering method can reliably predict the transaction state to prevent fraud. Thus, real transactions are never denied on the grounds of fraud in a collection of transactions [20].

Asha and Kumar (2021) developed a method for detecting CCF using Deep Learning. Additionally, they compared it to machine learning techniques such as SVM, K-nn, and ANN for predicting CCF. They have employed neural networks, even if it is difficult to train a model sufficient for CCF detection. This model is suited for detecting CCF since it utilizes (ANN), which has an accuracy of almost 100%. It is more precise than unsupervised learning techniques. Under-sampling is applied to address the challenges raised from an imbalanced dataset. Also, this study applied data pre-processing and normalization [21].

Du Shaohui et al. (2021) created a model for detecting CCF utilizing the RF algorithm. Using the IEEE CIS fraud data, the study's findings reveal that this model's technique is superior to model benchmarks such as LR and SVM. Finally, this model's accuracy was 97.4 percent, while the AUC ROC score was 92.7 percent [22].

X.Kewei et al. (2021) used Deep Learning to propose a more accurate and effective fraud detection method. They have used various techniques to improve this model’s performance, such as feature engineering, memory compression. Also used mixed-precision as well as ensemble loss. This model's training and evaluation use a dataset of more than a million records. This data is the IEEE-

CIS fraud dataset from Vesta Corporation. This model showed high results in performance when compared with conventional technologies such as SVM, NB [23].

Deng. Wenkai et al. (2021) produced a study that applies RF and manual detection to transaction fraud detection systems. The study results of the data of IEEE CIS indicate that this model's procedure is superior to the model benchmark, like LR and SVM. the accuracy of this model gave was 96.8%, while the score for the AUC ROC was 92.5%[24].

Shivanna et al. (2020) designed and developed a credit card detection system through two ML algorithms: the first classifier was Decision Forest and the second was Decision Jungle classifier. Authors have used the European cardholder dataset to model that new system. The results provided by the Decision Jungle were the best [25].

Sailusha et al. (2020) presented a project focused on using ML algorithms to detect CCF, specifically the random forest and AdaBoost algorithms. The results were evaluated separately for each algorithm relying on the accuracy, recall, F1score in addition to precision. Also, using the confusion matrix for drawing the ROC curve. Moreover, after comparing the results, Random Forest was preferred [26].

Jain et al. (2020) applied the following ML algorithm techniques: random forest, decision tree, and XGBoost to a bank card fraud data set. The algorithms provided similar results to each other, and they were all highly accurate. The best accuracy went to the XGBoost algorithm[27].

Nur-E-Arefin, Md and Sultan, Mohammad (2020) presented a study that analyzed and compared several classifiers that have different algorithms to detect credit card fraud. Then it became clear that the meta and tree classifiers performed better than their peers. The accuracy of the classification process was high, but despite that, the fraud detection process was at a weak level.

Performance measurement and evaluation are the major focus of this study by focusing on the accuracy of the classification and the rate of fraud detection [28]. This study also examines the performance and efficiency of classifiers based on a dataset taken from UCSD-FICO Data Mining Contest 2009.

S. Khatri et al. (2020) several supervised ML models were tested against an imbalanced dataset to see if they could accurately predict the likelihood of fraudulent transactions. In order to arrive at a given result, sensitivity and precision, as well as time, are considered. Accuracy was never utilized as a parameter since it is not sensitive and does not provide a conclusive response to unbalanced data. This research looked at the kNN, NB, Decision Tree (DT), LR, as well as RF models. These models were applied to forecast the probability of CCF. Using the DT model, they discovered that it is the best method for detecting this type of fraud. Analysis revealed kNN models' sensitivities to be greater than DT, but because kNN takes a long time to evaluate the data, they decided to utilize Decision Trees instead. As a result, using a DT was the most effective way to recognize CCF transactions. However, the Tree Model's performance must be examined using unsupervised ML if it is to yield a strong convincing outcome [29].

Saheed et al. (2020) suggested to use the Genetic Algorithm (GA) to implement the feature selection method. This procedure was divided into 2 phases. First, selected eight attributes as the most appropriate attributes. In the second step, selected another eight appropriate attributes. Used the following ML technologies: RF, SVM, additionally to NB. The dataset utilized was a German dataset, it is imbalanced dataset. The results showed the significant effect of the eight features that were selected. Aside from that, the results revealed that RF had higher accuracy and detection rates than the other algorithms tested [30].

Alhazmi and Aljehane (2020) presented a survey about the use of ML in the detection of CCF. They clarified the difficulties in the CCF detection process, conducted a review of numerous studies on this subject, and explained the techniques used in Yemen in this matter. The characteristics of fraud detection and comparison methods are mentioned, whether good or bad. Two forms of fraud detection, either supervised or unsupervised, mention the techniques. They cited the results of their experiments [31].

Fayaz Itoo et al. (2020) compared the ability of ML to determine how accurately fraud and legitimate transactions of credit card datasets are differentiated and classified by the method of RUS (random under-sampling) and to verify whether or not they have improved their performance. Three ML methods are used: LR, NB, and KNN. The LR results were the best. Accuracy, sensitivity, specificity, precision, F-measurement, as well as the area under the roc are the measurements of the performance of the algorithms. The project was performed in python [32].

Taha and Malebary (2020) proposed a methodology for detecting CCF depending on an optimized light gradient boosting machine (OLightGBM). In this work, two real data sets were used in several experiments. The evaluation was done after comparing results from other research and modern ML algorithms such as RF, LR, Radial SVM, Linear SVM, decision tree, KNN, and NB. The results showed that an effective parameter optimization strategy is important and valuable to enhance the predictive performance of the approach proposed [33].

Sriram Sasank et al. (2019) proposed a framework that balances the dataset. They used the following ML algorithms on a balanced dataset: RF, NB, SVM, KNN, and LR. Some sampling techniques were also applied for balancing operation, including Oversampling, Underdamping,

Both sampling, ROSE, and SMOTE. The AUC - ROC efficiency metric indicates that the performance of LR shows 97.04% in accuracy and 99.99% in precision [34].

Olawale Adepoju et al. (2019) examined SVM, NB, LR, and K-NN performance on exceptionally distorted CCF data. The performance examination relies on precision, sensitivity, accuracy, and specificity. The comparison results showed that LR has higher accuracy than other algorithms[35].

Pradheepan Raghavan and Neamat El Gayar (2019) benchmarked various ML methods like KNN, RF, and SVM, whereas deep learning methods like, convolutional neural networks, auto-encoders, restricted Boltzmann machine in addition to the deep belief networks. The datasets of European, Australian also German are the data sets that have been used. The performance evaluation relies on: First, the AUC, secondly Cost of Failure, and Thirdly, Matthews Correlation Coefficient (MCC)[36].

1.3 PROBLEM STATEMENT

Credit card fraud causes huge financial losses annually, and these losses are increasing year after year. So banks and institutions strive to build systems capable of detecting fraud and reducing it as much as possible. Despite these attempts, the issue requires greater efforts and continuous attempts. However, working in this field faces many challenges and problems, most notably:

i. The detection of fraud in credit cards is increasingly complex, as fraudulent card transactions are increasingly similar to legal ones.

ii. The lack of good datasets. The data sets available to researchers are unbalanced, so that the details of data fraud are often much less than the real class details. This makes understanding the dataset difficult for programmers to build the best model to solve the problem.

iii. Data overlap is another problem because many transactions appear to be fraudulent despite being legal, and also, many transactions appear legal despite being fraudulent.

1.4 THESIS OBJECTIVES

This study aims to reduce the economic losses of banks, institutions, and people resulting from credit card fraud by proposing a system capable of detecting fraud based on machine learning techniques.

1.5 MAIN CONTRIBUTIONS

First, this study contributes to reducing credit card fraud and improving the ability to detect fraud by proposing a system that works with the following Machine Learning methods LR, LDA, NB, as well as XGBoost. The techniques were used to create models qualified to detect the fraud even using a dataset that has an imbalance status. Moreover, the ability to develop the system efficiency by the fuzzy membership. Another contribution, the system can be employed for other classification tasks and more complex datasets.

1.6 THESIS STRUCTURE

Five chapters make up this thesis. The lines that follow give a brief about each chapter:

Chapter one: [introduction] is an introduction to the thesis. Chapter Two: [Background Information] presents the CCF detection background.
Chapter Three: [The Proposed System] explains the steps of the suggested system and the techniques used. A full description of the suggested system is given. Chapter Four: [Results and Discussion] contains the results of tests that have been implemented to estimate the system performance. The results of experimental examinations are discussed. Chapter Five: [Conclusions and Future Works] Concludes the thesis and provides directions for future works.

2. BACKGROUND INFORMATION

A credit card is a popular payment method in numerous emerging and developed nations. Inventing credit cards enables seamless, easy, convenient, and convenient online transactions. However, it has provided criminals with new scams and, in turn, increased swindles. CCF has worrying global implications, and so many enterprises and persons have lost millions. Moreover, cybercriminals are regularly innovative with sophisticated techniques. Therefore, improved and dynamic techniques must be developed urgently, which can be adapted to evolving patterns of fraud. It is a very daunting task to accomplish, mainly because the swindle is dynamic and because researchers do not have enough data set [37].

In this chapter, there is a description of credit card fraud, its types, and the challenges faced in the detection. Also there is an explanation of the techniques that are used in detection.

2.1 CCF

CCF is the fraudulent use of card details to make an online purchase without the cardholder's consent. Card activities can sometimes have carried performed physically or virtually. for the case of physical the buyer and seller engage physically. At the time of purchase, users must present a physical card. Virtual transactions mean internet or telephone transactions. For online purchases, users are required to focus on providing some info, including the CVV number, user password, and the answer to a particular question and others [38]. In addition to making online purchases easier, more comfortable, and more convenient, credit cards have also given criminals additional ways to commit fraud and raised the rate of fraud [11][39]. Credit card fraud has a disturbing effect and has significantly impacted the global economy. Many individuals and businesses have lost millions of US dollars. In 2009, online order values totaled approximately US$15 billion (for goods and services only). In addition, Additionally, 84 percent of these transactions were made online [39]. The scam was worth an estimated at around 23 billion dollars for US retailers in 2013 and around 32 billion dollars for fraud in 2014[37]. One of the main causes of CCF is weak credit and debit card security. In the United Kingdom, the estimated cost of non-present fraud was 183.2 million £ in 2011[40]. Card-not-present fraud is a danger for every credit card user, and retailers bear the expenses of abnormal transactions[37].

Fraud detection in credit cards is an issue of classification. The Luhn algorithm is used to generate credit card numbers. Users are not protected from online fraud by the algorithm, but it does help them authenticate their data. To protect their customers, some small businesses employ manual authentication measures, such as phone number validation, physical address verification, as well as a particular question with the answer. However, for large companies, these procedures are impractical; it is expensive and inefficient. In addition, most online businesses employ the Card Verification Value (CVV2) as an extra security step when approving card transactions that are not physically present. Although card-not-present fraud has not been reduced to a reasonable minimum by this added layer of security, it does not prevent scam caused by lost and stolen cards. The Address Verification Service can handle scam with a card that is not physically present. It is a digitized tool that checks transactions through the use of a cardholder's shipping address. While this strategy lowers scams, it results in a lack of selling, as some customers do not want to ship what they buy to directly the billing address [40].

Moreover, the Master, as well as VISA Card, have implemented a secure 3D protocol in the field of online banking. Include the secure code of the MasterCard and the VISA Verified. These protocols are authenticated via a digital certificate for online merchants and for customers using passwords [40]. Fraudsters use the Internet to commit fraud often, as it is easy to conceal their identities and locations[41]. Customers, as well as traders, suffer losses incurred from credit card fraud. Merchants carry most losses, and the customers are charged higher tax costs & membership dues. Also, traders lower the value of their incentives as well as promotions [11]. Fraud detection is critical to lowering losses for both financial organizations and consumers. Fraud detection systems have the primary goal of promptly identifying fraud[39]. A transaction normally involves four parties: the user, the trader, the finance enterprise, and also the VISA Center is located [42]. Security is required for all these parties. The majority of presently available scam detecting systems are rule-based [43]. Because rules are built relying on understood shapes, these systems are capable of identifying known scam shapes only; but the unknown or emerging patterns cannot be detected. On average, to discover a fraudulent transaction requires approximately 72 hours [44]. However, it is only useful for detecting false card fraud; it is not good for detecting lost or stolen card fraud. So for solving the problem, developers of fraud detection systems should be aware behavior of the fraudster and the user [43].

2.2 CCF DETECTION

The detection of fraud is a problem of data mining to separate transactions into different segments legal and fraudulent [43]. The recently developed fraud-sensing systems that are used by merchants and banking institutions aim to check transactions through the pattern of customers' expenditures and behavior [44]. For this, prediction algorithms are used by fraud detection systems in the classification of pattern observations [11]. The system will flag a transaction as fraudulent if it detects a deviation from the user's usual spending habits. Listed below are a few CCF detection methods [44]:

i. By using the customer's postal code, transaction verification is performed via the Address Verification System (AVS).

ii. Transaction verification is accomplished by the Card Verification Method, which relies on the customer's personal private number.

iii. Transaction verification is carried out using a secret number given by the client, and this is done through the use of the Personal Information Number (PIN).

iv. By using the customer's fingerprint, transaction verification is performed via biometrics.

There are two main ways for tackling fraud: prevention as well as detection. Prevention of fraud is intended to prevent fraudulent activity, while fraud detection identifies fraudulent transactions thus preventing them Permission for the transaction. The detection of fraud begins after the system fails to block the fraudster's action [41].

2.2.1 The Effective CCF Detection System

Fraud at high risk should be reduced to the lowest possible level by an effective fraud detection solution. Large sums of money are lost due to high-risk fraud. The entire credit limit is the most

common target when a credit card is stolen. Fraudsters can swoop in and snatch the money in 4-5 transactions [43]. Some benefits of an effective fraud detection system are as follows [37]:

i. The system should be capable of dealing with skewed distributions.

ii. The system should handle noise.

iii. The system must distinguish between fraudulent and lawful transactions that are quite similar in appearance.

iv. A strong fraud detection system should be versatile and flexible enough to accommodate variations in fraudulent shapes.

v. Evaluation of fraud detection requires a specific classification method as measurements such as accuracy will be incompetent in case of imbalanced distribution.

