Credit Card Fraud Detection using Machine Learning Techniques

A Thesis Submitted

In Partial Fulfillment of the Requirements
for the Degree of

MASTER OF TECHNOLOGY

in

Artificial Intelligence

by

Mayank Pathak

(Roll No. 2K23/AFI/30)

Under the Supervision of

Dr. Rohit Beniwal

(Dept of Computer Science & Engineering)

To the
Department of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042. India

May, 2025

ACKNOWLEDGEMENTS

I have taken efforts in this survey paper. However, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend my sincere thanks to all of them.

I am highly indebted to **Dr. Rohit Beniwal** for his guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in completing this review paper. I would like to express my gratitude towards the **Head of the Department (Computer Science and Engineering, Delhi Technological University)** for their kind cooperation and encouragement which helped me in the completion of this research survey. I would like to express my special gratitude and thanks to all the Computer Science and Engineering staff for giving me such attention and time.

My thanks and appreciation also go to my colleague in writing the survey paper and the people who have willingly helped me out with their abilities.

Mayank Pathak
31 May 2024

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I, Mayank Pathak, Roll No. 2K23/AFI/30 student of M.Tech (Artificial Intelligence), hereby certify that the work which is being presented in the thesis entitled "**Credit Card Fraud Detection using Machine Learning**" in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Artificial Intelligence in the Department of Computer Science and Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from August 2023 to Jun 2025 under the supervision of Dr Rohit Beniwal, Asst Prof, Dept of Computer Science and Engineering. The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Place: Delhi

Candidate's Signature

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE

Certified that **Mayank Pathak** (Roll No. 2K23/AFI/30) has carried out the research work presented in the thesis titled "**Credit Card Fraud Detection using Machine Learning**", for the award of Degree of Master of Technology from Department of Computer Science and Engineering, Delhi Technological University, Delhi under my supervision. The thesis embodies result of original work and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree for the candidate or submit else from the any other University /Institution.

Dr. Rohit Beniwal
(Supervisor)

Department of CSE

Date:

Delhi Technological University

ABSTRACT

The banks and financial industries are seriously threatened by credit card theft. This research investigates how well different Deep Learning (DL) and Machine Learning (ML) models identify fraudulent transactions. The primary objective is to analyze and compare different strategies for fraud detection to develop a more reliable and accurate decision-making system. The research reviews the challenges in fraud detection and presents solutions by highlighting both established and emerging fraud patterns.

In recent years, the rapid increase in online payments through credit cards and UPI has been accompanied by a corresponding rise in fraudulent activities. Fraudsters employ a wide range of techniques such as card theft, swapping, phishing, and large-scale data breaches to obtain sensitive card information. Due to the high volume of genuine transactions and the limited number of fraud cases, the transaction datasets are often extremely imbalanced, leading to challenges such as model bias, poor generalization, and misleading performance metrics.

In order to balance the dataset, the research uses Generative Adversarial Networks (GANs) to create artificial samples of the minority (fraudulent) class. The Synthetic Minority Oversampling Technique (SMOTE) is used to compare the performance of GANs. Experimental results show that GAN-based resampling yields better classification performance, particularly in terms of F1-score. Given its ability to manage high-dimensional and imbalanced data, GAN proves to be a powerful tool for financial fraud detection. The European Cardholders 2013 dataset is used to evaluate the models. This study demonstrates that incorporating GANs can significantly enhance fraud detection systems and emphasizes the need for continuous innovation to address evolving fraud patterns and improve financial security.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

GAN	Generative Adversarial Network
CCFD	Credit Card Fraud Detection
DL	Deep Learning
SLR	Systematic Literature Review
RNN	Recurrent Neural Network
LLM	Large Language Model
AI	Artificial Intelligence
SVM	Support Vector Machine
ML	Machine Learning
MLP	Multi Layer Perceptron
DT	Decision Tree
GRU	Gated Recurrent Unit
DNN	Deep Neural Network
LSTM	Long Short Term Memory
CNN	Convolutional Neural Network
KNN	K-Nearest Neighbor
SMOTE	Synthetic Minority Over-sampling Technique

CHAPTER 1
Introduction

1.1 OVERVIEW

Online e-commerce portals make a straight-up trajectory in terms of the number of online transactions. During the Corona period, the sudden rise of online transactions grew by 26 % in 2020. Due to the fact that e-commerce companies are now offering large discounts based on credit cards, consumers are shifting to online shopping, which causes the number of credit card transactions to suddenly increase. As a result, banks and e-commerce sites make a significant amount of money through these marketing strategies.

Juniper Research reports that in 2021, the amount of damages caused by fraud increased to $20 billion, from $17.5 billion in 2020. It is typical practice for fraudsters to get credit card details via phishing, phone calls, or text messages. Then, they use this information to make unauthorized transactions. This just serves to emphasize how critical it is for financial institutions to have robust fraud detection systems. The ever-increasing use of credit cards throughout the globe has made CCFD a crucial domain for ML models.

There are four main categories into which these models fall: supervised, unsupervised, semi-supervised, and reinforcement learning. The identification of fraudulent transactions is one area where supervised learning is very important. This method uses a labelled dataset for training ML models, which are then tested on an unlabelled dataset. A transaction may be classified as either valid or fraudulent with the use of a label.

Modern technology makes use of a sophisticated model called an Artificial Neural Network (ANN). In terms of both operation and structure, an ANN model mimics the way the human brain does it. A multilayer perceptron, or deep learning (DL), is an integral part of ANN. DL is a useful technique for CCFD since it can analyze complex patterns and extract high-level features. In fraud detection, banks are crucial in reviewing transactions for potential fraud before authorizing payments, often by verifying if the associated website appears on a block list. If the transaction is linked to a blacklisted site, it is flagged as fraudulent and rejected.

