College of Computer and Information Sciences
Computer Science Department

Enhance Credit Card Fraud Detection Using Generative AI

CSC 496– Final Report

Prepared by:

Fai Alharthi	442202288
Shoug Alsaleem	443200641
Lina Alsuwaylimi	443200998
Raghad Aldosari	443201004
Aliyah Aljarallah	443201214

Supervised by:

Dr.Amani Alajlan

Research project for the degree of Bachelor in Computer Science
Second Semester 1446

I.English Abstract

The increasing use of credit cards and digital transactions has made credit card fraud a major issue, costing the global economy billions of dollars annually. Detecting fraudulent transactions is difficult because there are fewer cases of fraud than normal transactions. This imbalance makes it challenging for traditional models to recognize fraud patterns. To solve this issue, our project uses VAE-GAN hybrid model to generate synthetic fraud data, balancing the dataset and improving detection accuracy. Additionally, we explore several anomaly detection techniques, a supervised learning algorithm, such as Random Forest, and an unsupervised learning algorithm, such as One-Class Support Vector Machine (OC-SVM), to identify unusual transactions. We will evaluate the performance of our model using accuracy, precision, recall, and F1-score. This project aims to improve fraud detection systems, reduce financial losses, and demonstrate how generative AI can help address real-world problems with unbalanced data.

II.Arabic Abstract

أسهم الانتشار الواسع لاستخدام بطاقات الائتمان والمعاملات الرقمية في تفاقم مشكلة الاحتيال في بطاقات الائتمان مما جعلها تحديا رئيسيا، يكلف الاقتصاد العالمي مليارات الدولارات سنويًا. يعد اكتشاف المعاملات الاحتيالية أمرًا صعبًا للغاية وذلك بسبب الفارق الكبير بين عدد معاملات الاحتيال وعدد المعاملات النظامية. هذا الاختلال يصعّب على النماذج التقليدية التعرف على أنماط الاحتيال.

لحل هذه المشكلة، يستخدم مشروعنا الشبكات التوليدية التنافسية (GANs) و المرمز التلقائي التباييني (VAE) معا لتوليد بيانات احتيالية اصطناعية، مما يساعد على تحقيق توازن بين أعداد المعاملات النظامية والاحتيال وتحسين دقة الاكتشاف. بالإضافة إلى ذلك، سنستكشف تقنيات مختلفة لاكتشاف القيم المتطرفة، بما في ذلك خوارزمية تعلم تحت إشراف مثل غابة العشوائية(Random Forest) ، وخوارزمية تعلم غير خاضعة للإشراف مثل آلة متجه الدعم أحادية التصنيف(OC-SVM) ، بما في ذلك التعلم الآلي التقليدي والتعلم العميق، لتحديد المعاملات غير النظامية. لمعرفة الطريقة الأكثر كفاءة، سيتم المقارنة بين نماذج اكتشاف الاحتيال التقليدية والنماذج القائمة على الذكاء الاصطناعي. ومن ثمّ تقييم أدائها باستخدام معايير الدقة، والتحديد (Precision)، والاسترجاع (Recall)، ومعامل F1. يهدف هذا المشروع إلى تحسين أنظمة اكتشاف الاحتيال، وتقليل الخسائر المالية، وإبراز كيف يمكن للذكاء الاصطناعي التوليدي المساعدة في معالجة المشكلات الواقعية التي تتعلق بالبيانات غير المتوازنة.

Table of Contents

List of Tables

List of Figures

List of Notations

GANs	Generative Adversarial Networks
VAE	Variational Auto-Encoder
PCA	Principal Component Analysis
SVM	Support vector machines
DNNs	Deep Neural Networks
RNNs	Recurrent Neural Networks
SMOTE	Synthetic Minority Oversampling Technigue
OC-SVM	One-Class Support Vector Machine
TP	True Positive
FP	False Positive
FN	False Negative
TN	True Negative