2.2.2 CCF Detection Challenges

Many challenges and problems are facing the detection process in this matter. The first of these challenges is the difficulty in being able to distinguish between real and fake activities, as most of the activities appear real. Secondly, the information card transactions are usable and unbalanced from time to time. Thirdly, the accuracy of CCF identification is unbelievably affected by the type of testing, variable choice, and identification methods applied. Fourthly, data are changing continuously over time, changing it over time and making frequent conduct and fraud behaviour, which in the past were true and might be actual fraud or vice versa[35]. Fifthly, one of the most important problems is the lack of realistic data. The scarcity of available data is a result of financial enterprises' fears regarding their client's sensitive info as well as the breach of their secrecy [45]. Also, the acceptance or refusal method must work in a very short period of time. Additionally, multiple transactions occur concurrently, so, each transaction cannot be monitored independently and hence fraud is determined. Therefore, an effective system of fraud detection must be

implemented to differentiate between a lawful and a phony kind. The system should be based on the user behavior card learning concept [29].

2.2.3 CCF Detection Features

The features of credit card detection system building classifiers are divided into three[44] :

i. Account-specific features. Which includes everything related to account matters such as the account number and type (whether the account is in dollars or other currencies), as well as the date of opening the account with specifying the date of the last transaction performed, with information about the balance that the account contains, and also the expiration date Card and other information rather.

ii. Transaction Features. includes everything related to the transactions, including the number of transaction and account, as well as the transaction class, the currency kind, the time with the date, and even information related to the place where the transaction took place. And other information.

iii. Customer-specific features. includes everything related to the customer, such as the customer's number, detailed information about the customer's branch, the type of this customer, and other information.

2.2.4 The Following are Several Fraudulent Shapes Characterize Fraud[44]:

i. PIN changing before a high-profile transaction is initiated. Many incorrect logins before the transaction started.

ii. The occurrence of a sudden new transaction on an inactive or neglectful account.

iii. Transactions from blacklisted locations.

iv. A large transaction occurred from a region different from the area in which the person treated it.

v. The process of requesting change the card occurs during a transaction.

2.2.5 Fraudsters' Techniques

These are some commonly used techniques of online fraud [37]:

i. Site cloning: by doing fake sites. Then users enter these sites (without knowing what they are), in addition to entering their own card details.

ii. Generators of credit cards: Sometimes, fraudsters transact their operations using credit card generators. They are software applications that provide credit card info including numbers and expiration date.

iii. Account Takeover occurs if fraudsters use unlawfully obtained information stolen from cardholders (such as the number of cards and CVV, or another) to control their accounts. Fraudsters can sometimes send a message to the card's issuer to ask for changes, such as an address change or a card change. In other cases, scammers can log onto the customer account by using stolen data (e.g., user name and password) and modifying customer detail; this makes the account owner's recovery efforts are complicated.

iv. Fake Card: when the cards have been stolen, fraudsters clone and use them to transact fraudulently. Fraudsters employ various methods of cloning. Fraudsters use magnets to delete the magnetic card strip in many of these case scenarios. The card details are then changed to the card information of a valid user. In another case, a fraudster copies the magnetic strip data of a legitimate card in a fraudulent card using the skimmer tools (involving an electronic magnetic strip reader). Sometimes, the data is transferred without the cardholders' knowledge. This case occurs because the skimming tools can be carried in a fraudster's pockets and then moved closer to a genuine card. As a result, the fraudster will carry out the card-non-present transaction using the cloned card.

v. Mail Theft: Many fraudsters will distribute emails with newly supplied card numbers.

vi. Fraudsters can also collect information from cloned websites or even card issuer staff.

2.2.6 Fraud Relating to Credit Card can be Split into Three Categories[37]:

i. Fraud is done by using a physical card.

ii. Merchant fraud. This fraud occurs by the merchant or by the employees of the merchant. In the case of merchants, these merchants share with each other the process of stealing the data of the cardholders who are their customers and then giving it to the fraudsters.

iii. Internet fraudsters. In this case, fraudsters use various types of fraud for the purpose of gaining access to card owners' information illegally, and then they try to use those cards in online purchases or money transfers.

2.2.7 Types of Fraud[46]:

i. CCF: it is split into two categories. The first is the offline card, which occurs by physically stealing the card, then the fraudster uses it and withdraws money from it. The second type is online, where it occurs when the fraudster steals someone's card information and then makes purchases and withdraws money using this information.

ii. Telecommunication Fraud: The number of companies specializing in telecommunications has expanded in recent years, particularly as the impression of using phones has improved [47]. This high demand for the use of phones and telecommunication technologies has helped the growth and diversity of frauds in all parts of the world. It is carried out by relying on the services provided by telecommunications services, whereby users of these services are vulnerable to fraud, whether they are companies, customers, or even those working in providing telecommunications services. This problem is a large part of the cause of losses

for many companies and telecom providers. For fraudsters, this type of fraud is considered easy and quick to obtain illegal money, and it is less dangerous for them. Telecommunication fraud can be classified into two categories: The first is Subscription, and the second is superimposed. The subscription occurs in cases where fraudsters acquire an invoice, which shows at the stage of the phone number, without having any intention of paying the invoice. All transactions from the number obtained are fraudulent. The most likely use of these accounts is to sell calls or some other illegal use. On the other hand, when criminals rob a valid account, this is referred to as superimposed fraud. In that case, in addition to the regular use of genuine clients, the great use is superimposed. Cellular cloning is an example of overlapping fraud. Telecommunications fraud may also be committed by an insider if a staff provides confidential information to hustler individuals for unlawful objectives.

iii. Computer Intrusion**:** it is a cyber-crime involving hacking or the act of altering data into pcs, mobiles, and some other smart devices for illicit reasons [48]. In other words, the endeavor to get info is unauthorized. This form of fraud is mostly directed at people or even companies and maybe exceedingly destructive and have serious effects. A person who knows infrastructure and system layout often commits computer intrusion. The intruder may be an external person (hacker).

iv. BANKRUPTCY Fraud**:** The aim is to allow a person or enterprise to restructure its business and resolve its financial problems [49]. Fraud in the case of bankruptcy means that the customer uses a credit card while declaring bankruptcy. A bankruptcy scam is difficult to foresee since credit card issuers cannot know their clients' actual economic status. So in case of bankruptcy, the issuers will have to cover the loss themselves. One

way to prevent this type of fraud is to check the financial history of clients with the credit office.

v. THEFT/ COUNTERFEIT: scam through theft entails the use of a card that you do not possess. The fraudulent person obtains the victim's bank card then attempts to use it as frequently as possible before the legitimate customer reports suspicious activity and asks that the card issuer ban it. The sooner the actual customer reports, the quicker the bank will respond. A false scam happens if a scammer obtains and uses card details remotely over particular Internet sites, where no signature or actual card is required to complete the transaction.

vi. Application Fraud: that happens when the person asks for a bank card using the wrong info [50], or sometimes correct but stolen info [51]. In [51] Identification analytics identified over 300 million phony registration and found that 88 percent of these phony profiles were established using identity fraud strategies.

vii. Application fraud can be classified in two different situations: double application and fraudster identity. Duplicate applications are submitted by the same client using similar info. Cross-matching procedures are advised in this case since they can help in discovering the duplications. Cross-compliance technologies are designed to produce a suspicious numerical result using tacit connections across credit applications. Another situation (fraudster identity) occurs if different applications happen with similar details come from various individuals. However, the majority of banks need applicants to complete a form requesting certain details in order to assess their qualifying for a credit card. For search and identification purposes, most of the information requested is used.

viii. Behavioral Fraud occurs when "cardholder" sales are made, and legally binding card details have been gained fraudulently.

Figure 2.1 illustrates these types.

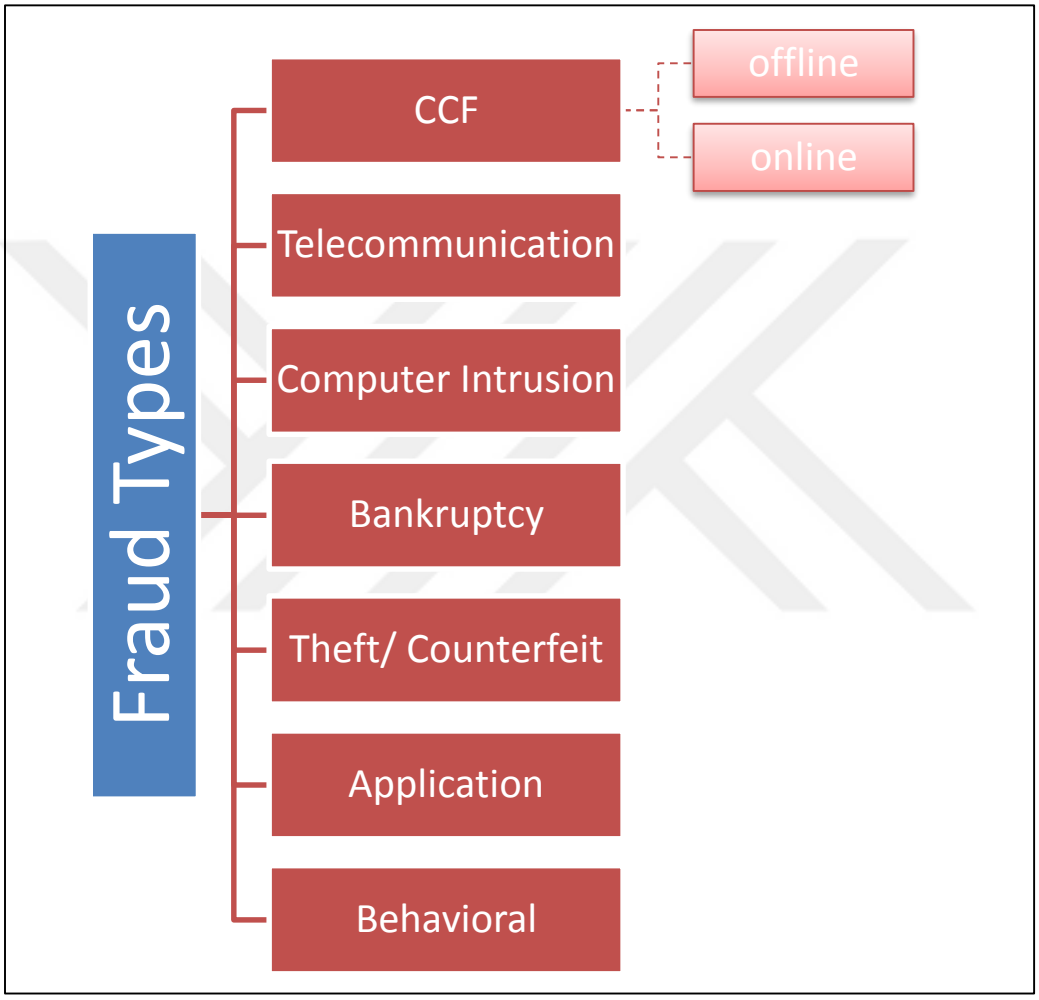

Figure 2.1: Fraud Types [46]

2.3 THE PROCESSING STEPS OF THE TRANSACTION BY CREDIT CARD[37]:

a. The user is required first to swipe a card (in the case of physical transactions) or enter card information (in the case of virtual transactions).

b. Moreover, if the card info is correct as well as the credit limit is acceptable, the activity is accepted.

c. The next step includes the process of sending the request for the transaction to a specific file, where this request remains for approximately 5 days for the merchant during this period to verify the validity of the transaction.

d. A key part of the transaction is transaction authorization. If the transaction is approved, the cardholder's information will be entered into a database, and the transaction will be recorded. The first level of safety is also involved. At this level, the cardholder's credit card limits are regulated.

e. After the important authorization step, the transaction cost deduction occurs from the total amount available to the card owner. Moreover, after this deduction, the process of sending this deducted amount to the merchant's account takes place.

f. All cardholder transactions and their balance are included in a monthly statement that the financial institution sends each month's end. After that, the user pays the total balance.

2.4 TECHNIQUES OF DETECTION

Different techniques of detection of fraud are used, such as ML and Nature Inspired(NI). We will try to explain some of them in this section. ML provides many techniques: HMM, NN, SVM, Random Forest(RF), etc… Sometimes more than one of these technologies is combined to generate a better classifier. The algorithm resulting from merging some ML algorithms together is called the ensemble. Hybridization techniques are generally more effective than independent techniques[37].

2.4.1 ML Techniques

ML is an area of Artificial Intelligence(AI) that can make a specific system learn by itself. Without being programmed to perform a given operation, the system analyzes, develops, and reacts dynamically to the situation. The field is the development of programs that can handle data by themselves, it is capable of accessing and modifying the data given in accordance with the user's specifications. The ML is divided into three main groups, the first is Supervised Learning, the second is Unsupervised Learning, and the third is Reinforcement Learning [29].

ML gives the system learning ability. The machine would use and interpret previously acquired data without being specifically instructed to do so. This feature is essentially useful in CCF detection. This allows for the successful implementation of ML algorithms to determine the transactions that have the potential to be harmful. Upwards of one million transactions are performed each day, all of which must be authenticated. The system can be trained to distinguish both the scams and non-scam transactions to achieve this goal. This is done with data from previous transactions, particularly from phony transactions, to label every newly coming transaction as either regular or hazardous. The hazardous ones would subsequently be separated for more inquiry [29]. There are several stages when ML are used in the fraud detection process. First, the training process starts from a core machine-learning algorithm that forms training data to dissect the relations with impartial consideration of various components. The objective is clearly indicated in the training stage for the machine learning algorithm[35]. Then, when the model is trained, it can then be used to predict unknown target values for other data instances. Machine learning may be categorized as Supervised or Unsupervised depending on whether the training data presented are labeled [35].

Supervised learning aims at discovering a link between an input value and an output value to predict extra input value if there is more input. A supervised learning problem may be further grouped into classification or regression. Classification concerns categorize output (for example, fraud and non-fraud), and regression concerns output as a particular value (for example, height). ML algorithms, which do not generate output but rather evaluate the input-output relationship, are called Unsupervised, as training data are not marked or classified[35].

Design dependent on features of fake and legal operations in supervised fraud detection methods to identify the transactions as scam or real, and transactions of outliers are detected in unsupervised fraud detection as possible instances of fraudulent transactions. There are many methods in ML used to detect and characterize CCF, such as the Artificial Neural Network model, Bayesian model, LR, SVM, DT, K-NN, etc.[35].