Fraud detection makes use of a number of approaches, such as statistical methods, ML, and DL.Imbalanced datasets are one of the most challenging issues in detecting fraud transactions. In imbalanced data, very small chunks of transactions are considered fraud, and almost all transactions are considered not fraud. Due to this data, it can pose a biased model result and suboptimal fraud detection capabilities.

1.2 MOTIVATION

The increasing evolution of digital platforms and transactions offer ease of use to consumers;

By employing machine learning techniques such as data balancing, undersampling, and oversampling to handle unbalanced credit card data, several studies attempt to address this issue. Comprehensive research on these methods' efficacy is still lacking, though. To identify a fraudulent transaction, ensemble learning models and techniques are crucial. Several models are combined in ensemble learning to produce detection that is more reliable and accurate.

however, security risks continue to remain a major threat to institutions. Credit cards have especially become a target for fraud and financial scam activities, causing a great amount of concern for banks, clients, and merchants. Each year, the amounts being scammed fortify and become more intimidating to established financial systems.

The structure of transaction data, consisting of verified users and fraudulent users, is chiefly disproportionate, making the detection of fraud much harder than it already is. Likely, causing a lack of performance in classification model procedures. With the changing patterns employed by criminals trying to outsmart the system- capturing new and old fraud patterns calls for intelligent detection systems.

This thesis focuses on the need to address the imbalance issue within class models and defy the patterns used by scammers employing advanced machine learning algorithms. Markedly, innovation can stem from implementing Generative Adversarial Networks (GANs) for model generalization to tackle with balance class problems. In particular, a potential path toward innovation is the use of Generative Adversarial Networks (GANs) to rectify class imbalance and enhance model generalization. The project intends to aid in the creation of reliable fraud detection systems that can instantly adjust to changing threats by fusing traditional machine learning algorithms with deep learning techniques and data balancing tactics.

The motivation behind this research stems from both the technical challenge and the real-world impact: improving financial security, reducing economic losses, and safeguarding consumer trust in digital transactions.

CHAPTER 2

LITERATURE REVIEW

As shown in Table 1, this section offers a thorough literature review of the methods and resources used for successful transaction fraud detection. As gathered for this systematic literature review (SLR), it provides an overview of the most current studies on the use of classification techniques in CCFD. Author information, publication year, datasets, methodology, targeted domains, key performance indicators (KPIs), accuracy, and other features are collected from the chosen research, as described in the data extraction step. Additional performance indicators include confusion matrix (CM), F score/F1, recall (R), accuracy (A), specificity (S), area under the curve (AUC), precision (P), and F score/F1 score (F) and geometric mean (GM), it also contains specific information about various dataset types, including European Card, Brazilian, IEEE-CIS fraud detection, and Chinese commercial banks.

To improve CCFD, Prasad Chowdary et al. suggest using the ensemble learning technique [1]. By incorporating DL models, identification errors can be fixed and false negatives can be decreased. The author used DT, LR, DT, SVM, RF, and a Gradient Boosting Classifier (XGBoost) in this work. The author compares various algorithms using a number of assessment parameters and finds that DT has the greatest 100% recall percentage. XGBoost, LR, Random forest (RF), and SVM have the next best recall values, with 85.0%, 74.50%, 75.90%, and 69.0%, respectively.
The CCFD algorithm was developed by Sahithi et al. in a 2022 paper [2]. A Weighted Average Ensemble was utilized by the author to integrate RF, LR, KNN, Adaboost, and Bagging. Adaboost (97.91%), LR (98.900%), RF Bagging (98.91%), KNN (97.81%), and Bagging (95.37%) were all followed by their model, which demonstrated a 99 percent accuracy. This author used data from the European Credit Card Company.

Qaddoura et al. [3] tested the efficacy of the SVM oversampling algorithm for CCFD, as well as the oversampling techniques SMOTE, ADASYN, borderline 1, and borderline 2. The study made use of LR, RF, NB, KNN, DT, and SVM. Oversampling may improve model performance, according to the paper's author, however the exact method depends on the machine learning algorithm. The computational overhead has an impact on this model's applicability in real-world situations.

To identify the fraud activity, Forough et al. [4] proposed a unique voting mechanism based on the ANN and an ensemble model based on the sequential modeling of deep recurrent neural networks. Recurrent Neural Networks (RNNs) are used as voting mechanisms in the author's proposed model, which combines the output of these networks using Feed Forward Neural Networks (FFNNs) with either LSTM or GRU networks acting as base classifiers. The GRU ensemble model, which used two base classifiers, produced the best results on the Brazilian and European card datasets. In the majority of measures, both the baseline ensemble model and the single GRU model perform well. Nevertheless, the limitations of the author's suggested ensemble model-based sequential data modeling with deep RNN and a unique voting mechanism cannot be adequately addressed.
Esenogho et al. [5] combined an ensemble classifier based on neural networks with a hybrid

data resampling technique in their ground-breaking approach to CCFD. The Adaboost framework and LSTM neural networks are used to create their ensemble classifier. The SMOTE-ENN hybrid resampling method would rectify any imbalance in the data. This method combines the oversampling strategy of SMOTE with the undersampling approach of ENN to produce a more balanced dataset. For the minority class, it creates synthetic samples, and for the majority class, it eliminates noisy samples. Although SMOTE-ENN is utilized to solve data imbalance concerns, researchers have not investigated the impact of altering the neural network topology or hyperparameters on the performance of the proposed model.

The effectiveness of the machine learning method was examined by Varmedja et al. [6] in relation to CCFD. The European Credit Cardholder Dataset was a source of information for the writer of this article. The author tackled the problem of class imbalance by using the SMOTE oversampling approach. The performance of the suggested method was evaluated using the following machine learning techniques: NB, RF, and multilayer perceptron (MLP). With a fraud accuracy of 99.96%, the RF algorithm performed well in this paper. The accuracy of the other algorithms, NB and MLP, were 99.23% and 99.93% respectively. The author acknowledges the need for additional research to apply feature selection techniques that could increase the precision of other machine learning techniques.
CNN, a DL technique related to text processing and baseline model, was performed in CCFD by F. K. Alarfaj et al. [7]. The results of using these methods in CCFD are superior to those of traditional algorithms. With an accuracy of 99.79%, CNN with 20 layers and the baseline model is the best method when comparing the performance of all the methods side by side.