Chapter 1:
Introduction

Fraud detection is the process of detecting and preventing unauthorized or suspicious activities that could lead to financial losses. When it comes to credit card transactions, it is all about analyzing spending patterns, identifying unusual behavior, and distinguishing between real and fraudulent transactions. With fraudsters constantly finding new ways to exploit financial systems, banks and businesses rely on advanced fraud detection methods like machine learning and AI to catch suspicious activity in real time and keep customers safe. As credit cards and digital payments become more popular, credit card fraud is also increasing. Credit card fraud happens when someone uses another person's card without permission, leading to unexpected financial losses for both the cardholder and the bank [1]. The problem has grown in recent years, especially during the COVID-19 pandemic when online shopping increased, making people more vulnerable to fraud [2]. One of the biggest challenges in fraud detection is that fraudulent transactions make up only a tiny portion of all transactions. Because of the imbalance nature of data, traditional detection systems sometimes struggle to accurately identify fraud. Previous studies [3] highlighted that addressing this imbalance is crucial for improving detection accuracy.

To tackle this issue, we are exploring how deep learning and generative AI can make fraud detection more effective by improving anomaly detection techniques. Our project will use a publicly available dataset for European cardholder transactions[3]. First, we will clean and normalize the data to ensure consistency. Since fraud cases are rare, we will use resampling techniques to balance the dataset. Generative adversarial networks (GANs) offer an exciting solution. GANs can create synthetic fraud data to improve detection accuracy [4]. Our approach combines traditional fraud detection methods with advanced AI-driven techniques [1] provide a strong foundation for anomaly detection strategies, which we will build on to explore deep learning algorithms and hybrid models. Ultimately, our goal is to develop a fraud detection system that is both accurate and efficient. By leveraging these advanced techniques, we hope to improve financial security, reduce losses, and contribute to a safer digital payment landscape.

1.1 Problem Statement

The use of credit cards and digital purchases has increased in recent years. People use it to shop, pay bills and for online transactions. As the number of credit card users increases, cases of credit card fraud have also risen. The definition of credit card fraud is "When an individual uses another individual's credit card for personal use while the owner of the card as well as the card issuer are not aware of the thing that the card is being used"[5]. According to statistics, credit card fraud costs the global economy billions of dollars annually[6]. Consequently, banks and financial institutions need to improve their fraud detection strategies and increase customer awareness about the related risk.

To address this problem, we will use anomaly detection techniques while also analyzing and comparing deep learning algorithms. Additionally, generative AI techniques offer a promising solution to solve class imbalance by generating synthetic data to build up the minority class samples and enhance feature representations. Our project aims to enhance system reliability, reduce financial losses, and advance the application of generative AI in credit card fraud detection. This solution will help financial institutions to minimize potential losses and improve customer trust.

1.2 Goals and Objectives

The main goal of this project is to build a fraud detection system for credit card transactions using various anomaly detection techniques. We aim to help financial institutions to identify fraudulent transactions and reduce false alarms, making transactions safer.

Objectives

In order to achieve our intended goal, the following objectives are required:

- Review Related Work: read and understand previous research on fraud detection and anomaly detection techniques to define accurately the problem.
- Study and investigate existing anomaly detection techniques, then select the most suitable technique for our problem.

- Select a suitable public data set for credit card fraud detection.
- Data Preprocessing and Normalization: clean and standardize the data to make sure all features are on the same scale.
- Use Generative AI Techniques to generate synthetic fraud data.
- Build and Train Anomaly Detection Models.
- Evaluate the performance of our proposed algorithm and compare it with previous studies in the field.
- Final Reporting.

1.3 Proposed Solution

We propose a methodology that combines anomaly detection techniques, and generative AI models to address the issue of detecting credit card fraud in highly imbalanced datasets. By preprocessing and normalizing the data and using generative AI techniques as data augmentation methods to create synthetic fraud data followed by resampling techniques, to further improve the training process we hope to reduce the imbalance and enhance the model's ability to distinguish between fraudulent and non-fraudulent transactions.

In addition, Anomaly detection techniques will then be used to detect anomalies by learning patterns in the data. The effectiveness of our approach will be validated using metrics like accuracy, recall, precision, and F1-score, and compared with previous studies. This research aims to demonstrate how combining generative AI and traditional methods can significantly enhance anomaly detection in credit card fraud detection.