2.4.2 Selection of features:

The key basis for the detection of CCF is to analyze the spending behaviour of the cardholder. This spending behaviour is evaluated by using appropriate features to determine the behaviour of a single credit card, which generally changes an actual or fake transaction profile. These features come from a historical mixture of past transactions. The basic inputs categorize all features, including operational information, geographical data, merchant type data, quantity time-based data, and transaction time [52].

The variables of all transactions under the statistical type show an overall use pattern for the card. The card's habit of spending takes geographic areas into account all the variables under the type of geographical data, and variables show card use under a type of trader data in separate trading classes; Data variables based on time describe the usage history of cards as compared to time periods for the quantity of use. In general, some papers focused on the card owner's data rather than on card details. It is also obvious that a person may have at least two payment cards for different reasons. Therefore, the spending habits and behaviour of the person will appear on those cards. During this examination, the core is used towards the card. In contrast to the cardholder, one of the cards can show an interesting cost habit while the card owner will demonstrate several practices with different credit cards [35].

2.4.3 Some Types of ML Algorithms that are Used in CCF Detection:

i. Supervised Algorithms, such as:

   a. Random Forest(RF) [19][9][53][22].

b. Neural networks and Artificial Neural Networks(ANN)[54][21].

c. Deep Learning(DL)[36][8][55].

d. SVM [17][54][56].

e. NB [57][58].

f. LR [58][59].

g. Linear Discriminant Analysis (LDA)[60][61].

h. Extended Gradient Boosted Tree (XGBoost)[62][63].

i. K-Nearest Neighbours (KNN)[57][21][53].

ii. Unsupervised algorithms, such as:

j. Self-Organizing Maps (SOM) [54].

k. K-means clustering algorithm[64].

l. Isolation Forest[62].

m. The local outlier depends on a local density concept where the location of which the distance is used by comparing the outliers to estimate the density is provided by closest neighbors [62].

In our work, we used the following supervised classification techniques LR, NB, LDA, and XGBoost, which are described in section 2.4.5 in this chapter.

2.4.4 Classification Algorithms

Algorithms of classification are used if there is needing to categorize data into specific categories. Many types of classification are available depending on the count of divisions into which the data will be sorted. The classification kinds include binary, multiclass, and multi-labeled [65]. The classification algorithms are used in different applications. One of these applications is CCF detection. The following stages can indeed be utilized to create a classification model [66]:

i. Dataset gathering and then preprocessing.

ii. Initialize the model of the classifier.

iii. Divided the dataset into train and test, then provided training data to the classifier model.

iv. To get new observation data, predict the label.

v. Assess the error rate on the test dataset for the classifier model.

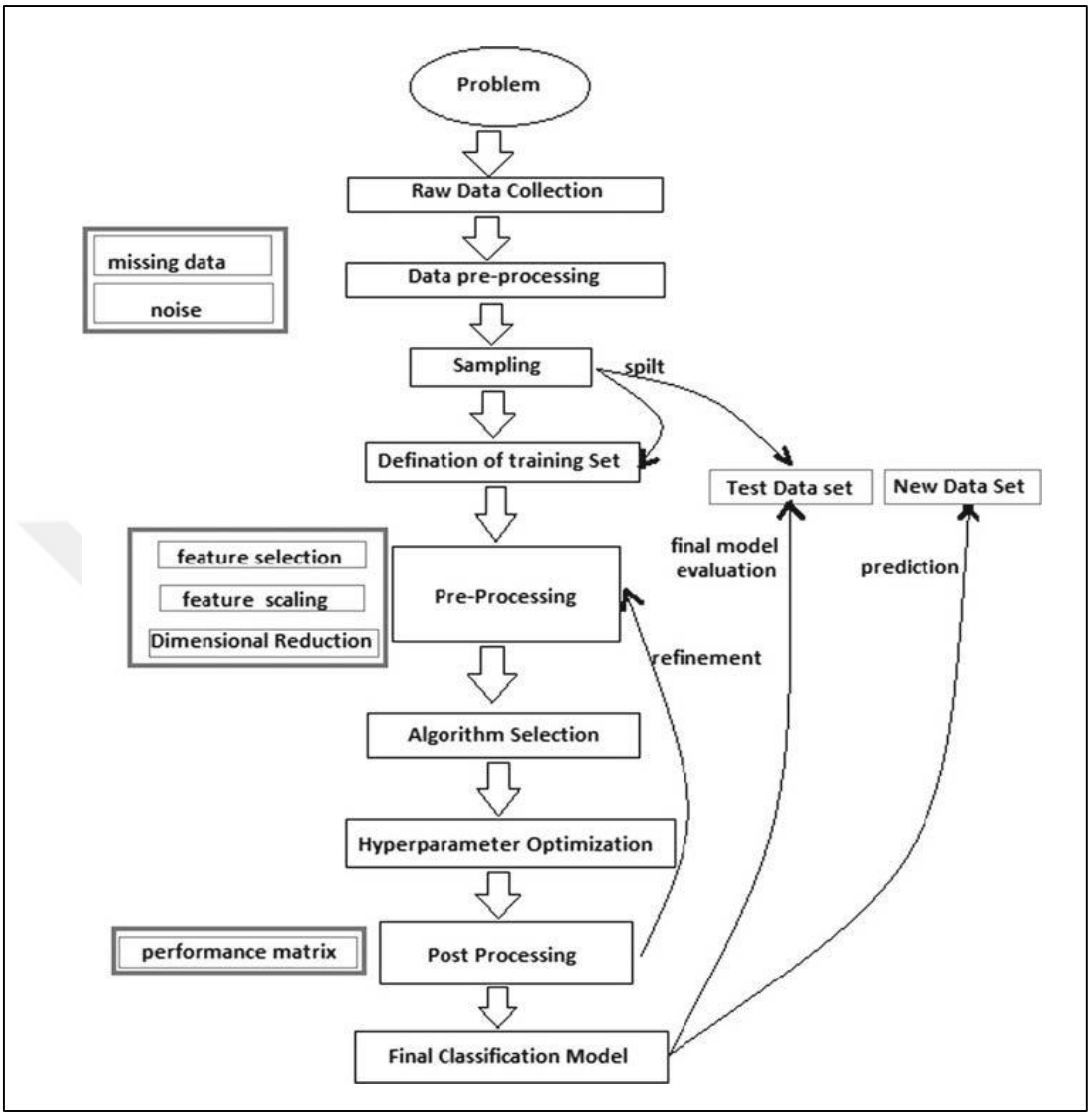

Figure 2.2: Diagram of classification model[66]

Thus, the initial step is the dataset acquisition. Next, you have to find out the most pertinent attributes/fields/features, it is referred to as feature extraction. One of the easiest methods that extracts the most relevant/informational properties by measuring all accessible elements is the Brute-force method. A subset feature selection procedure also detects and reduces the largest possible number of redundant, irrelevant, and superfluous attributes [66]. Secondly, there is a need for doing data pre-processing[67]. The dataset frequently includes noise (outliers) and sometimes missing feature values, and also there is a need to turn some classes into dummy variables. There are several ways of handling missing data that must be taken into account. Therefore, a good pre-processing of data is required[68][69].

2.4.5 ML Techniques Used in Our Work

i. LR is an extensively used algorithm for classifying. It is easy to use LR when the dependent variable is binary or categorical, and the predictors are both continuous. The dependent variable could have a binary nature. Some predictors can tell whether or not something will happen. The likelihood of attending each class is predicted [70].

ii. NB is used for classifying data. It works using the Bayesian theorem posterior probability approach to classifications. The NB model performs admirably, particularly when the predictors are divided into distinct classes. However, even when predictors lack a distinct, independent class, they performed well. In the NB method, data is classified twice. Predictors are assumed to be conditionally independent in this phase, which means that data from a training dataset is used to compute probability distribution parameters. Next, the unfamiliar data (test dataset) is fed into the NB classifier to predict and evaluate individual classes' posterior probabilities. After that, the dataset is classified according to the posterior probability with the highest score[30]. NB is a good method for large datasets, since there's no necessity for sophisticated iterative parameter approximation [71]. When it comes to training and classification, it does a great job[72]. Feature independence and essentiality are assumed to be equal for all features in the feature vector[30].

iii. LDA: Ronald A. Fisher first proposed the technique in 1936 for a two-class problem, and C.R. Rao later proposed the multiclass LDA in 1948[73]. The primary aim of this model is to represent the dependent variables as a linear combination of the independent variables. To get the classification result, these linear equations must be processed further [74].

iv. XGBoost: an effective machine learning technique is required to address a range of machine learning issues including classification and others [53]. XGBoost is a Gradient Boosting framework optimized for efficiency, flexibility, and portability developed by Chen and Guestrin. The basic idea behind boosting is really to conjoin a collection of low-accuracy poor classifiers in order to construct a strong classifier gives the improving classification performance. XGBoost is a fast and scalable Gradient Boosting Machine (GBM) implementation with easy parallelism and strong predictive accuracy. It can also deal with transient stability prediction because of the following advantages. First, It is possible to automatically use multithreading parallel processing in the XGBoost model to predict transient grid stability with huge datasets more quickly than the traditional ensemble learning method. Second, adding a regularization term to XGBoost improves the generalization ability, making up for the decision tree's tendency to become overfitted. Third, data acquired by the Phasor Measurement Unit (PMU) in the power system need not be normalized for XGBoost's tree structure model. It is capable of handling values that are missing [75].

2.5 SAMPLING TECHNIQUES

Numerous strategies exist to cope with an unbalanced dataset[76]. These strategies are described and classified into four broad categories: algorithmic, data-level, cost-sensitive, and ensemble-based.

Oversampling, under-sampling, and hybrid methods are all types of data level sampling. Using the under-sampling techniques, balancing dataset by eliminating items of the large proportion class until to make it equal to the other class items. As a result, a significant amount of the original dataset's value is lost. On the other hand, the oversampling duplicates items of the small proportion

class at random, increasing the volume of the dataset and, as a result, the ML model training time. When using a hybrid method, we combine the advantages of both sampling methods[19].

In our work, we apply the SMOTE technique to deal with the unbalanced dataset to tackle these hurdles. The SMOTE approach is well-known and may be put to various uses [77][78].

Synthetic Minority Oversampling Technique (SMOTE) is a popular and successful method of oversampling. The SMOTE is used to duplicate existing instances of the minority class and to create new artificial instances of the minority class using knowledge about the neighbors of every sample in the minority class[79].

2.6 EVALUATION PERFORMANCE METRICS

The metric of evaluation has an important role to play in classification training to ensure optimal classification. Therefore, selecting appropriate assessment metrics is a vital component to discrimination and optimum classification [80].

To assess the performance of the model, metrics are utilized. Many metrics are available. A classifier is assessed on the basis of performance metrics calculated after model training. Different classifiers for machine learning are available[81][82][83]. There is no general agreement among researchers on specific metrics to be used to evaluate the success of a classifier[84][85]. Sometimes the classifier to be very good when evaluated by a certain performance metric, but the same classifier appears not good when evaluated by another performance metric [86]. Usually, the choice of the analyzer falls on the use of the most common and widely used, such as classification accuracy, the area under the ROC (receiver operating characteristic), and F-measure [87][84][85]. Several research may create models then provide conclusions by relying on their chosen metrics. This might result in redundancies in performance space assessment, neglecting some aspects of performance or making obstructing the determination of a clear winner in the event of conflicting performance measurements[86]. A classifier should be assessed by a number of performance

measurements that independently capture a distinct part of the classifier range. This is especially essential because most performance metrics are generally dependent on the four values of the confusion matrix, i.e., number of true positives(TP), number of true negatives(TN), number of false positives(FP), and number of false negatives(FN). As a result, it is reasonable to believe some of the metrics produced are likely connected [86].

3. DESIGN AND IMPLEMENTATION

In our project, we tried to take advantage of ML techniques in order to design a CCF detection system. Supervised algorithms were used. The classical algorithms were used in addition to one of the Boost algorithms. The task required in this problem is the classification problem, so these algorithms were chosen as classification algorithms. Relying on these algorithms, system models are designed, trained, and tested. In machine learning, datasets play a crucial role. Moreover, because our study needs data in credit card fraud, the datasets from Kaggle [88] and IEE Data Port [89] were used.

The next sections of this chapter explain the details of the used datasets and the methodology.

3.1 DATASET DESCRIPTION

The model's efficacy is highly dependent on the sort and type of data that it is working with. We 0used two credit card data sets in our work. One was a dataset of European cardholders that we got from Kaggle. The second dataset was the Turkish dataset obtained from IEEE Data Port.

3.1.1 The First Dataset

It was created by using the data of European card users' transactions in Sept 2013 for a 2 days. and was derived from data collected throughout that month. Because the problem of confidentiality is a critical one that must not be disregarded, it is difficult to show and exhibit all of the data's fundamental qualities and information without compromising its confidentiality. As a result, the numerical values in the datasets presented are the outcome of the PCA transformation and are not included in the datasets themselves. This dataset has 31 columns and is organized as follows: Each column indicates a different characteristic. A new label was assigned to each column due to the confidentiality principle being observed. First and foremost, there is a column named Time, and

from the second column on up to column number 29, the labeling began with V1, V2..., V28, and was followed by the last two columns, which were labeled Amount and Class, respectively. Besides, excluding the 2 columns values (Amount and Time), every one of the values in the dataset, were given as numerical values after they were converted to PCA values. As for the final column, No. 31 (Class), its values were restricted to the numbers 1 and 0, as these numbers indicate whether or not the transaction is genuine. If the value is zero, the transaction is regarded as valid. On the other hand, If the value is 1, it is regarded as illegal. The total number of transactions is 284,807, with 492 of those being fraudulent.