In this work [8], ensemble learning strategies such gradient boosting (LightGBM and LiteMORT) are evaluated by combining weighted and basic averaging techniques. We can reduce the amount of errors committed while increasing efficiency and accuracy by integrating them. We compared the models using weighted averaging and the measures of AUC, F1-score, Accuracy, Precision, and Recall. The greatest percentages were 95.20%, 91.66\%, 91.67%, and 99.44% for the combination of LightGBM and LiteMORT. Using information from the Kaggle website, the IEEE-CIS fraudulent dataset was analyzed. In an experiment conducted by Palak.G et al. [9] on an imbalanced dataset, the XGBoost model achieved a precision score of 0.91 and an accuracy score of 0.99. Also, XGBoost was able to control data imbalance when used in conjunction with the Random Oversampling method, as it achieved precision and accuracy scores of 0.99.
The authors of [10] presented a method that successfully models sequential data by combining LSTM-based deep recurrent neural networks with attention mechanisms to enhance CCFD. This approach allows the model to focus on the most pertinent transactions within the series by taking into account the temporal nature of transaction sequences. In comparison to previous techniques, the strategy seeks to achieve more accuracy by concentrating on identifying significant transactions that are suggestive of fraudulent conduct.

S.No	Author	Publication	Year	Techniques	Data	Domain	Performane Metrics	Accuracy	Tools
1	Prasad, P.Y. et al.	IEEE	2023	CNN, SVM, DT, RF, LR	European Credit Card	Finance	Precision, Recall, Accuracy, F1-score	CNN model's performance F1 score, recall, accuracy, and precision (AUC): 99.9\%, 85.71\%, 93\%, and 98\%.	Scikit-learn
2	Sahithi, G.L. et al.	IEEE	2022	RF, Adaboost, LR, KNN, Bagging	European Credit Card	Finance	RF, Adaboost, LR, KNN, Bagging	Proposed model(Weighted average)score 99\%accuracy	Scikit-learn
3	Qaddoura, R.; Biltawi, M.M	IEEE	2022	SMOTE, ADASYN, RF, KNN, SVM	European Credit Card	Finance	Geometric Mean, Recall, Accuracy, F1-score	Best F1: 88\% (RF with SVM SMOTE)	Scikit-learn
4	Bakhtiari et al.	Springer	2023	LightGBM, LiteMORT (Weighted Avg)		Finance	AUC, Precision, Recall, Accuracy, F1-score	By applying weighted averaging, the optimal results for combining LightGBM and LiteMORT were achieved,95.20, 90.65, 91.67, 92.79, and 99.44 for AUC, Recall, F1-score, Precision, and Accuracy, respectively	Scikit-learn
5	Forough, J. et al.	Elsiver	2021	LSTM	European Credit Card	Finance	F1-score, Precision, Recall, AUC and Confusion Matrix	LSTM based model show improvement compared to their l model	Keras, Scikit-learn
6	Esenogho et al.	IEEE	2022	Proposed LSTM Ensemble	European Credit Card		Specificity, Recall, AUC .	Proposed LSTM ensemble with SMOTE-ENN AUC score0.99	Keras, Scikit-learn

S.No	Author	Publication	Year	Techniques	Data	Domain	Performane Metrics	Accuracy	Tools
7	Varmedja et al.	IEEE	2019	LR, RF, Multilayer Perceptron	European Credit Card	Finance	Precision, Recall, Accuracy, Confusion Matrix	RF shows the highest accuracy among all models about 99.96\%.	Keras, Scikit-learn
8	I.Malik et al.	IEEE	2022	CNN	European,Brazilian ,China Card dataset	Finance	Accuracy, F1-score, AUC, Precision	CNN with a balanced dataset has a 96\%accuracy.	Keras
9	Qaddoura, R.; Biltawi, M.M	Elsiver	2023	SMOTE, ADASYN, RF, KNN, SVM	European Credit Card	Finance	Precision, Recall, Accuracy, F1-score,Confusion matrix	Using XGboost for random oversampling produced an F1 score of 99.98\%.	Scikit-learn ,XGboost classifer
10	I.Benchaji et al.	Springer	2021	LSTM, Attention Mechanism	European Credit Card	Finance	Precision, Recall, Specificity, Accuracy, AUC,Confusion Matrix	Attention mechanism and LSTM achieved an accuracy of 94	Keras

Table .1 Summary of the studies undertaken for review

CHAPTER 3

Methodology

3.1 DATASET

A. European Credit Card

Sourced via Kaggle in 2013, the dataset includes 284K credit card transaction records from European transac- tions. The data has been anonymized in order to preserve cardholder confidentiality. This dataset's main goal is to aid in the creation of models and algorithms that can identify transactions that might be fraudulent.

B. Brazilian Dataset

A prominent Brazilian bank provided the data, which comprises 374,823 transaction records, 3.74 percent of which are false. Seventeen numerical attributes are included in a record, including the following: merchant category, transaction type, card type (e.g., Visa), transac- tion status (local or worldwide), prior fraud score, time since last transaction, zip codes for both the current and previous transactions and transaction amounts. This dataset, in spite of its age, may provide useful insights on fraud detection and transaction behavior.

C. IEEE-CIS Fraud Detection Database

The IEEE-CIS dataset includes both fraudulent and non-fraudulent transaction records throughout time. It was made publicly available in 2019. Additionally, Vesta Corporation's transaction data is included in the dataset.