1.4 Research Scope

In our project we aim to improve the way of detecting credit card fraud using anomaly detection and generative artificial intelligence. By exploring public data from European credit card holder’s transactions [7] .This dataset comprises 284,807 transactions, of which only 492 are fraudulent. The dataset includes 30 features, such as transaction time, amount, and anonymized principal component analysis (PCA) transformed features.

1.5 Research Significance

The significance of this research is rooted in its ability to reduce financial losses and improve security for both financial institutions and their clients. By employing generative AI techniques, the project enhances fraud detection accuracy, improves data balance, and reduces financial losses, leading to more reliable anomaly detection models. The project seeks to enhance the precision of fraud detection, particularly in imbalanced datasets where conventional approaches often fall short. This initiative is crucial for pushing the boundaries of anomaly detection, fostering customer confidence, and reducing risks tied to credit card fraud. The results could make a significant impact on the financial sector by establishing new standards for fraud detection systems.

1.6 Ethical and Social Implications

Ethical issues are critical in fraud detection research because online transactions have become a part of daily life. Fraud detection research typically involves sensitive data in its processes. This research project considers the ethical and social implications of applying AI in financial transactions and fraud detection. We avoid ethical concerns related to data privacy and consent by using a public database.

1.7 Report Organization

This report is organized into six main chapters. Chapter 2 provides an overview of key concepts related to fraud, including the imbalance problem, anomaly detection and generative AI. Chapter 3 reviews existing studies, categorizing them into anomaly detection-based approaches and those that combine anomaly detection with generative AI. Chapter 4 outlines the methodology that we used to address our problem. Chapter 5 describes the experimental setup for improving credit card fraud detection. Finally, Chapter 6 concludes the report by summarizing key findings.

Chapter 2: Background

In this chapter, we introduce credit card fraud detection and its challenges. We explore the imbalance problem and methods to handle it. We also explain different types of anomalies techniques and their role in fraud detection. Finally, we explore Generative AI as a data augmentation technique.

2.1 Credit card fraud

Credit card fraud is a growing issue in the financial industry, resulting in considerable financial losses and endangering the safety of digital transactions. With the rise in electronic transactions, fraud methods are becoming increasingly advanced, highlighting the importance of detecting and preventing fraud. Credit card fraud occurs when a person utilizes a credit card or its details without permission to carry out transactions without the cardholder's knowledge [8]. To tackle this problem, anomaly detection methods and generative AI models are employed to enhance the accuracy of fraud detection. In this research, our goal is to improve credit card fraud detection effectiveness by utilizing generative AI models, including Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs).

2.2 The Imbalance Problem

Class imbalance occurs when there is a huge difference between the number of samples in each class. For example, in a dataset that detects lung cancer from X-ray images, the majority of X-rays (99.99%) show normal lungs, while only a small fraction (0.01%) contains cancerous cells [9], In fraud detection, fraudulent transactions typically make up less than 1% of all transactions [7].

Machine learning, and deep learning models have been successfully applied to many application domains. However, imbalanced class distribution of a dataset leads to a significant challenge to these classifiers, because most algorithms assume a relatively balanced distribution of classes [10].

There are three main reasons why class imbalance makes learning difficult. First, the model often lacks enough examples from the minority class to train or learn from, leading to poor detection or cannot recognize these classes. So, if there are no instances of the rare class in the training data, the model may assume these classes do not exist.

Second, class imbalance can cause the model to rely on simple methods, such as assuming the label for the majority class for all instances, which leads to high accuracy but fails to recognize the fundamental patterns in the data. Lastly, class imbalance leads to unequal error costs, where making a mistake with a rare case can have much worse consequences than with a common case. If the model does not consider this imbalance, it might treat all cases the same, missing the more important and costly errors. These issues are common in various tasks like fraud detection, disease screening, and object detection, where detecting rare cases is often more important, even though they are less frequent [9].