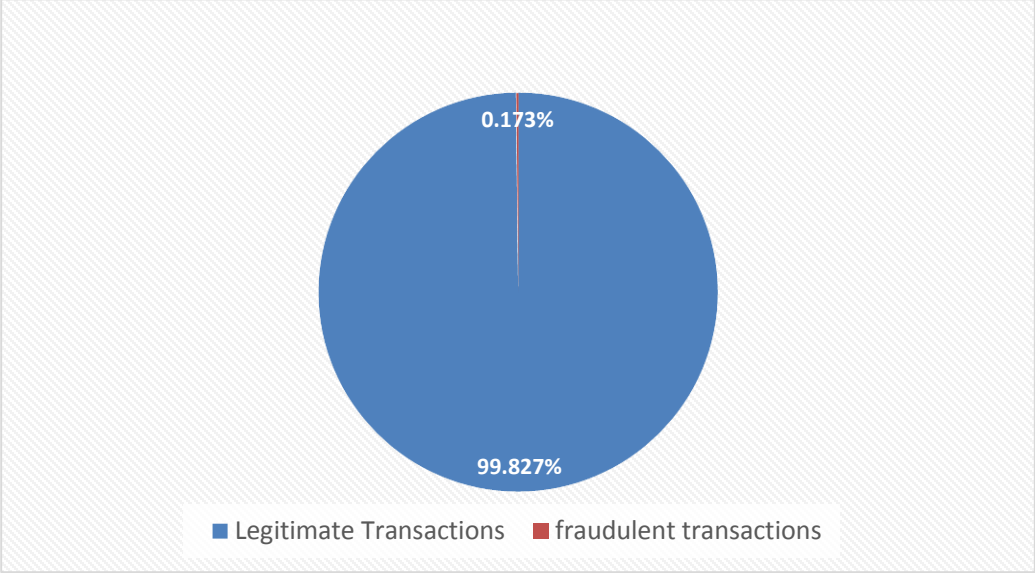

Figure 3.1: Dataset 1 Distribution

As illustrated in Figure 3.1, the data is so skewed, with the percentage of fraudulent transactions being only 0.173 percent of the total, while the percentage of legitimate transactions is 99.827 percent. As a result, we employed SMOTE to balance the dataset.

Table 3.1: Description of Dataset 1 Features

feature	Description
Time	Shows the range of seconds among the occurrence of each activity and the first activity.
V1, V2, …, V28	Various information about the transaction has been converted into PCA format for confidentiality purposes.
Amount	The amount of Transaction
Class	Transaction status, value (1) if fraudulent, and (0) if legal.

3.1.2 The Second Dataset

This dataset is available on the IEEE Data Port [89]. The dataset comes from Yapi Kredi Teknoloji and includes 62,380 transactions. 61,572 of which are legitimate (98.70%) and 808 (1.30%) are fraudulent. Each transaction has 28 features a TimeDate, a class label (1 for fraud, 0 otherwise), Amount, CardNo, and 24 extra explaining features titled from feature0, feature1…, to feature 23. Even though the transactions in this dataset are real, the variables are numeric due to Principal Component Analysis (PCA) preprocessing and lack an explanation because of confidentiality issues. This dataset was also split into two portions; one is for training/testing and the second for validation. In our work, we used the part of training/testing. This part consists of 31,190 transactions. The number of genuine transactions is 30,400, and the number of fraudulent is 790.

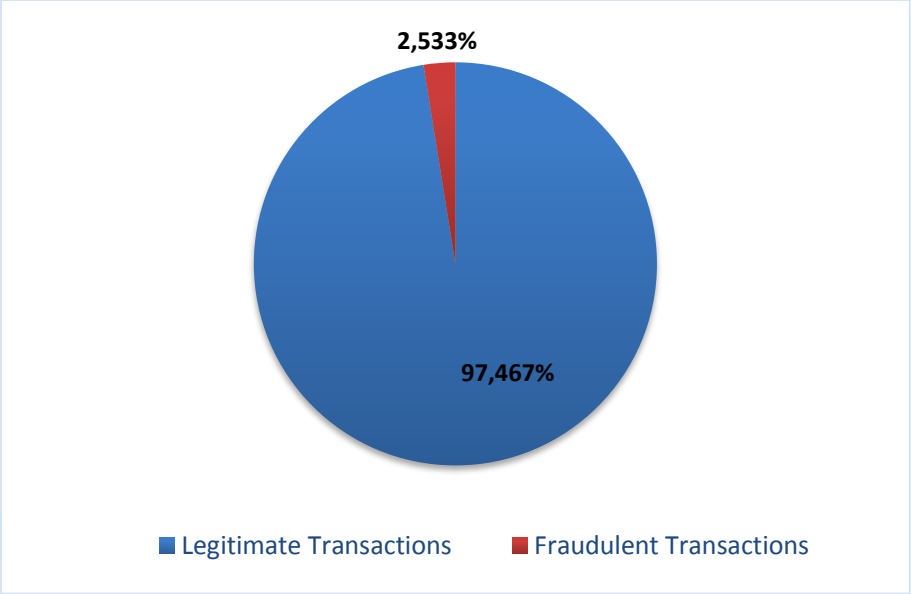

Figure 3.2: Dataset 2 Distribution

As illustrated in Figure 3.2, the data is extremely skewed; the scam transactions are 2.53 percent of the total. Therefore, SMOTE was used.

3.2 METHODOLOGY

This work proposes a CCF detection system using ML algorithms. In our work, Logistic regression(LR), Linear Discriminant Analysis (LDA), and Naïve Bayes(NB), additionally the boosting algorithm XGB have been implemented in Python language three times for each dataset. The first time was on an original dataset (before applying SMOTE). The second time on the dataset after applying SMOTE. The third time, after applying the fuzzy membership function with SMOTE. Then, we presented their results. Finally, we compared these results with other papers' results. The general steps of our methodology are illustrated in Figure 3.2.

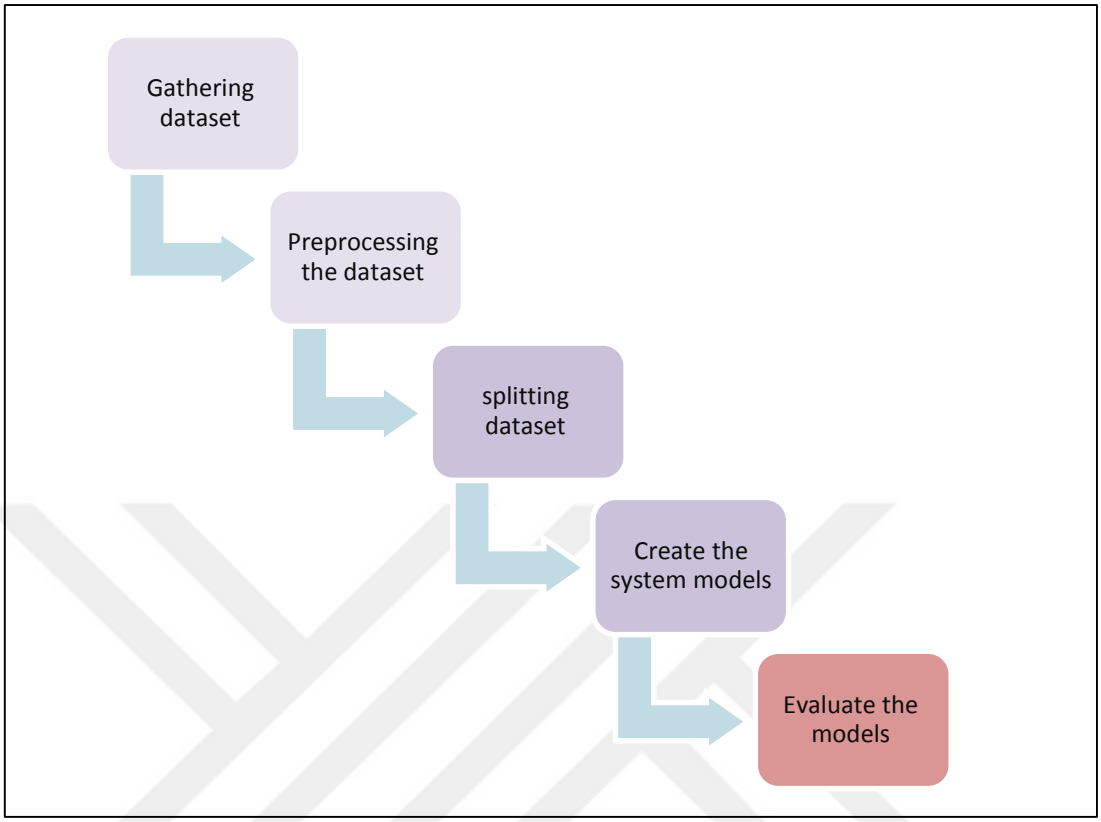

Figure 3.3: General implementation steps

The use of machine learning techniques in the field of CCF detection requires going through several steps:

i. The first step is to bring and collect the data used. A dearth of credit card dataset is one of the most significant challenges researchers confront, and to implement our system, we used datasets from the Kaggle website and the IEEE Data Port.

ii. The second step is dedicated to preprocessing the collected dataset, where the dataset is cleaned to give good results when used in the next steps. In this step, the nulls are eliminated, the appropriate feature selection is performed, and a scale for feature values is done when needed. Also, at this stage, the problem of data imbalance is solved in case the data is unbalanced.

iii. The third step is Splitting the dataset into test and train datasets.

iv. The fourth step is Choosing a model according to our dataset and the type of the task. In the case of credit card fraud detection, the task is binary classification.

v. The fifth step is evaluating the performance of the proposed system models. The proposed system is shown in Figure 3.3.

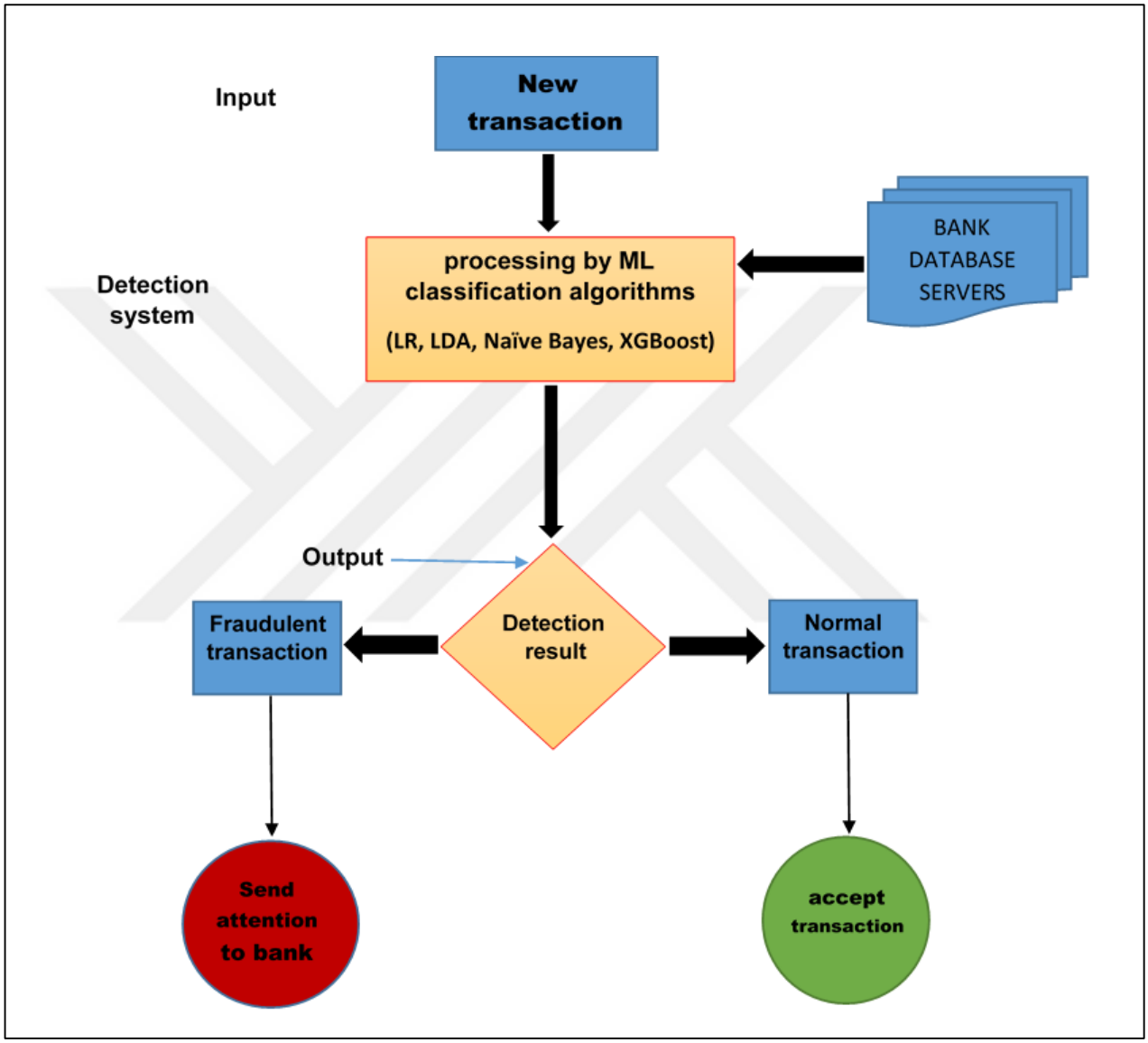

Figure 3.4: Proposed CCF Detection System Architecture

3.2.1 Dataset Gathering

The datasets we used were obtained from Kaggle's website and IEEE Data Port. The first dataset is a European Cardholder, and it is a reliable and well-known dataset that many researchers have used. The second dataset is Turkish, provided by Yapi Kredi company. These datasets contain a

label with two values, either 0 or 1, 1 for the fraudulent transaction, and 0 for the legal transaction. The two datasets suffer from a high imbalance between normal and fraudulent classes.

3.2.2 Dataset Preprocessing

The datasets were ready for use and were not needed to clean. So several dataset preprocessing has been done: scaling, SMOTE, and fuzzy membership function. The values of DateTime feature in the Turkish dataset were converted to numeric because it was a string.

3.2.2.1 Scaling

To scale the dataset, we used Standard Scaler processing (Standardization). It is the most famous process. It is done using a standardization procedure to make the feature columns have the same properties as an ordinary normal distribution, with mean 0 and standard deviation 1. For the European cardholders dataset, a scale was created for the features Amount and Time, and then the resulting values were placed in two new columns for each feature, where the values are ranged between 1 and -1. A normalizedtime replaced the Time column, and a normalizedamount replaced the Amount. After that, we dropped the two main columns, Amount and Time. Also, we dropped the Class column. A scale was created for all the features in the Turkish dataset, except the label feature. As for the scaled features, they were replaced by other features for each one. After that, we dropped all the main columns. Also, we dropped the Class column.

3.2.2.2 Balancing

There is a significant imbalance between fraudulent transactions (class 1) and legal transactions (class 0) in the two datasets. In the case of the first dataset, only 0.173 percent of total transactions are fraudulent, whereas 99.827 percent of all transactions are legitimate. In the case of the second dataset, only 2.533 percent of total transactions are fraudulent, whereas 97.467 percent of all transactions are legitimate. To balance the datasets, we employ the SMOTE approach. This technique increases the data of the small class and makes it equivalent to the data of the large class. In the first dataset, the number of fraud transactions before applying SMOTE was 492, and after applying SMOTE increased to 284315. Also, in the other dataset, the fraud transactions before applying SMOTE were 790, and after applying SMOTE increased to 30400.