The dataset, which includes roughly 590,000 transactions, is divided into two files: one containing identity information and the other containing transaction details. A tiny percentage of the dataset represents fraudulent transactions, making it extremely unbalanced.

3.2 DATA PRE-PROCESSING

- **Addressing Gaps in Data:** Gaps in data are a frequent problem. Imputation can be done statistically through mean, median, or mode calculation. Value could also be removed alongside other rows and columns. More complex imputation methodologies such as using predictive modeling can be applied.

- **Removing Noise and Outliers:** Outliers often distort machine learning algorithms, increasing training time and ultimately degrading overall performance.

- **Normalization and Scaling:** This includes standardization where variables are set to a mean of 0 and variance equal to 1, or scaling between two limits like 0 to 1. Methods like neural network gradient descent optimization or k-nearest neighbors (KNN) rely heavily on data being normally distributed and thus require this step.

- **Data Splitting:** Separating the entire dataset into testing, validation, and training sets creates more order. It is beneficial to train the model on a piece of reserved data and evaluate it on the rest to determine the model's generalization capabilities.

3.3 Imbalance Dataset

Banks are typically reluctant to make credit card fraud datasets publicly available, making them difficult to obtain. The publicly available Credit Card dataset, which comprises 284,807 credit card transactions made by European cardholders over two days in September 2013, was used in our investigation. Only 492 of these transactions, or 0.172 percent of the total, were fraudulent. The dataset is made up of numerical characteristics called V1 through V28 that were acquired by applying Principal Component Analysis (PCA) on the original attributes due to confidentiality restrictions imposed by the data supplier.Other features include the transaction amount, the transaction time in seconds, and the target variable, Class, which shows if a transaction is legitimate (0) or fraudulent (1). To prepare the dataset for analysis, we removed duplicates and normalized all features to fall within the range [0, 1]. The processed dataset included 446 fraudulent transactions out of a total of 283,726 records. It was then divided into a training set, which comprised two-thirds of the data, and a testing set, which included the remaining third. The training set contained 315 fraudulent transactions out of 170,236 records, reflecting an incidence rate of 0.185%. The testing set included the remaining 131 fraudulent transactions out of 113,490 records, with an incidence rate of 0.115%.

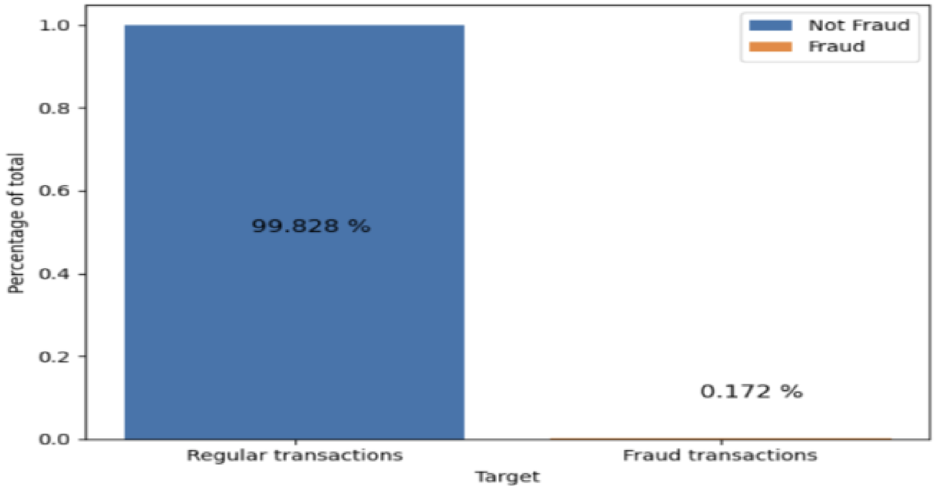

Fig. 1. Imbalanced Data

3.4 GAN Oversampling Model

This study examines a number of oversampling methods for resolving class imbalance issues using a systematic approach. Doing data cleaning, dataset partitioning, oversampling using different methods, constructing the ML algorithm, and assessing performance are notable components of the process. As a measure of enhancing data quality, duplicate records were removed in the first stage of the dataset's construction. This step resulted in the dataset's records being reduced from 284,807 to 275,663. Following the cleaning step, the data was partitioned into training and testing datasets, with 80% allocated to training and 20% reserved for testing, which provided a good boundary for model evaluation.

To balance the classes, SMOTE and GAN approaches in fig [2] were used to oversample the training data. SMOTE generates synthetic samples by efficiently enhancing the dataset by interpolating data points from minority classes. In contrast, to produce high-quality synthetic samples, GAN uses two neural networks—a discriminator and a generator—that compete with one another. The generator creates new samples, while the discriminator evaluates their production or legitimacy. The generator may provide data that is realistic and closely reflects the distribution of minority groups because of this iterative approach. After the oversampling ,we implement various method DT, XGBoost, RF, KNN, MLP, SVM, and LR were the seven machine learning models that were subsequently put into practice. Finally, metrics including accuracy, recall, and F1-score were used to assess how well the model handled class imbalance. When compared to more conventional methods like SMOTE, this approach demonstrates the promise of GAN-based oversampling as a reliable way to improve the performance of classification models.