Imbalanced datasets are a significant challenge in fraud detection, as fraudulent transactions represent only a small fraction of total transactions [11]. Traditional machine learning models tend to focus on the majority class, leading to high accuracy but poor detection on minority class instances. To address this issue, various techniques have been developed to balance the dataset and improve model performance in detecting the minority class patterns. The primary methods for handling class imbalance include resampling techniques (modifying the dataset), anomaly detection techniques (adjusting model training strategies).

2.2.1 Resampling techniques

Resampling techniques modify the dataset distribution to ensure fraud cases are better represented during model training. These methods help improve recall and reduce false negatives by making fraudulent transactions more detectable. The three main approaches to resampling are oversampling, undersampling, and hybrid resampling [12]:

- Oversampling increases the number of fraud cases in the dataset by duplicating existing fraud transactions [13], [14]. Oversampling ensures that the machine learning model is exposed to more fraud cases, improving its ability to recognize fraudulent behavior.

Figure 1 below illustrates oversampling in a fraud detection dataset. The left side figure shows the dataset before oversampling, where fraud cases are significantly outnumbered. The right side figure demonstrates the effect of oversampling, where additional synthetic fraud cases have been generated to balance the dataset.

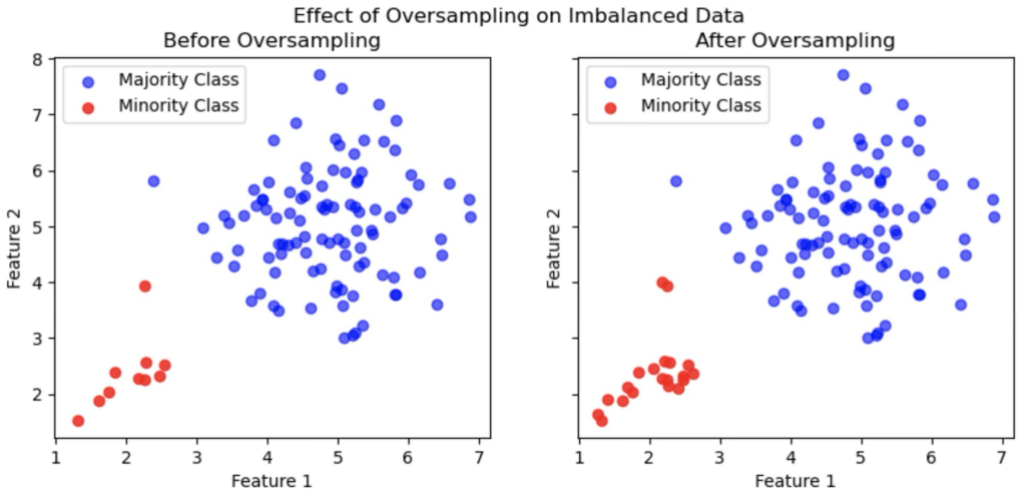
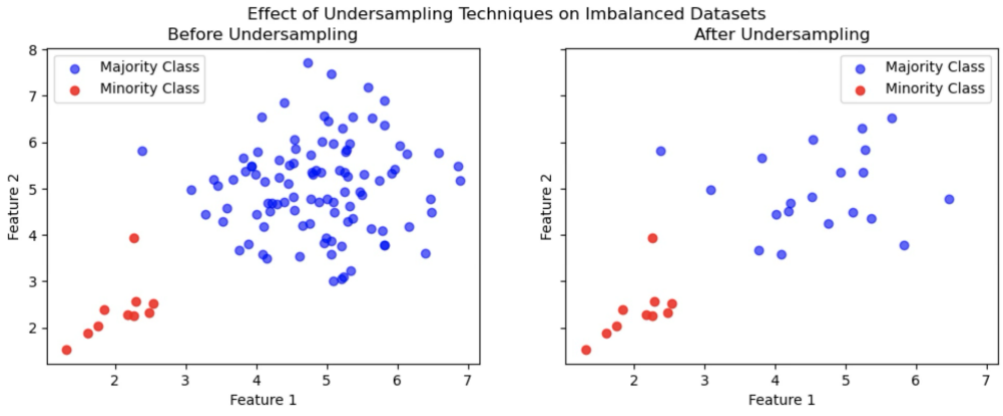


Figure 1 : Effect of OverSampling on Fraud Detection Dataset

- Undersampling reduces the number of Majority Class to create a more balanced dataset. Instead of increasing the number of Minority Class, it removes a portion of the Majority Class to match the count of Minority Class[12], [14]. This technique forces the model to focus more on fraud detection rather than learning patterns dominated by Majority Class.