3.2.2.3 Applying the Fuzzy

The fuzzy membership function was performed to expand the number of dataset columns (features).

i. Fuzzy membership function: fuzzy sets are an excellent recommendation for datasets that have classification issues[90]. When fuzzy logic is used, sets represent linguistic concepts, and membership functions represent the true values of individual linguistic claims. The degree of membership of an object in a fuzzy set specifies a function with a domain equal to the set of potential values for the object and a range equal to [0,1]. This is a membership function that is defined[91]. The triangular membership function defined in equation 3.1 is depicted in Figure 3.4. A fuzzy space is a collection of fuzzy sets that determine the fuzzy linguistic values or fuzzy classes that can be associated with an attribute.

$$f(x,a,b,c) = \begin{cases} 0 & if\ x \le a \\ \frac{x-a}{b-a} & if\ a \le x \le b \\ \frac{c-x}{c-b} & if\ b \le x \le c \\ 0 & if\ x \ge c \end{cases} \quad (3.1)$$

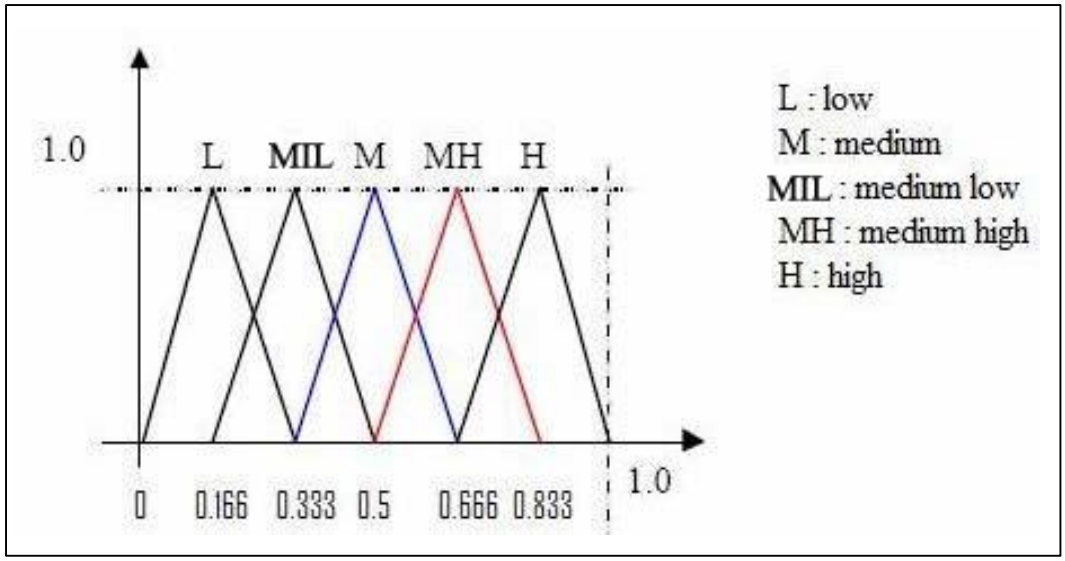

Figure 3.5: The Function of Membership [90]

Fuzzy spaces and fuzzy logic allow for the concurrent membership of an object in many classes. By using the fuzzy, we were able to create five new columns from each column.

In the first dataset, after processing the features V1 toV28 via the fuzzy function, the dataset has 171 columns. While in the other dataset, after processing the features (feature0 to feature23 addition to Amount) via the fuzzy function, the dataset has 153 columns. The goal of utilizing the fuzzy function is to be an aid to improve the performance of the results as much as possible.

3.2.3 Splitting Dataset

At this stage, we divided the data into two parts. One is to train the models of the proposed system, and the other is to test those models. The percentage of the test dataset was 30%, and the percentage of the training dataset was 70%.

3.2.4 Create the System Models

We selected a model based on our dataset as well as the task characteristics. The task in CCF detection is classification. So the LR, LDA, NB, and XGBoost are all used.

3.2.5 Performance Evaluation

After the training and testing of the systems' models, we must evaluate the performance of these models. There are many performance metrics used for that. We used confusion matrix, accuracy, recall, precision, f1, and AUC.

4. RESULTS AND DISCUSSIONS

Here we presented all the obtained results in detail and discussed. After that, we compared the results of our work and then also compared it with other works.

In our work, we used several performance measures to determine the efficiency of the proposed system models.

4.1 THE PERFORMANCE METRICS USED IN OUR WORK

Performance metrics are often used to evaluate a model's performance. There are many types of metrics. Some of them were used in this study; what interests us is the classification metrics because the fraud detection problem is considered a classification problem. As datasets classes are highly unbalanced, common performance metrics such as accuracy cannot determine the actual fraud rate. Classification metrics must therefore be appropriately selected [62]. The following lines describe the details of all metrics used.

4.1.1 Confusion Matrix

it is the simplest and effective way to apply in terms of performance evaluation. The amount of correctly identified data instances can be determined by using a confusion matrix [92]. In the confusion matrix, the model becomes increasingly accurate as of the values of TP and TN increase. The details of the confusion matrix are shown in Table 4.1.

Table 4.1: Confusion Matrix

Predicted class	Actual class	
	positive	Negative
Positive	True positive (TP)	False negative (FN)
Negative	False positive (FP)	True negative (TN)

a. TP values: an actual class of the data that matches the predicted class of the data.

b. FP: the actual class of the data was 1, but the model predicted it to be 0.

c. FN: the actual value of the class was 0, but the model predicted it to be 1.

d. TN: the actual value was 1, and the model predicted it to be 1.

4.1.2 Accuracy

It is employed for evaluating the number of TP. That means The larger the number, the higher the accuracy. Sometimes accuracy measures might be deceiving when the data set is excessively imbalanced. For the upper instance class, the true positive is high. Classification methods should therefore be appropriately selected. Accuracy can be computed as follows:

$$\text{Accuracy} = (TP + TN)/ (TP+ FP+ TN+ FN) \quad (4.1)$$

4.1.3 Precision

It is used for measuring the positive patterns successfully forecasted in a positive class from the total forecast patterns. It can be computed as follows:

$$\text{Precision} = TP/(TP+FP) \quad (4.2)$$

4.1.4 Recall (Sensitivity)

It is used for calculating the fraction of positively categorized patterns. It can be computed as follows:

$$\text{Recall} = TP/(TP+FN) \quad (4.3)$$

4.1.5 F1

It can be computed if the values of recall and precision are available. It shows the harmonic mean between the values of recall and precision. Its calculation formula is:

$$\text{F-1 score} = 2*\text{precision} * \text{recall} / (\text{Precision} + \text{recall}) \quad (4.4)$$

4.1.6 AUC

It is a statistical metric for assessing how widely a sample is dispersed. It evaluates whether or not the categorization model is faultless. With a greater AUC, a model's ability to distinguish between different classes is better understood. As the AUC score improves, the model's ability to distinguish between fakes and genuine.

4.2 RESULTS AND DISCUSSIONS OF THE FIRST DATASET

The results were presented and discussed in three sections: the results of implementation on the original dataset, the implementation results after applying SMOTE to balance the dataset, and the implementation results after applying SMOTE and fuzzy membership function.

4.2.1 Results and Discussions of Implementation on the Original Dataset

the number of TP, TN, FP, and FN for all classifiers are illustrated in Figure 4.1**.** Table 4.2 shows the results of the classifiers before SMOTE.

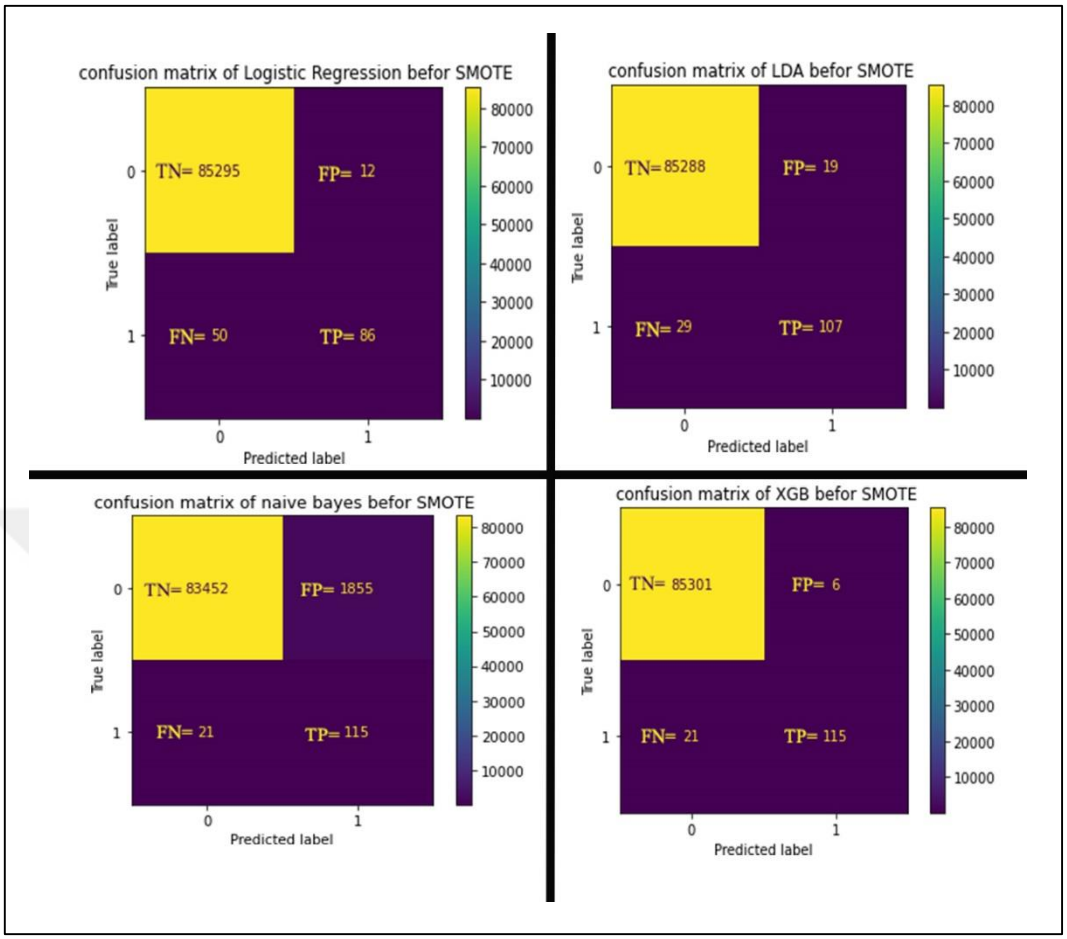

Figure 4.1: The confusion matrix before applying SMOTE

Table 4.2: The results before applying SMOTE

Classifier	accuracy %	precision %	Recall%	F1%	AUC%
LR	99.926	87.755	78.676	81.679	81.611
LDA	99.944	84.921	84.559	10.921	89.327
NB	97.804	5.838	85	10	91.192
XGBoost	99.968	95.041	84.559	89.494	92.276

The results produced from all classifiers prior to using the SMOTE are unconvincing since the dataset has a significant imbalance between its two classes, with fraud data accounting for 0.173 percent of the whole data set. That is, the models will be trained and tested exclusively on the largest dataset class (non-fraud). The confusion matrix for all models is shown in Figure 4.1. Each

model contains a much big number of TN. We saw that the lowest accuracy value was 98 percent while looking at the findings in Table 4.2, which shows the outcomes of the classifiers prior to applying the SMOTE. Other metrics such as precision, f1, and recall are insignificant compared to accuracy. This demonstrates that the model's training was insufficient. Thus, in order to achieve true accuracy, the dataset must be balanced. There are numerous approaches for resolving the issue of imbalance. In our work, we used SMOTE to balance the dataset.

4.2.2 Results and Discussions of Implementation after Applying SMOTE

The number of TP, TN, FP, and FN for all classifiers are illustrated in Figure 4.2. Table 4.3 shows the results of the classifiers after applying SMOTE.

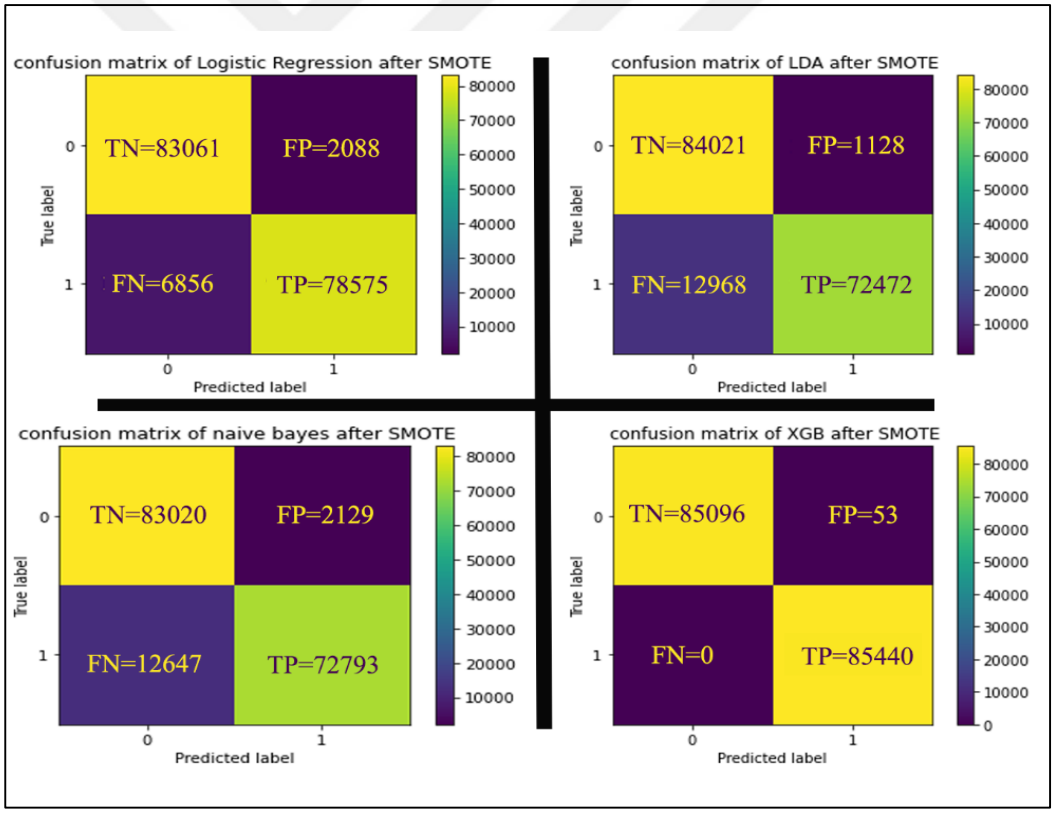

Figure 4.2: The confusion matrix after utilizing SMOTE

Table 4.3: The results after utilizing SMOTE

Classifier	accuracy %	precision %	Recall%	F1%	AUC%
LR	94.752	97.411	91.965	94.61	94.756
LDA	91.737	98.467	84.822	91.137	91.749
NB	91.338	97.158	85.198	90.786	91.349
XGBoost	99.969	99.938	100	99.969	99.969

Following the application of SMOTE, which equalizes the two divisions of the dataset via increasing the data from the small class, we can see from Table 4.3, which displays the outcomes of the classifiers following the application of SMOTE, that the results for all classifiers were good. There were no statistically significant differences in the LR, LDA, and NB outcomes, except for the recall values of LDA and NB, which were lesser than the other metrics values. Overall, XGBoost produced the best results of any of the models tested; the accuracy, f1, and AUC values were all 99.969 percent, precision was 99.938 percent, and recall was 100 percent for this model. After applying the SMOTE to the dataset, we discovered that the greatest classifier performance was achieved in the following order: XGBoost, LR, LDA, then NB. From best to worst, the accuracy rates were 99.969 percent, 94.752 percent, 91.737 percent, and 91.338 percent, respectively.