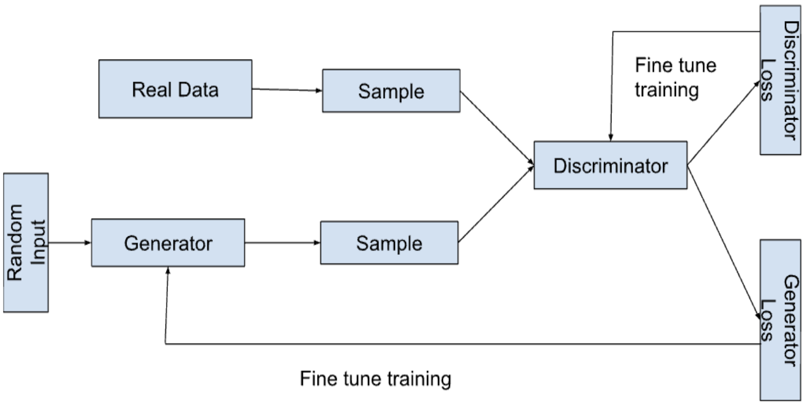

Fig-2 GAN ARCHITECTURE

GAN are neural networks often employed for unsupervised learning applications. Two main components make up they are the generator and the discriminator. The generator generates synthetic data where the discriminator evaluates and differentiates the created data from actual data using adversarial training. This dynamic interaction helps the generator produce synthetic outputs that very nearly reflect real data.

The first step in using GAN to solve data imbalance is data preparation. This phase separates minority class samples to serve as the GAN target distribution. Following that, a generator is taught to generate synthetic data samples by transforming random noise inputs into outputs that very nearly reflect the features of the minority class. At the same time, a discriminator becomes able to differentiate between the synthetic data of the generator and the real minority class samples, therefore assigning each a probability score. Through adversarial training, the generator and discriminator both get better iteratively. The discriminator becomes stronger at distinguishing actual data from fake, while the generator gets better at producing realistic samples that can trick the discriminator. Upon completion of enough training, the generator generates superior synthetic samples that closely mimic the actual minority class data. These artificial samples are added to the original dataset to improve its usefulness for training machine learning models by balancing the distribution of classes.

3.5 SMOTE

One way to tackle class imbalance problems in datasets is by using the Synthetic Minority Over-Sampling Technique (SMOTE). As one of the solutions to the problem, SMOTE aims at achieving a balanced class distribution by creating samples in the minority classes. It generates new instances within the region defined by the feature space of the minority class.

Working Procedure of SMOTE

Find Instances of Minority Classes: SMOTE works with datasets that include one or more classes that are substantially underrepresented in relation to the rest. Finding the minority class or classes in the dataset is the initial step.

Nearest Neighbor Selection: SMOTE finds the k closest neighbors of each minority class instance in the feature space. The user-specified parameter is the number of closest neighbors, represented by the letter k.

Synthetic Sample Generation: SMOTE chooses one of its k closest neighbors at random for every occurrence of a minority class. Then, along the line segment between the minority class instance and the chosen closest neighbor in the feature space, it creates synthetic samples.

Controlled Oversampling: A parameter known as the oversampling ratio, which indicates the ideal proportion of synthetic samples to actual minority class samples, regulates the degree of oversampling. By creating synthetic samples until the minority class equals the size of the majority class, SMOTE normally attempts to balance the class distribution by default.

Repeat for All Minority Class Instances: To create synthetic samples to supplement the

minority class, steps 2-4 are carried out again for every minority class case in the dataset.
Produce a Balanced Dataset: The dataset that is produced after creating synthetic examples for the minority class is more balanced, with a more equal distribution of occurrences across classes.

3.6 Classification Model

Logistic Regression (LR)

Binary classification problems are the main application for the supervised learning method known as logistic regression. It operates by calculating the likelihood that a certain input point falls into a specified category. This is accomplished by applying a linear combination of input characteristics to the sigmoid (logistic) function, which yields an output between 0 and 1 that may be understood as a probability. The instance is allocated to one class if this probability is higher than a threshold, usually 0.5; if not, it is assigned to the other class. Despite its simplicity, Logistic Regression is a good baseline model for many machine learning tasks and performs well on datasets that are linearly separable.

Support Vector Machine (SVM)

Finding the ideal hyperplane that divides classes with the greatest margin is the goal of the potent classification method known as support vector machines. Stated differently, support vector machines (SVM) aim to optimize the distance between the closest data points of various classes. SVM converts the input space into a higher-dimensional space where a linear separation is feasible for non-linearly separable data using kernel functions (such as polynomial, RBF, or sigmoid). SVMs are renowned for their resilience to overfitting, particularly when there are more features than samples, and they perform especially well in high-dimensional domains.

k-Nearest Neighbors (KNN)

A straightforward, non-parametric classification technique called k-Nearest Neighbors bases its predictions on the majority class of the 'k' nearest training samples in the feature space. It saves the whole training dataset instead of requiring an explicit training step. KNN chooses the most frequent label among the closest neighbors after calculating the distance (often Euclidean) between each new data point and every training point. KNN may be computationally costly during prediction and is sensitive to the choice of 'k' and the distance metric used, while being simple to construct and understand.

Decision Tree

A decision tree is a tree structure that resembles a flowchart, with internal nodes standing in for feature choices, branches for the decisions' results, and leaf nodes for class labels. In order to optimize the homogeneity of the resultant subgroups, the dataset is recursively separated according to parameters such as information gain or Gini impurity. Decision trees are helpful for exploratory data analysis because they are simple to understand and comprehend. They may, however, readily overfit the training set, particularly when deep trees are built. Pruning methods are often used to overcome this constraint.

Random Forest

An ensemble learning technique called Random Forest constructs many decision trees and aggregates their results to provide a final prediction, often by majority vote. To provide variety among the trees and lessen overfitting, each tree is trained on a random fraction of the data and employs a random subset of features at each split. High accuracy, resilience to noise and outliers, and capacity to manage huge datasets with high complexity are all attributes of Random Forest models. When class weights or resampling approaches are used, they are also comparatively impervious to the issue of unbalanced datasets.

Extreme Gradient Boosting - XGBoost

An effective and scalable solution in gradient boosting, XGBoost builds an ensemble of decision trees in a sequential manner. While it is possible to further improve the model performance, Each new tree tries to correct the shortcomings of the previous ensemble. To avoid overfitting, XGBoost incorporates a number of regularization techniques (L1 and L2). Furthermore, it supports the handling of missing values, tree pruning and parallel computing. Due to its remarkable performance and flexibility, XGBoost has earned popularity in data science challenges and real-world scenarios.