Figure 2 below illustrates undersampling in fraud detection. The left side figure shows the dataset before undersampling, where Majority Class outnumber the Minority Class. The right side figure demonstrates how undersampling removes a portion of Majority Class to create a more balanced dataset.



- Hybrid approach combines both oversampling and undersampling. This approach applies oversampling to Minority Class to increase their representation while reducing the number of Majority Class through undersampling[14]. The goal is to balance the dataset without excessively increasing its size or losing too much Majority Class data.

Figure 3 below illustrates hybrid resampling in imbalanced dataset. The left side shows the dataset before resampling, where Majority Class dominate. The right side shows the combined effect of oversampling fraud cases and undersampling on Majority Class, achieving a more balanced dataset.

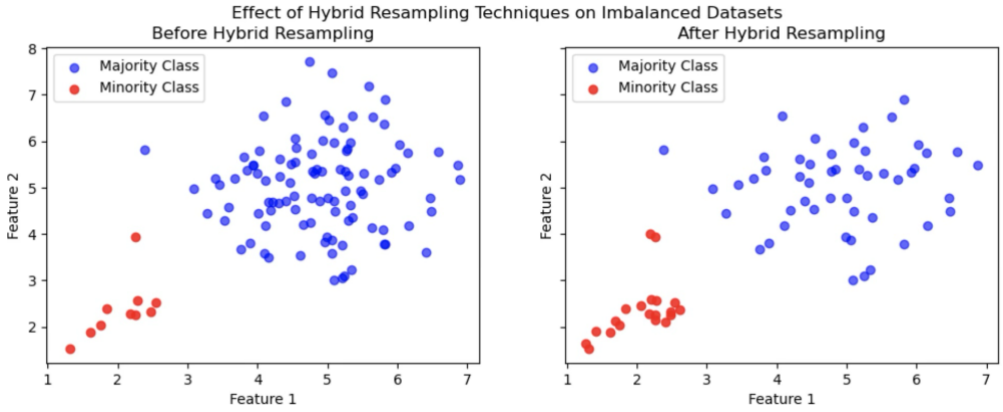


Figure 3 : Effect of Hybrid Resampling on Fraud Detection Dataset

2.2.2 Anomaly detection

Anomaly detection is a powerful approach for detecting fraud when fraudulent transactions are rare and exhibit unusual patterns[1]. Unlike traditional classification models that require balanced datasets, anomaly detection methods assume that fraudulent transactions differ significantly from legitimate ones and can be identified as statistical outliers.

Anomaly detection models analyze transaction features (e.g., transaction amount, frequency, location) and identify deviations from normal patterns (see Section 2.3: Anomaly Detection.). These deviations are flagged as potential fraud cases. Unlike

standard classification models, anomaly detection does not rely on labeled fraud data, making it particularly useful when fraud patterns evolve over time.

2.2.3 Challenges and limitations in handling imbalance

Handling imbalanced datasets presents several challenges that can impact model performance. Over-sampling methods can lead to overfitting [13], where the model memorizes synthetic minority class instances instead of learning general patterns. On the other hand, while undersampling is effective in addressing the imbalance, it has major drawbacks that can lead to information loss if valuable majority class patterns are removed [12]. Additionally, anomaly detection techniques, while useful, often suffer from high false positive rates, as rare but legitimate instances may be incorrectly flagged as anomalies. Selecting the right approach requires balancing improved detection with the risk of overfitting, data loss, and false alarms.