4.2.3 Results and Discussions of Implementation after Employing the Fuzzy

The number of TP, TN, FP, and FN for all classifiers are illustrated in Figure 4.3. Table 4.4 shows the results of the classifiers after applying fuzzy.

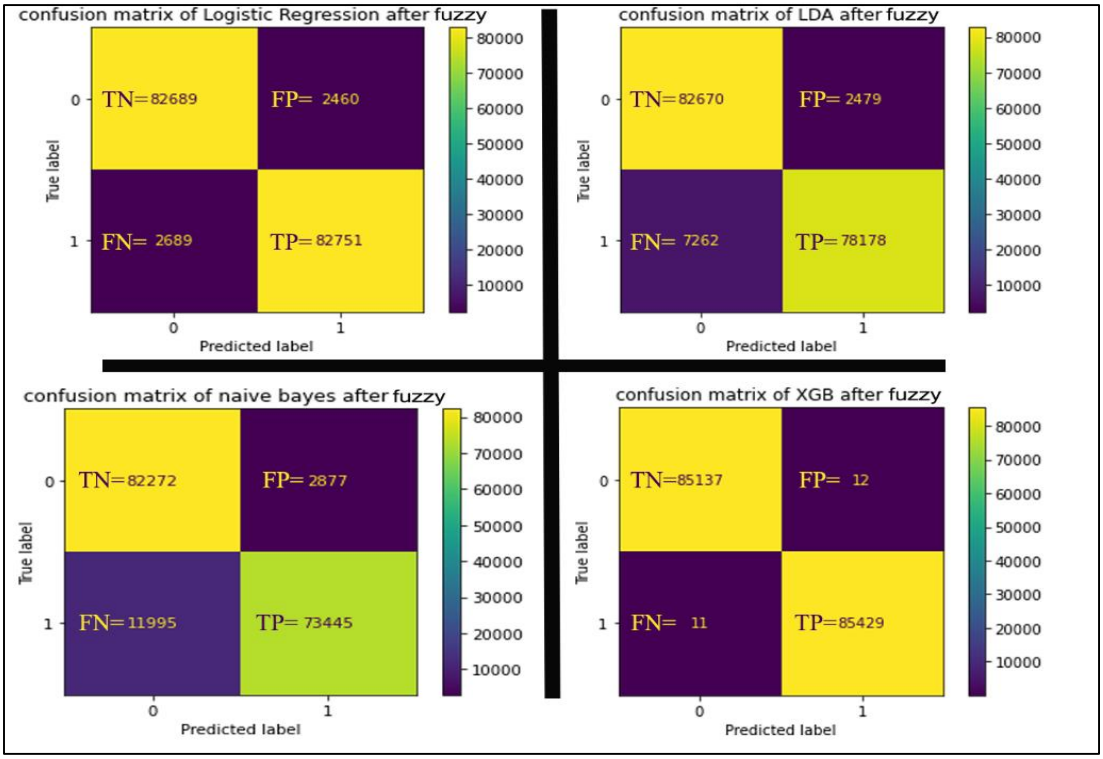

Figure 4.3: The confusion matrix after applying fuzzy membership

Table 4.4: Results of employing the fuzzy

model	accuracy%	precision%	recall%	F1%	AUC%
LR	96.982	97.113	96.853	96.983	96.982
LDA	94.290	96.926	91.500	94.135	94.295
NB	91.282	96.230	85.961	90.806	91.291
XGBoost	99.987	99.986	99.987	99.987	99.987

Because the suggested system's purpose is to achieve the best feasible results, we attempted to determine how much the fuzzy membership function influences the system's efficiency. The outcomes showed a significant improvement. As shown in Table 4.4, we obtained the following results. As an LR's findings, the accuracy is improved to 96.982 percent, whereas it was 94.752 percent before applying the fuzzy; recall improved from 91.965 percent to 96.853 percent; F1

improved from 94.61 percent to 96.983 percent, and AUC improved from 94.756 percent to 96.982 percent after applying the fuzzy. Concerning the results of the LDA, the accuracy is improved to 94.290 percent, whereas it was 91.737 percent before applying the fuzzy membership; recall improved from 84.822 percent to 91.5 percent; F1 improved from 91.137 percent to 94.135 percent, and AUC improved from 91.749 percent to 94.295 percent. For the results of the NB, the changes are almost non-existent. Where all the results remained almost constant. As for the results of XGBoost, they were very high, close to 100%, so the improvements that occurred were very, very slight, with values close to parts of a thousand.

The final results extracted from all of the above prove that the best capabilities of the proposed system models were for XGBoost, then LR, LDA, and lastly NB.

4.3 RESULTS AND DISCUSSIONS OF THE SECOND DATASET

The findings are provided and discussed in three sections: the implementation results on the original dataset, the implementation results after using SMOTE, and the implementation results after applying the fuzzy membership function.

4.3.1 Results and Discussions of Implementation on the Original Dataset

Figure 4.4 illustrates the total TP, TN, FP, and FN for all classifiers. The results of the classifiers before employing SMOTE are shown in Table 4.5.

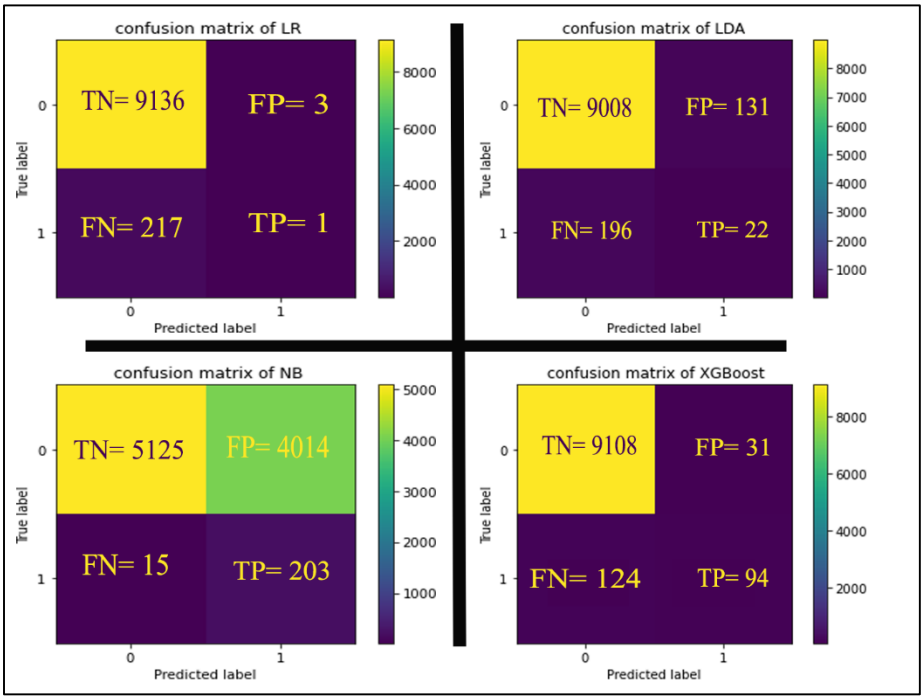

Figure 4.4: The confusion matrix before applying SMOTE

Table 4.5: The results before applying SMOTE

Classifier	accuracy %	precision %	Recall%	F1%	AUC%
LR	97.649	25	0.459	0.901	50.213
LDA	96.505	14.379	10.092	11.86	54.329
NB	56.941	4.814	93.119	9.154	74.599
XGBoost	98.343	75.2	43.119	54.81	71.39

Prior to implementing the SMOTE, all classifiers provided unconvincing results due to the dataset's large imbalance between its two classes, with fraud data accounting for 2.533 percent of the whole data set. Models will be trained and evaluated exclusively on the class with the largest dataset (non-fraud). Figure 4.4 depicts the confusion matrix for all models. This demonstrates that the model was not adequately trained. When a dataset is skewed, the accuracy is bogus. Thus, the dataset must be balanced in order to obtain actual accuracy.

4.3.2 Results and Discussions of Implementation after Applying SMOTE

Figure 4.5 illustrates the total TP, TN, FP, and FN for all classifiers. The findings of the classifiers after using SMOTE are displayed in Table 4.6.

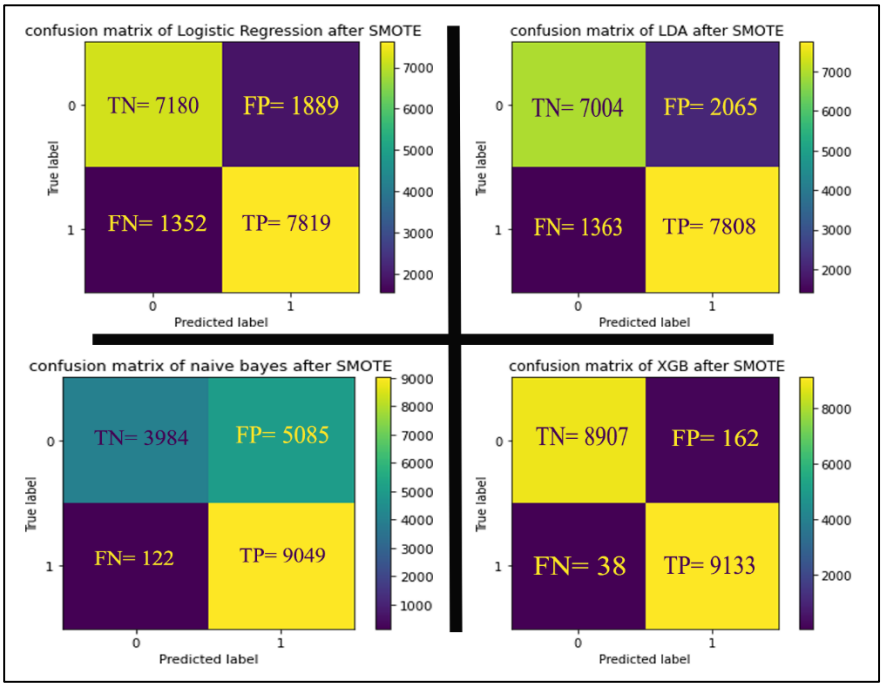

Figure 4.5: The confusion matrix after applying SMOTE

Table 4.6: The results after applying SMOTE

Classifier	accuracy %	precision %	Recall%	F1%	AUC%
LR	82.231	80.542	85.258	82.833	82.214
LDA	81.206	79.084	85.138	82	81.184
NB	71.453	64.023	98.67	77.657	71.3
XGBoost	98.904	98.257	99.586	98.917	98.9

Following SMOTE, which equalizes the two classes of the dataset, we can see from Table 4.6, which displays the results of the classifiers following the applying of SMOTE, that the results for

all classifiers were good excepted the NB. There were no statistically significant differences in the LR and LDA outcomes, for the nb results were the lowest among all the other classifiers. While XGBoost produced the best results of any of the models tested. The accuracy of XGBoost was 98.904%, precision was 98.257%, recall was 99.586%, f1 was 98.917%, and AUC was 98.9%. After applying the SMOTE, the classifier performance was achieved in the following order: XGBoost, LR, LDA, then NB. From best to worst, the accuracy rates were 98.904 percent, 82.231 percent, 81.206 percent, and 71.453 percent, respectively.

4.3.3 Results and Discussions of Implementation after Applying the Fuzzy Membership Function

Figure 4.6 illustrates the total TP, TN, FP, and FN for all classifiers. The findings of the classifiers after employing fuzzy are displayed in Table 4.7.