MLP- Multilayer Perceptron

A kind of artificial neural network is a Multilayer Perceptron, which contains an input layer, one or multiple hidden layers and an output layer. In a Multilayer Perceptron every neuron in one layer is connected to every neuron in the succeeding layer termed as fully connected. During a training session, these connections are assigned weights which are modified through backpropagation. The incorporation of non-linear activation functions such as sigmoid or ReLU, allows the MLP to master complex non-linear patterns. MLPs can process single or multiclass classification problems and they have proven to perform well when used alongside generalization-enhancing techniques, such as batch normalization and dropout.

3.7 PERFORMANCE PARAMETERS

The performance parameters used to evaluate sentiment analysis systems can vary depending on the specific task and application. However, some common metrics is represented in Table3:

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

Table 2 Confusion Matrix

The number of times the model accurately predicts the positive class is known as the True Positive (TP).

The number of times the model accurately predicts the negative class is known as the True Negative (TN).

The number of times the model predicts the positive class wrongly is known as the False Positive (FP).

The number of times the model predicts the negative class wrongly is known as the False Negative (FN).

KPI	Formula Used
Accuracy	(TP+TN)/(TP+TN+FN+FP)
Precision	(TP)/(TP+FP)
Recall	(TP)/(TP+FN)
F1score	(2*Precision*Recall)/(Precision+Recall)

Table-3 Performance Parameter

- **Accuracy** is the most commonly used evaluation metric in classification tasks. It measures the overall correctness of the model by calculating the ratio of correctly predicted instances (both positives and negatives) to the total number of instances. While accuracy provides a general idea of performance, it can be misleading in cases where the data is imbalanced (i.e., one class is significantly more frequent than others).
- **Precision** focuses on the quality of positive predictions. It is the ratio of correctly predicted positive observations to the total predicted positives. Precision is particularly important in scenarios where **false positives** are costly, such as in spam detection, where we want to minimize the number of non-spam emails incorrectly marked as spam.
- **Recall,** or sensitivity/true positive rate, measures how well a model identifies all relevant cases. It calculates the proportion of true positives against all actual positive cases. Recall becomes crucial when false negatives can be more damaging, such as

in medical checks, where not detecting an existing ailment may lead to severe repercussions.

- **F1-score** refers to the weighted average of precision and recall. It attempts to balance both concerns by providing one value, particularly useful when there is a compromise between precision and recall. F1-score proves to be efficient with imbalanced datasets since it considers the entire performance of the model beyond mere accuracy.

CHAPTER 4

RESULT ANALYSIS

4.1 Evaluation Measures:

The models were evaluated on F1-score, Precision, Recall, and three other relevant criteria. It will be evident from the computations that follow, these criteria were chosen in such a way that they have considered the problem from every angle. Recall is determined, by true positive (TP) in relation to the total positive instances available in the dataset. Equation (1) defines recall(R). In equation (2) Precision(P) is used to calculate the proportion of true positive (TP) to all predicated positive cases. The F1 score is the average of the two extremes of precision and recall, and is given by equation (3). These equations employ the terms True Positives (TP), False Negatives (FN), and False Positives (FP). Intentionally omitted from the criteria was Accuracy, due to the dataset imbalance, in order to steer clear from misleading outcomes.

$$R = TP/(TP + FN) \quad (1)$$
$$P = TP/(TP + FP) \quad (2)$$
$$F1 = (2 * R * P)/(R + P) \quad (3)$$

4.2 Result and Discussion:

In Table [3] we specify the each method parameters used in a decision making of test data where we see that logistic regression need no parameter which is very simple model and remaining other model has the parameter like number of iteration, depth, learning rate etc.

Method	Parameters
Logistic Regression	
Suport Vector Machine	C=1.0, random_state=42, max_iter=1000
k-nearest-neighbors(KNN)	n_neighbors=3
Decision Tree	criterion='entropy', random_state=42, max_depth=50
Random Forest	max_depth=50, random_state=42, n_estimators=100

Method	Parameters
XGBoost	max_depth=50, learning_rate=0.1, random_state=42, n_estimators=100
Multilayer Perceptron (MLP)	Adam (learning_rate=0.001), metrics=['accuracy'], loss='binary_crossentropy'

Table-3 Learning Parameter

In Table [4], after balancing the data using various oversampling techniques, the F1 scores from different models were analyzed. The results indicate that models with GAN-based oversampling consistently outperformed those with SMOTE. This observation demonstrates that GAN sampling is more effective than SMOTE across all parameters.

Technique	SMOTE(%)	GAN(%)
Logistic Regression	11	71
k-Nearest Neighbour	64	83
Suport Vector Machine	14	75
Decision Tree	57	73
Random Forest	81	83
XGBoost	79	82
Multilayer Perceptron	76	78

Table-4 Comparison of F1 Scores for Different Techniques.

Table[5] reveals that models using GAN-based sampling consistently outperform those with SMOTE in terms of precision. Notably, Random Forest achieved the highest precision score of approximately 96% with GAN sampling.

Technique	SMOTE(%)	GAN(%)
Logistic Regression	6	76
k-Nearest Neighbour	53	91
Suport Vector Machine	8	85
Decision Tree	45	73
Random Forest	84	96
XGBoost	75	94
Multilayer Perceptron	76	88

Table-5 Comparison of Precision for Different Techniques.

In Table[6], the analysis of recall values shows a smaller gap between the two oversampling techniques. Most models exhibit higher recall values with SMOTE compared to GAN sampling, except for Random Forest, where both techniques achieve an equal recall of approximately 74%.

Technique	SMOTE(%)	GAN(%)
Logistic Regression	6	76
k-Nearest Neighbour	53	91
Suport Vector Machine	8	85
Decision Tree	45	73
Random Forest	84	96
XGBoost	75	94
Multilayer Perceptron	76	88

Table-6 Comparison of Recall for Different Techniques.