2.3 Anomaly Detection

Anomaly detection is the process of identifying patterns or behaviors that significantly deviate from expected or normal values. These unusual patterns, often called anomalies, outliers, exceptions, or deviations as presented in Figure 4, can indicate important events across various fields[4]. For example, cybersecurity, anomaly detection is crucial for identifying unusual network activities, such as unauthorized access attempts or potential cyberattacks. By monitoring network traffic and user behavior, it enables the early detection of threats, thereby enhancing the security posture of organizations. In manufacturing and industrial processes, anomaly detection plays a key role in ensuring product quality and operational efficiency. It helps detect defects on production lines, identify potential machine failures, and facilitate predictive maintenance. By analyzing sensor data and equipment performance metrics, it aids in preventing costly downtime and maintaining consistent product standards. In healthcare, anomaly detection is used to monitor patient health indicators to detect early signs of medical conditions. It assists in identifying irregularities in vital signs, lab

results, or imaging data, enabling timely interventions and improving patient outcomes [13].

Beyond these fields, anomaly detection has applications in various other domains. In finance, it is used for fraud detection by identifying unusual transaction patterns. In environmental monitoring, it 20helps detect abnormal changes in climate data or pollution levels. General, anomaly detection serves as a critical tool across multiple sectors, enabling the identification of irregular patterns that may signal significant events or conditions. Its applications enhance decision-making processes, improve operational efficiency, and contribute to safety and security in various industries.

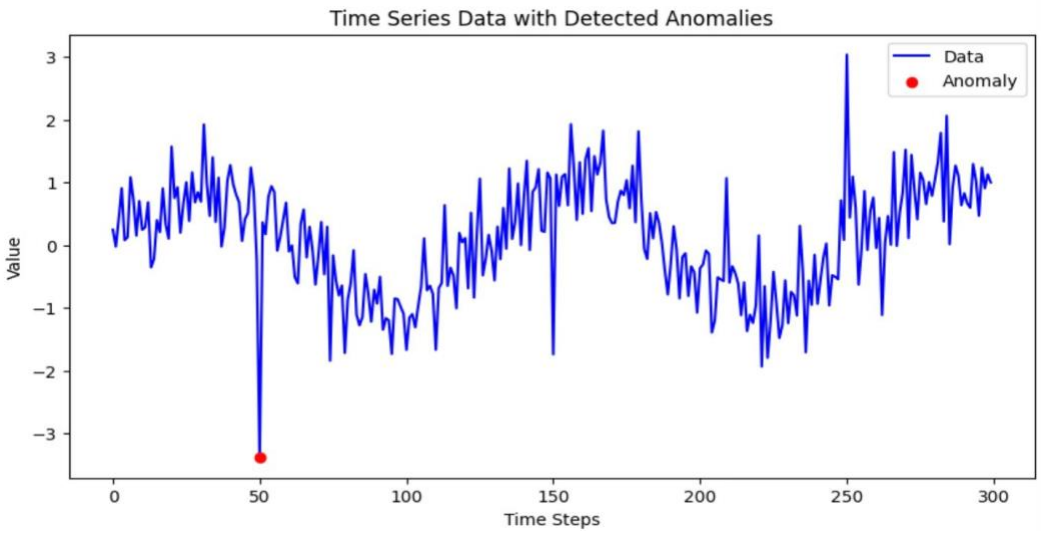


Figure 4 : Anomaly Detection

2.3.1 Types of Anomalies:

1. Point Anomalies: This simplest type of anomaly is defined as an instance of individual data that has been identified as anomalous compared to the rest of the dataset.
2. Contextual Anomalies: A contextual anomaly occurs when a data point is normal overall but unusual within a specific context. It depends on:

   A. Contextual Attributes (Specify the environment, such as the time in stock market data).

B. Behavioral Attributes (Describe the data properties, like infection rates in disease tracking)

3. Collective Anomalies: It is a collection of related data instances that are anomalous compared to the entire data set.