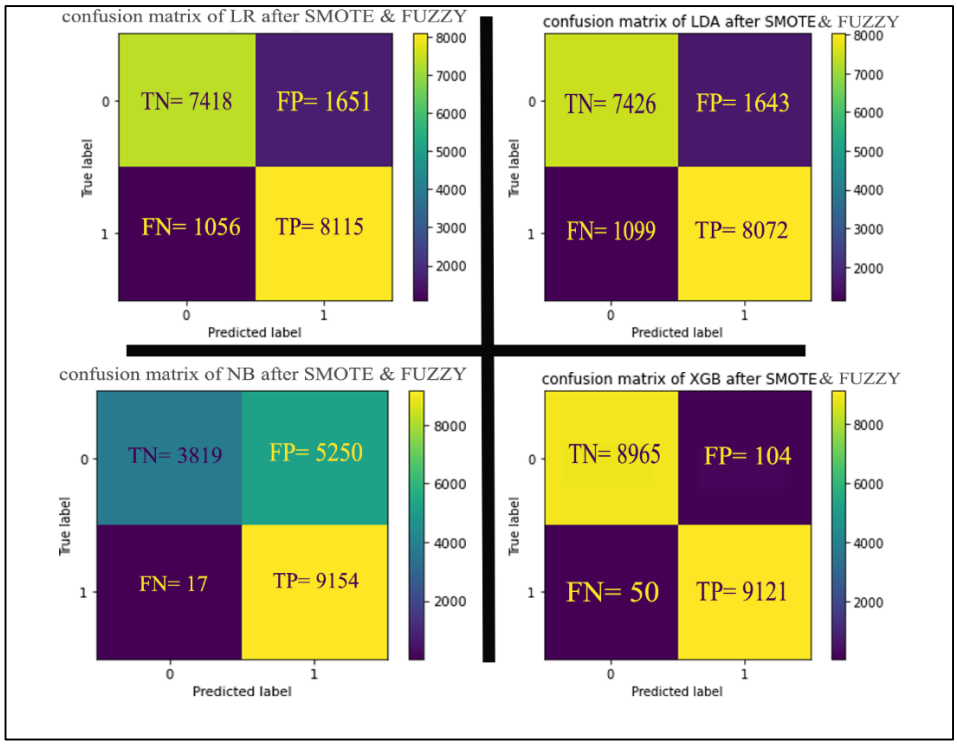

Figure 4.6: The confusion matrix after applying fuzzy membership

Table 4.7: Results of employing the fuzzy

model	accuracy%	precision%	recall%	F1%	AUC%
LR	85.159	83.094	88.485	85.705	85.14
LDA	84.967	83.088	88.017	85.481	84.95
NB	71.124	63.552	99.815	77.659	70.963
XGBoost	99.156	98.873	99.455	99.163	99.154

Because the suggested system's purpose is to achieve the best possible results, we attempted to determine how much the fuzzy membership function influences the system's efficiency. The outcomes showed a good improvement. As shown in Table 4.7, we obtained the following results. With the LR findings, accuracy is improved to 85.159 percent, whereas it was 82.231 percent before applying the fuzzy; recall improved from 85.258 percent to 88.485 percent; precision improved from 80.542 percent to 83.094; F1 improved from 82.833 percent to 85.705 percent, and AUC improved from 82.214 percent to 85.14 percent. With regard to the results of the LDA, the accuracy is improved to 84.967 percent, whereas it was 81.206 percent before applying the fuzzy membership; recall improved from 85.138 percent to 88.017 percent; precision improved from 79.084 percent to 83.088; F1 improved from 82 percent to 85.481 percent, and AUC improved from 81.184 percent to 84.95 percent. For the results of the NB, the changes are almost non-existent. Where all the results remained almost constant. As for the results of XGBoost, they were basically very high, so the improvements that occurred were very slight, with values close to parts of a hundred.

The final results extracted from all of the above prove that the best capabilities of the proposed system models were for XGBoost, then LR, LDA, and lastly NB.

4.4 COMPARISONS

We compared the performance of our work's top two classifiers, XGBoost and LR, to the performance of the top two classifiers in [63], [27], and [58]. By calculating accuracy, AUC, and precision for the same dataset, as indicated in Table 4.8.

Table 4.8: A comparing with other works

Author	Classifier	Dataset	Accuracy	Precision	AUC
[63]	XGBoost	European Cardholders	95.1%	88%	98%
[63]	Gradient boosting	European Cardhoders	0.947%	0.87%	0.92%
[27]	XGBoost	European Cardholders	99.962%		
[27]	Random Forest	European Cardholders	99.957%		
[58]	Random Forest	European Cardhoders	96.77%	100%	95.55%
[58]	LR	European Cardhoders	95.16%	95.34%	94.28%
our work	XGBoost	European Cardhoders	99.969%	99.938%	99.969%
our work	LR	European Cardhoders	94.752%	97.411%	94.756%

A comparison of the findings in Table 4.8 revealed that our study's XGBoost results are superior to those reported in [63], [27], and [58] using the same dataset, indicating that our results are superior.

5. CONCLUSIONS AND FUTURE WORKS

5.1 CONCLUSIONS

In this work, we presented a system for CCF detection. LR, LDA, NB, and the boosting technique XGBoost were employed to construct the system models and trained using the credit card dataset to detect fraud. The dataset used was two. The first was a European cardholder dataset, and the second was a Turkish dataset from IEEE Data Port. these datasets were extremely imbalanced, and a balance was achieved using SMOTE. Also, the fuzzy membership function was applied to the datasets. After that, the data sets were divided into two parts: 30 percent for testing and 70 percent for training. Following that, we analyzed the performance of the Algorithms. The top performance was by XGBoost in the two datasets. The accuracy, precision, recall, and AUC of the XGBoost model were all excellent. With the first dataset, the accuracy was 99.987 percent, precision was 99.986 percent, a recall was 99.987 percent, F1 was 99.987 percent, and AUC was 99.987 percent. The second dataset's accuracy was 99.156 percent, precision was 98.873 percent, the recall was 99.455 percent, F1 was 99.163 percent, and AUC was 99.154 percent.

The fuzzy membership function was applied to the features' values to acquire the best possible results. That resulted in an increase in system models performance. The disadvantage of using fuzzy membership is that it increases the time required because the number of features has grown. The employment of the fuzzy membership function is one of the accessible solutions from which scientists and researchers can profit in their respective research, specifically in the classification task.

At the end of our conclusion, we can say that the performance of the proposed system was good, but if the dataset used were larger and more realistic and without the need to employ strategies to fix the imbalance; The results will be better and higher.

5.2 FUTURE WORKS

Several suggestions are listed below to increase the system performance for future:

i. Using another balancing technique.

ii. Using more ML techniques, specifically more boosting algorithms such as Gradient Boosting, AdaBoost, and LightGBM.

iii. Reduce the number of features as much as possible by using the best feature selection techniques to decrease the time.

iv. Needing to get a larger dataset.

v. Needing to get a real balancing dataset.

vi. Needing to get a more real dataset from banks.

REFERENCES

[1] B. Al Smadi and M. Min, "A Critical review of Credit Card Fraud Detection Techniques," *2020 11th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2020*, pp. 0732–0736, 2020, doi: 10.1109/UEMCON51285.2020.9298075.

[2] M. Puh and L. Brkić, "Detecting credit card fraud using selected machine learning algorithms," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019 - Proceedings*, May 2019, pp. 1250–1255, doi: 10.23919/MIPRO.2019.8757212.

[3] A. Dal Pozzolo, "Adaptive Machine Learning for Credit Card Fraud Detection Declaration of Authorship," PhD Thesis, Department of Computer Science, Université Libre de Bruxelles, 2015.

[4] The Nilson Report, "Nilson Report – Card Fraud Losses Reach $27.85 Billion." 2019, Accessed: Mar. 22, 2021. [Online]. Available: https://nilsonreport.com/mention/407/1link/.

[5] A. Salazar, G. Safont, and L. Vergara, "Semi-Supervised Learning for Imbalanced Classification of Credit Card Transaction," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2018-July, pp. 1–7, 2018, doi: 10.1109/IJCNN.2018.8489755.

[6] J. Gao, Z. Zhou, J. Ai, B. Xia, and S. Coggeshall, "Predicting Credit Card Transaction Fraud Using Machine Learning Algorithms," *J. Intell. Learn. Syst. Appl.*, vol. 11, no. 03, pp. 33–63, Aug. 2019, doi: 10.4236/jilsa.2019.113003.

[7] O. S. Yee, S. Sagadevan, and N. H. A. H. Malim, "Credit card fraud detection using machine learning as data mining technique," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 1–

4, pp. 23–27, 2018.

[8] A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, and P. Beling, “Deep learning detecting fraud in credit card transactions,” *2018 Syst. Inf. Eng. Des. Symp. SIEDS 2018*, pp. 129–134, 2018, doi: 10.1109/SIEDS.2018.8374722.

[9] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, “Random forest for credit card fraud detection,” *ICNSC 2018 - 15th IEEE Int. Conf. Networking, Sens. Control*, pp. 1–6, 2018, doi: 10.1109/ICNSC.2018.8361343.

[10] F. Carcillo, Y. A. Le Borgne, O. Caelen, and G. Bontempi, “Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization,” *Int. J. Data Sci. Anal.*, vol. 5, no. 4, pp. 285–300, 2018, doi: 10.1007/s41060-018-0116-z.

[11] S. Maes, K. Tuyls, and B. Vanschoenwinkel, “Credit Card Fraud Detection Using Bayesian and Neural Networks,” *Maciunas RJ, Ed. Interact. image-guided neurosurgery. Am. Assoc. Neurol. Surg.*, 2002.

[12] G. Singh, R. Gupta, A. Rastogi, M. D. S. Chandel, and A. Riyaz, “SVM01:A Machine Learning Approach for Detection of Fraud based on SVM,” *Int. J. Sci. Eng. Technol.*, vol. 198, no. 1, pp. 194–198, 2012.

[13] K. Ramakalyani and D. Umadevi, “Fraud Detection of Credit Card Payment System by Genetic Algorithm,” *Int. J. Sci. Eng. Res.*, vol. 3, no. 7, pp. 1–6, 2012, [Online]. Available: http://www.ijser.org/researchpaper/Fraud-Detection-of-Credit-Card-Payment-System-by-Genetic-Algorithm.pdf.

[14] J. R. Gaikwad, A. B. Deshmane, H. V Somavanshi, S. V Patil, and R. A. Badgujar, “Credit

Card Fraud Detection using Decision Tree Induction Algorithm," *Int. J. Innov. Technol. Explor. Eng.*, no. 6, pp. 2278–3075, 2014.

[15] Ogwueleka and F. Nonyelum, "Data mining application in credit card fraud detection system," *J. Eng. Sci. Technol.*, vol. 6, pp. 311–322, 2011.

[16] A. C. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten, "Improving credit card fraud detection with calibrated probabilities," *SIAM Int. Conf. Data Min. 2014, SDM 2014*, vol. 2, pp. 677–685, 2014, doi: 10.1137/1.9781611973440.78.

[17] C. Sudha and D. Akila, "Credit card fraud detection system based on operational transaction features using SVM and random forest classifiers," in *Proceedings of 2nd International Conference on Computation, Automation and Knowledge Management, ICCAKM 2021*, Jan. 2021, pp. 133–138, doi: 10.1109/ICCAKM50778.2021.9357709.

[18] M. R. Dileep, A. V Navaneeth, and M. Abhishek, "A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms," in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Feb. 2021, pp. 1025–1028, doi: 10.1109/ICICV50876.2021.9388431.

[19] T. C. Tran and T. K. Dang, "Machine Learning for Prediction of Imbalanced Data: Credit Fraud Detection," in *Proceedings of the 2021 15th International Conference on Ubiquitous Information Management and Communication, IMCOM 2021*, Jan. 2021, pp. 1–7, doi: 10.1109/IMCOM51814.2021.9377352.

[20] K. Janvitha, C. R. S. Vasavi, A. Sruthi, K. Praharshitha, and D. K. Anguraj, "Survey on Detection of Credit Card Frauds using Hmm and various Clustering Approaches," in

Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021, Jan. 2021, pp. 101–107, doi: 10.1109/ICICT50816.2021.9358773.

[21] A. RB and S. K. KR, "Credit Card Fraud Detection Using Artificial Neural Network," *Glob. Transitions Proc.*, Jan. 2021, doi: 10.1016/j.gltp.2021.01.006.

[22] D. Shaohui, G. W. Qiu, H. Mai, and H. Yu, "Customer Transaction Fraud Detection Using Random Forest," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, Jan. 2021, pp. 144–147, doi: 10.1109/ICCECE51280.2021.9342259.

[23] X. Kewei, B. Peng, Y. Jiang, and T. Lu, "A Hybrid Deep Learning Model for Online Fraud Detection," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, Jan. 2021, pp. 431–434, doi: 10.1109/ICCECE51280.2021.9342110.

[24] W. Deng, Z. Huang, J. Zhang, and J. Xu, "A Data Mining Based System for Transaction Fraud Detection," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering, ICCECE 2021*, Jan. 2021, pp. 542–545, doi: 10.1109/ICCECE51280.2021.9342376.

[25] A. Shivanna, S. Ray, K. Alshouiliy, and D. P. Agrawal, "Detection of Fraudulence in Credit Card Transactions using Machine Learning on Azure ML," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2020*, Oct. 2020, pp. 0268–0273, doi: 10.1109/UEMCON51285.2020.9298129.

[26] R. Sailusha, V. Gnaneswar, R. Ramesh, and G. Ramakoteswara Rao, "Credit Card Fraud

Detection Using Machine Learning," *Proc. Int. Conf. Intell. Comput. Control Syst. ICICCS 2020*, no. Iciccs, pp. 1264–1270, 2020, doi: 10.1109/ICICCS48265.2020.9121114.

[27] V. Jain, M. Agrawal, and A. Kumar, "Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection," *ICRITO 2020 - IEEE 8th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir.*, pp. 86–88, 2020, doi: 10.1109/ICRITO48877.2020.9197762.

[28] M. Nur-E-Arefin and M. Sultan, "A Comparative Study of Machine Learning Classifiers for Credit Card Fraud Detection," *Int. J. Innov. Technol. Interdiscip. Sci.*, vol. 3, no. 1, pp. 395–406, 2020.

[29] S. Khatri, A. Arora, and A. P. Agrawal, "Supervised Machine Learning Algorithms for Credit Card Fraud Detection: A Comparison," *2020 10th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 680–683, 2020.

[30] Y. K. Saheed, M. A. Hambali, M. O. Arowolo, and Y. A. Olasupo, "Application of GA Feature Selection on Naive Bayes, Random Forest and SVM for Credit Card Fraud Detection," in *2020 International Conference on Decision Aid Sciences and Application, DASA 2020*, Nov. 2020, pp. 1091–1097, doi: 10.1109/DASA51403.2020.9317228.

[31] A. H. Alhazmi and N. Aljehane, "A Survey of Credit Card Fraud Detection Use Machine Learning," in *2020 International Conference on Computing and Information Technology, ICCIT 2020*, Sep. 2020, doi: 10.1109/ICCIT-144147971.2020.9213809.

[32] F. Itoo, Meenakshi, and S. Singh, "Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection," *Int. J. Inf.*

Technol., pp. 1–9, Feb. 2020, doi: 10.1007/s41870-020-00430-y.

[33] 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.

[34] J. V. V. Sriram Sasank, G. R. Sahith, K. Abhinav, and M. Belwal, "Credit Card Fraud Detection Using Various Classification and Sampling Techniques: A Comparative Study," in *Proceedings of the 4th International Conference on Communication and Electronics Systems, ICCES 2019*, Jul. 2019, pp. 1713–1718, doi: 10.1109/ICCES45898.2019.9002289.