The bar graph demonstrates that GAN-based oversampling consistently outperforms SMOTE across all models, indicating that GAN can be a superior alternative to SMOTE for oversampling, especially with large datasets. Among the evaluated techniques, K-Nearest Neighbors and Random Forest achieved the highest F1 score of 83\%. On the other hand, Logistic Regression with SMOTE resulted in the lowest F1 score (11\%), highlighting SMOTE's limitations when paired with Logistic Regression.

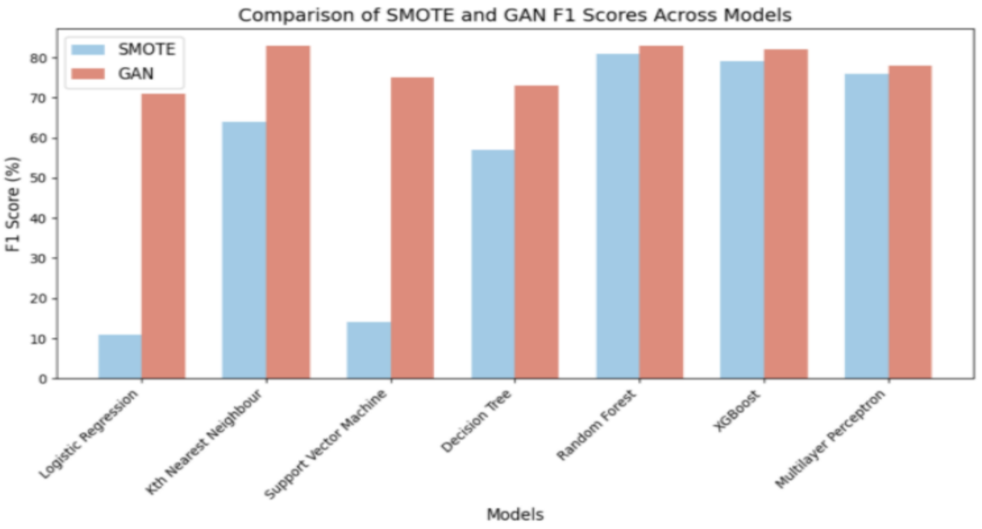

Fig-3 Comparison between the SMOTE and GAN

CHAPTER 5

CHALLENGES

Machine learning–driven credit card fraud detection confronts a range of technical and practical hurdles that stem from the nature of financial data and the evolving tactics of fraudsters. These challenges span data scarcity, evaluation difficulties, real-time constraints, and the shifting behavior of legitimate and fraudulent transactions.

Challenges in Machine Learning–Based Credit Card Fraud Detection

- **High transaction volume**
  Financial institutions process millions of transactions daily, creating massive datasets that pose storage, processing, and real-time analysis challenges3.
- **Uniqueness and ingenuity of frauds**
  Fraudulent schemes continually evolve in novel ways-through skimming, phishing, social engineering, or synthetic identity fraud-making it hard for static models to generalize.
- **Extreme class imbalance**
  Fraudulent transactions typically account for less than 0.5% of all transactions, biasing models toward the majority class and undermining their ability to detect rare frauds.
- **Concept drift**
  The statistical properties of transaction data and fraud patterns change over time, requiring models to adapt continuously to maintain detection accuracy4.
- **Verification latency**
  Fraud detection systems must flag suspicious transactions in milliseconds to prevent losses without delaying customer experience, imposing strict performance requirements.
- **Lack of public benchmarks and standard metrics**
  The absence of widely accepted datasets and evaluation frameworks makes it difficult to compare models and reproduce results across studies3.
- **Data confidentiality and limited sharing**
  Privacy concerns and regulatory restrictions prevent researchers from accessing real transaction data, forcing reliance on anonymized or synthetic datasets that may not reflect real-world complexity.
- **Human annotation errors**
  Labeling fraud is error-prone-fraudulent cases may be misclassified as legitimate and vice versa-compounding model training and evaluation errors.

CHAPTER 6

CONCLUSION AND FUTURE WORK

CONCLUSION

Credit card fraud detection remains a critical challenge due to the rarity and evolving nature of fraudulent transactions. This study demonstrates that advanced oversampling techniques, particularly GAN-based methods, significantly enhance the performance of various classifiers compared to traditional approaches like SMOTE. GANs improve precision and F1-scores across all tested models, making them a promising solution for addressing extreme class imbalance in real-world financial datasets.

Despite these advances, challenges such as concept drift, real-time detection requirements, and data confidentiality persist. No single model or technique guarantees perfect detection, and ongoing adaptation is necessary to keep pace with sophisticated fraud tactics.

FUTURE WORK

To further improve fraud detection systems, future research should explore:

- Real-time Adaptive Models: Developing models that can learn and adapt to new fraud patterns on-the-fly, possibly using online or incremental learning techniques.
- Hybrid Approaches: Combining multiple oversampling methods (e.g., SMOTE, GANs) and ensemble classifiers to leverage their complementary strengths.
- Explainability and Transparency: Integrating explainable AI to help analysts understand and trust model decisions, which is crucial in financial applications.
- Robustness to Concept Drift: Implementing mechanisms to detect and adapt to changes in transaction patterns and fraud strategies over time.
- Data Privacy and Collaboration: Creating privacy-preserving frameworks that allow institutions to share insights or synthetic data without compromising sensitive information.
- Benchmark Datasets: Establishing standardized, realistic benchmark datasets and evaluation protocols to facilitate reproducible research and fair comparison of methods.

By addressing these directions, future work can help build more accurate, robust, and trustworthy fraud detection systems that better protect both consumers and financial institutions.