2.3.2 Anomaly Detection Methods:

We will discuss five anomaly detections as presented in Figure 5

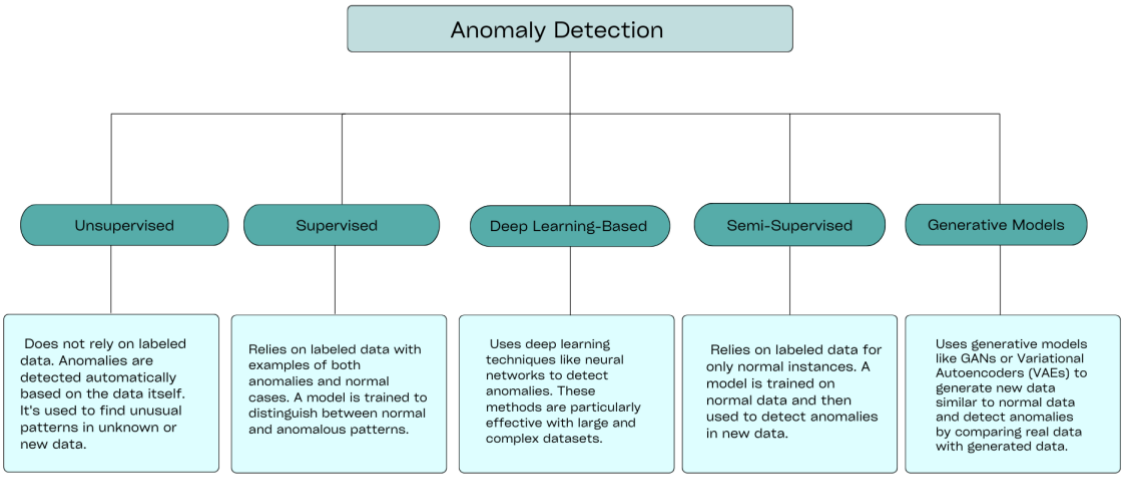


Figure 5: Types Anomaly Detection Techniques

1. **Unsupervised Learning:** This method does not require labeled data, making it useful when anomalies are rare or unknown in advance. Instead of learning from predefined labels, the model analyzes normal patterns in the data and flags instances that significantly deviate[15]. The most common techniques include:
   - Clustering: Groups similar data points together, considering instances that do not fit any cluster as anomalies.
   - Principal Component Analysis (PCA): Reduces data dimensionality while identifying outliers.

- Isolation Forests: Efficiently isolates rare instances by breaking the data into smaller partitions.
- Auto encoders: Train Neural networks to reconstruct normal patterns, where high reconstruction errors indicate anomalies.

2. **Supervised Learning:** Unlike unsupervised methods, supervised learning relies on labeled datasets where each instance is classified as either normal or anomalous. This allows the model to learn specific characteristics associated with anomalies. Common algorithms include:
   - Decision Trees: Rule-based models that identify patterns from labeled data.
   - Logistic Regression: A statistical model used to estimate the probability of an anomaly.
   - Support Vector Machines (SVM): Separates normal and anomalous data points using a hyperplane.
   - Ensemble Methods: Combine multiple models to improve accuracy.

3. **Deep Learning:** Deep learning techniques are highly effective for anomaly detection, particularly in large and complex datasets [17]. Some widely used methods include:
   - Deep Neural Networks (DNNs): Capable of recognizing complex patterns in high-dimensional data.
   - Recurrent Neural Networks (RNNs): Ideal for analyzing sequential data, such as time-series financial transactions or industrial sensor readings.

4. **Semi-Supervised:** To increase detection accuracy, semi-supervised learning blends a significant amount of unlabeled data with a small amount of labeled data. Because labeled fraud cases are few and difficult to obtain, this method is helpful in the detection of fraud. The model can discover new patterns and identify fraud more successfully by learning from both labeled and unlabeled data. To label the unlabeled data, the model first learns from the classified data. This aids in its ability to identify novel forms of fraud. Typical techniques include "co-training," in which several models learn from various aspects to improve accuracy, and "self-training," in which the model classifies the data

itself. These models are frequently used in conjunction with rule-based systems to identify both new and intricate fraud patterns as well as well-known ones [18]. This combination aids in the system's ability to adjust to evolving fraud strategies. According to research, semi-supervised learning enhances the detection of novel fraud types and lowers false positives[18]. This method maintains excellent accuracy while requiring less labeled data, making it even more helpful as transaction data grows.