[35] O. Adepoju, J. Wosowei, S. Lawte, and H. Jaiman, "Comparative Evaluation of Credit Card Fraud Detection Using Machine Learning Techniques," *2019 Glob. Conf. Adv. Technol. GCAT 2019*, pp. 1–6, 2019, doi: 10.1109/GCAT47503.2019.8978372.

[36] P. Raghavan and N. El Gayar, "Fraud Detection using Machine Learning and Deep Learning," in *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019*, Dec. 2019, pp. 334–339, doi: 10.1109/ICCIKE47802.2019.9004231.

[37] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *Int. J. Syst. Assur. Eng. Manag.*, vol. 8, no. 2, pp. 937–953, Nov. 2017, doi: 10.1007/s13198-016-0551-y.

[38] 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, doi: 10.5120/8184-1538.

[39] J. Pun and Y. Lawryshyn, “Improving Credit Card Fraud Detection using a Meta-Classification Strategy,” *Int. J. Comput. Appl.*, vol. 56, no. 10, pp. 41–46, Dec. 2012, doi: 10.5120/8930-3007.

[40] N. Wong, P. Ray, G. Stephens, and L. Lewis, “Artificial immune systems for the detection of credit card fraud: An architecture, prototype and preliminary results,” *Inf. Syst. J.*, vol. 22, no. 1, pp. 53–76, Jan. 2012, doi: 10.1111/j.1365-2575.2011.00369.x.

[41] Y. Sahin and E. Duman, “Detecting credit card fraud by decision trees and support vector machines,” in *IMECS 2011 - International MultiConference of Engineers and Computer Scientists 2011*, 2011, vol. 1, pp. 442–447, Accessed: Apr. 18, 2021. [Online]. Available: https://openaccess.dogus.edu.tr/xmlui/handle/11376/2366.

[42] Ehramikar, “The Enhancemeat of Credit Card Fraud Detection Systems using Machine Learning Methodology,” *collectionscanada.gc.ca*2000 „ Accessed: May 16, 2021. [Online]. Available: https://www.collectionscanada.gc.ca/obj/s4/f2/dsk1/tape3/PQDD_0023/MQ50338.pdf.

[43] E. Duman and M. H. Ozcelik, “Detecting credit card fraud by genetic algorithm and scatter search,” *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13057–13063, Sep. 2011, doi: 10.1016/j.eswa.2011.04.110.

[44] J. T. S. Quah and M. Sriganesh, “Real-time credit card fraud detection using computational intelligence,” *Expert Syst. Appl.*, vol. 35, no. 4, pp. 1721–1732, Nov. 2008, doi: 10.1016/j.eswa.2007.08.093.

[45] K. Modi and R. Dayma, “Review on fraud detection methods in credit card transactions,”

in *Proceedings of 2017 International Conference on Intelligent Computing and Control, I2C2 2017*, Mar. 2018, vol. 2018-Janua, pp. 1–5, doi: 10.1109/I2C2.2017.8321781.

[46] K. Chaudhary, J. Yadav, and B. Mallick, "A review of Fraud Detection Techniques: Credit Card," in *International Journal of Computer Applications*, 2012, vol. 45, no. 1, pp. 39--44.

[47] J. Pieprzyk, H. Ghodosi, and E. Dawson, Eds., *Information Security and Privacy*, vol. 4586. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

[48] P. Richhariya, "A Survey on Financial Fraud Detection Methodologies," *Int. J. Commer. Bus. Manag.*, vol. 1, no. 1, pp. 14–24, 2012, Accessed: May 08, 2021. [Online]. Available: https://www.academia.edu/download/46599355/pxc3879373.pdf.

[49] L. Delamaire, H. Abdou, and J. Pointon, "Credit card fraud and detection techniques: A review," *Banks and Bank Systems*, vol. 4, no. 2. pp. 57–68, 2009, Accessed: May 09, 2021. [Online]. Available: http://eprints.hud.ac.uk/19069/1/AbdouCredit.pdf.

[50] P. JRana and J. Baria, "A Survey on Fraud Detection Techniques in Ecommerce," *Int. J. Comput. Appl.*, vol. 113, no. 14, pp. 5–7, 2015, doi: 10.5120/19892-1898.

[51] C. Phua, R. W. Gayler, V. C. S. Lee, and K. Smith-Miles, "On the Approximate Communal Fraud Scoring of Credit Applications," in *Credit Scoring and Credit Control IX*, 2005, pp. 1–10, Accessed: May 10, 2021. [Online]. Available: https://www.researchgate.net/publication/215991909.

[52] A. C. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten, "Cost sensitive credit card fraud detection using bayes minimum risk," in *Proceedings - 2013 12th International Conference on Machine Learning and Applications, ICMLA 2013*, 2013, vol. 1, pp. 333–338, doi:

10.1109/ICMLA.2013.68.

[53] X. Niu, L. Wang, and X. Yang, “A comparison study of credit card fraud detection: Supervised versus unsupervised,” *arXiv*, 2019.

[54] N. Sethi and A. Gera, “A Revived Survey of Various Credit Card Fraud Detection Techniques,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 3, no. 4, pp. 780–791, 2014.

[55] A. Pumsirirat and L. Yan, “Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 1, pp. 18–25, 2018, doi: 10.14569/IJACSA.2018.090103.

[56] R. Jain, B. Gour, and S. Dubey, “A Hybrid Approach for Credit Card Fraud Detection using Rough Set and Decision Tree Technique,” *Int. J. Comput. Appl.*, vol. 139, no. 10, pp. 1–6, Apr. 2016, doi: 10.5120/ijca2016909325.

[57] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, “Credit card fraud detection using machine learning techniques: A comparative analysis,” in *Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017*, 2017, vol. 2017-Janua, pp. 1–9, doi: 10.1109/ICCNI.2017.8123782.

[58] D. Tanouz, R. R. Subramanian, D. Eswar, G. V. P. Reddy, A. R. Kumar, and C. H. V. N. M. Praneeth, “Credit card fraud detection using machine learning,” *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, no. Iciccs, pp. 967–972, 2021, doi: 10.1109/ICICCS51141.2021.9432308.

[59] P. Roy, P. Rao, J. Gajre, K. Katake, A. Jagtap, and Y. Gajmal, “Comprehensive analysis for fraud detection of credit card through machine learning,” *2021 Int. Conf. Emerg. Smart*

Comput. Informatics, ESCI 2021, pp. 765–769, 2021, doi: 10.1109/ESCI50559.2021.9397029.

[60] B. K. Padhi, S. Chakravarty, and B. N. Biswal, "Anonymized credit card transaction using machine learning techniques," in *Lecture Notes in Networks and Systems*, vol. 109, Springer, Singapore, 2020, pp. 413–423.

[61] A. Bhushan Sharma and B. Singh, "Performance Evaluation and Identification of Optimal Classifier for Credit Card Fraudulent Detection," in *Studies in Computational Intelligence*, Springer, Cham, 2021, pp. 137–155.

[62] S. Mittal and S. Tyagi, "Performance evaluation of machine learning algorithms for credit card fraud detection," *Proc. 9th Int. Conf. Cloud Comput. Data Sci. Eng. Conflu. 2019*, pp. 320–324, 2019, doi: 10.1109/CONFLUENCE.2019.8776925.

[63] H. Feng, "Ensemble learning in credit card fraud detection using boosting methods," *Proc. - 2021 2nd Int. Conf. Comput. Data Sci. CDS 2021*, pp. 7–11, Jan. 2021, doi: 10.1109/CDS52072.2021.00009.

[64] V. Vaishali, "Fraud Detection in Credit Card by Clustering Approach," *Int. J. Comput. Appl.*, vol. 98, no. 3, pp. 29–32, Jul. 2014, doi: 10.5120/17164-7225.

[65] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.

[66] P. C. Sen, M. Hajra, and M. Ghosh, "Supervised Classification Algorithms in Machine Learning: A Survey and Review," in *Advances in Intelligent Systems and Computing*, 2020,

vol. 937, pp. 99–111, doi: 10.1007/978-981-13-7403-6_11.

[67] J. Basak and R. Kothari, “A classification paradigm for distributed vertically partitioned data,” *Neural Comput.*, vol. 16, no. 7, pp. 1525–1544, Jul. 2004, doi: 10.1162/089976604323057470.

[68] R. Lopez De Mantaras and E. Armengol, “Machine learning from examples: Inductive and Lazy methods,” *Data Knowl. Eng.*, vol. 25, no. 1–2, pp. 99–123, Mar. 1998, doi: 10.1016/S0169-023X(97)00053-0.

[69] G. E. A. P. A. Batista and M. C. Monard, “An analysis of four missing data treatment methods for supervised learning,” *Appl. Artif. Intell.*, vol. 17, no. 5–6, pp. 519–533, May 2003, doi: 10.1080/713827181.

[70] F. Ahmed and R. Shamsuddin, “A comparative study of credit card fraud detection using the combination of machine learning techniques with data imbalance solution,” *Proc. - 2021 2nd Int. Conf. Comput. Data Sci. CDS 2021*, pp. 112–118, 2021, doi: 10.1109/CDS52072.2021.00026.

[71] K. Suresh and R. Dillibabu, “Designing a Machine Learning Based Software Risk Assessment Model Using Naïve Bayes Algorithm,” *TAGA J.*, vol. 14, pp. 3141–3147, 2018.

[72] I. D. Dinov, “Probabilistic Learning: Classification Using Naive Bayes,” in *Data Science and Predictive Analytics*, Springer, Cham, 2018, pp. 289–305.

[73] A. Tripathy and S. K. Rath, “Classification of Sentiment of Reviews using Supervised Machine Learning Techniques,” *Int. J. Rough Sets Data Anal.*, vol. 4, no. 1, pp. 56–74, 2016, doi: 10.4018/ijrsda.2017010104.

[74] A. Balakrishnama, Suresh and Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Inst. Signal Inf. Process.*, vol. 18, pp. 1--8, 1998.

[75] O. RAJU, "CREDIT CARD FRAUD DETECTION USING XGBOOST CLASSIFIER," *Int. J. Techno-Engineering*, vol. XIII, no. III, pp. 175–187, 2021, [Online]. Available: http://ijte.uk/archive/2021/volume3/PAPER-Credit-card-fraud-detection.pdf.

[76] A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, "Learning from Imbalanced Data Sets," *Learn. from Imbalanced Data Sets*, 2018, doi: 10.1007/978-3-319-98074-4.

[77] K. Li, W. Zhang, Q. Lu, and X. Fang, "An improved SMOTE imbalanced data classification method based on support degree," in *Proceedings - 2014 International Conference on Identification, Information and Knowledge in the Internet of Things, IIKI 2014*, Mar. 2014, pp. 34–38, doi: 10.1109/IIKI.2014.14.

[78] L. Demidova and I. Klyueva, "SVM classification: Optimization with the SMOTE algorithm for the class imbalance problem," *2017 6th Mediterr. Conf. Embed. Comput. MECO 2017 - Incl. ECYPS 2017, Proc.*, Jul. 2017, doi: 10.1109/MECO.2017.7977136.

[79] N. V. Chawla, A. Lazarevic, L. O. Hall, and K. W. Bowyer, "SMOTEBoost: Improving prediction of the minority class in boosting," in *Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science)*, 2003, vol. 2838, pp. 107–119, doi: 10.1007/978-3-540-39804-2_12.

[80] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015, doi:

10.5121/ijdkp.2015.5201.

[81] T. M. Khoshgoftaar and N. Seliya, “Comparative assessment of software quality classification techniques: An empirical case study,” *Empir. Softw. Eng.*, vol. 9, no. 3, pp. 229–257, Sep. 2004, doi: 10.1023/B:EMSE.0000027781.18360.9b.

[82] M. Kubat, R. C. Holte, and S. Matwin, “Machine learning for the detection of oil spills in satellite radar images,” *Mach. Learn.*, vol. 30, no. 2–3, pp. 195–215, 1998, doi: 10.1023/a:1007452223027.

[83] N. J. Pizzi, A. R. Summers, and W. Pedrycz, “Software quality prediction using median-adjusted class labels,” in *Proceedings of the International Joint Conference on Neural Networks*, 2002, vol. 3, pp. 2405–2409, doi: 10.1109/ijcnn.2002.1007518.

[84] N. J.-I. P. of the 3rd W. on Evaluation and undefined 2008, “Classifier evaluation: A need for better education and restructuring,” *site.uottawa.ca*, Accessed: Jun. 25, 2021. [Online]. Available: http://www.site.uottawa.ca/ICML08WS/papers/N_Japkowicz.pdf.

[85] M. Sokolova, N. Japkowicz, and S. Szpakowicz, “Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation,” in *AAAI Workshop - Technical Report*, 2006, vol. WS-06-06, pp. 24–29, doi: 10.1007/11941439_114.

[86] N. Seliya, T. M. Khoshgoftaar, and J. Van Hulse, “A study on the relationships of classifier performance metrics,” in *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 2009, pp. 59–66, doi: 10.1109/ICTAI.2009.25.

[87] J. Davis and M. Goadrich, “The relationship between precision-recall and ROC curves,” in *ACM International Conference Proceeding Series*, 2006, vol. 148, pp. 233–240, doi:

10.1145/1143844.1143874.

[88] A. D. Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, “Calibrating probability with undersampling for unbalanced classification,” in *Proceedings - 2015 IEEE Symposium Series on Computational Intelligence, SSCI 2015*, 2015, pp. 159–166, doi: 10.1109/SSCI.2015.33.

[89] Yapı Kredi Teknoloji Company, “Credit Card Fraud Detection - Turkish Data,” *IEEE Dataport*, 2021. https://dx.doi.org/10.21227/88a1-zm60 (accessed Nov. 28, 2021).

[90] K. M. A. Alheeti, A. Gruebler, and K. D. McDonald-Maier, “An intrusion detection system against malicious attacks on the communication network of driverless cars,” in *2015 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015*, Jul. 2015, pp. 916–921, doi: 10.1109/CCNC.2015.7158098.

[91] K. M. Ali, W. Venus, and M. S. Al Rababaa, “The affect of fuzzification on neural networks intrusion detection system,” in *2009 4th IEEE Conference on Industrial Electronics and Applications, ICIEA 2009*, 2009, pp. 1236–1241, doi: 10.1109/ICIEA.2009.5138399.

[92] M. Hossin, Mohammad Sulaiman, “A review on evaluation metrics for data classification evaluations,” *Int. J. Data Min. \& Knowl. Manag. Process*, vol. 5, 2015.