LIST OF PUBLICATIONS AND THEIR PROOFS

[1] Mayank Pathak, and R. Beniwal, “A Systematic Literature Survey on Credit Card Fraud Detection Using Machine Learning Techniques”,in proceedings of the International Conference on Pervasive Computing and Social Networking(ICPCSN) 2025,Mar. 3,2025 .[Accepted]

[2] Mayank Pathak, and R. Beniwal, “Implementation On Credit Card Fraud Detection Using GAN Oversampling”,in proceedings of the International Conference on Pervasive Computing and Social Networking (ICPCSN) 2025,Mar 3,2025.[Accepted]

REFERENCE

[1] P. Y. Prasad *et al.*, “A comparison study of fraud detection in usage of credit cards using machine learning,” in *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 1204- 1209, IEEE, Apr. 2023.

[2] R. Qaddoura and M. M. Biltawi, “Improving fraud detection in an imbalanced class distribution using different oversampling techniques,” in *2022 International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*, pp. 1-5, IEEE, Nov. 2022.

[3] K. G. Al-Hashedi and P. Magalingam, “Financial fraud detection ap- plying data mining techniques: A comprehensive review from 2009 to 2019,” *Computer Science Review*, vol. 40, p. 100402, 2021.

[4] E. Esenogho *et al.*, “A neural network ensemble with feature engineering for improved credit card fraud detection,” *IEEE Access*, vol. 10, pp. 16400-16407, 2022.

[5] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, “Credit card fraud detection-machine learning methods,” in *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pp. 1-5, IEEE, Mar. 2019.

[6] F. K. Alarfaj *et al.*, “Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms,” *IEEE Access*, vol. 10, pp. 39700-39715, 2022.

[7] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, “Enhanced credit card fraud detection based on attention mechanism and LSTM deep model,” *Journal of Big Data*, vol. 8, p. 1-21, 2021.

[8] S. P. Maniraj, A. Saini, S. Ahmed, and S. Sarkar, ”Credit card fraud detection using machine learning and data science,” *Int. J. Eng. Res.*, vol. 8, no. 9, pp. 110-115, 2019.

[9] A. Mishra and C. Ghorpade, ”Credit card fraud detection on the skewed data using various classification and ensemble techniques,” in *Proc. IEEE Int. Students’ Conf. Electr., Electron. Comput. Sci. (SCEECS)*, Feb. 2018, pp. 1-5.

[10] S. Bakhtiari, Z. Nasiri, and J. Vahidi, “Credit card fraud detection using ensemble data mining methods,” *Multimedia Tools and Applications*, vol. 82, no. 19, pp. 29057-29075, 2023.

[11]S. Carta, G. Fenu, D. R. Recupero, and R. Saia, “Fraud detection for E-commerce transactions by employing a prudential Multiple Consensus model,” *Journal of Information Security and Applications*, vol. 46, pp. 13-22, 2019.

[12] E. A. L. M. Btoush *et al.*, “A systematic review of literature on credit card cyber fraud detection using machine and deep learning,” *PeerJ Computer Science*, vol. 9, p. e1278, 2023.

[13] P. Gupta, A. Varshney, M. R. Khan, R. Ahmed, M. Shuaib, and S. Alam, “Unbalanced credit card fraud detection data: a machine learning- oriented comparative study of balancing techniques,” *Procedia Computer Science*, vol. 218, pp. 2575-2584, 2023.

[14] I. A. Mondal, M. E. Haque, A. M. Hassan, and S. Shatabda, “Handling imbalanced data for credit card fraud detection,” in *2021 24th Interna- tional Conference on Computer and Information Technology (ICCIT)*, pp. 1-6, IEEE, Dec. 2021.

[15] H.Ahmad,B.Kasasbeh,B.Aldabaybah,andE.Rawashdeh,“Classbal- ancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS),” *International Journal of Information Technology*, vol. 15, no. 1, pp. 325-333, 2023.

[16] A.R.Khalid*etal.*,“Enhancingcreditcardfrauddetection:anensemble machine learning approach,” *Big Data and Cognitive Computing*, vol. 8, no. 1, p. 6, 2024.

[17] N.F.Ryman-Tubb,P.Krause,andW.Garn,“HowArtificialIntelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark,” *Engineering Applications of Artificial Intelligence*, vol. 76, pp. 130-157, 2018.

[18] M. Alamri and M. Ykhlef, “Survey of credit card anomaly and fraud detection using sampling techniques,” *Electronics*, vol. 11, no. 23, p. 4003, Dec. 2022.

[19] J. Forough and S. Momtazi, “Ensemble of deep sequential models for credit card fraud detection,” *Applied Soft Computing*, vol. 99, p. 106883, 2021.

[20] H.Najadat,O.Altiti,A.A.Aqouleh,andM.Younes,”Creditcardfraud detection based on machine and deep learning,” in *Proc. 11th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2020, pp. 204-208.

[21] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. S. Hacid, and H. Zeined- dine, ”An experimental study with imbalanced classification approaches for credit card fraud detection,” *IEEE Access*, vol. 7, pp. 93010-93022, 2019.

[22] N.S.AlfaizandS.M.Fati,”Enhancedcreditcardfrauddetectionmodel using machine learning,” *Electronics*, vol. 11, no. 4, p. 662, 2022.

[23] K. A. K. Saputra, M. Mu’ah, J. Jurana, C. W. M. Korompis, and D.T. Manurung, ”Fraud Prevention Determinants: A Balinese Cultural Overview,” *Australas. Account. Bus. Finance J.*, vol. 16, no. 3, pp. 167- 181, 2022.

[24] S. Rajora *et al.*, ”A comparative study of machine learning techniques for credit cardfraud detection based on time variance,” in *Proc. IEEE Symp. Comput. Intell. (SSCI)*, Nov. 2018, pp. 1958-1963.

[25] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, ”Fraud detection in banking data by machine learning techniques,” *IEEE Access*, vol. 11, pp. 3034-3043, 2022.