2.4 Generative AI

Generative AI is a powerful data augmentation technique that enhances the feature representation of the minority class by generating synthetic data [19]. Unlike traditional resampling methods that duplicate or remove existing samples, they create new**,** diverse examples. By introducing realistic synthetic samples, Generative AI improves model performance, reduces overfitting, and strengthens the detection of rare patterns in imbalanced datasets.

2.4.1 Generative Adversarial Networks

Generative Adversarial Networks (GANs) are part of deep learning models designed to generate synthetic data by learning from real examples. It was first introduced by Ian Goodfellow et al [20], GANs have been widely used in so many fields, including image synthesis, text generation, medical data augmentation, and anomaly detection. They are a powerful tool for applications where data is limited, imbalanced, or difficult to obtain.

GANs are composed of two neural networks that engage in a competitive process. as presented in Figure 6, the Generator and the Discriminator. The role of the Generator is to produce synthetic samples that closely mimic real data. It initiates this process by using random noise as input and incrementally enhances its capability to generate output that appears realistic. On the contrary, the Discriminator functions as a classifier, tasked with differentiating between real data (Training set) and the synthetic data produced by the Generator. This continuous improvement in both networks; the Generator creates realistic samples, while the Discriminator identifies synthetic data [21].

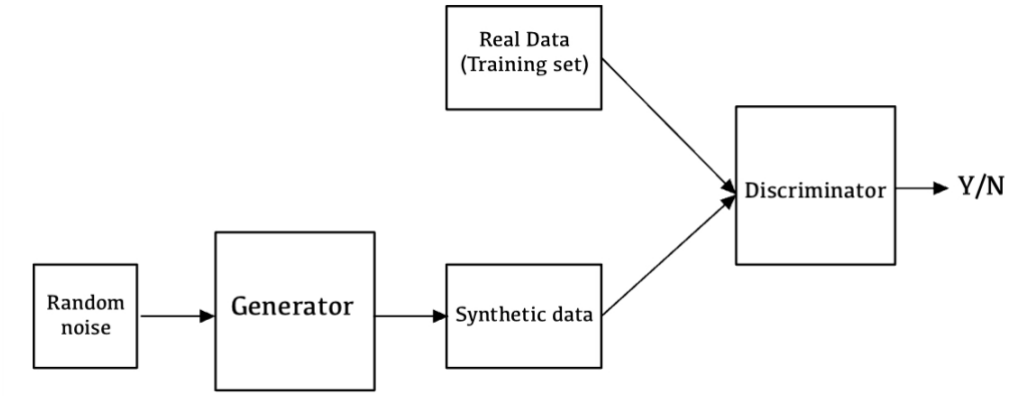


Figure 6: Generative Adversarial Networks(GANs)

2.4.2 Variational Auto-encoder

Autoencoders are neural networks that take the input data and reconstruct it or generate new data by using encoders, which are the part that compresses the input data into smaller forms called latent vectors, and decoding, which tries to reconstruct data based on a similar representation to the original data.

As our goal is to generate new realistic data, the autoencoder could fall into overfitting because it maps each input to a single latent vector.
So, the number of latent vectors that are generated exactly matches the number of input data points, and that leads to too much similarity between the generated data and the data in the dataset.

A variational autoencoder (VAE), which is considered an extension of the autoencoder, consists of the same phases encoders that segment the input into smaller forms. But for the decoder instead of generating a single latent vector for each input, the decoder generates the mean and covariance that defines a chosen distribution of latent vectors, so the generated data will not be exactly the same but will be new while at the same time remaining similar to the typical pattern of the data in the dataset[22].

2.4.3 Diffusion Model

Diffusion Models are a type of generative model used to create synthetic data by progressively adding noise to a dataset and then reversing this process to reconstruct realistic samples. This approach is particularly useful in fraud detection, as it helps address the issue of imbalanced datasets by generating synthetic fraudulent transactions that closely mimic real fraud patterns[23].

These models operate in two main phases:

Forward Diffusion Process: Where Gaussian noise is gradually added to the data, transforming it into a standard normal distribution.

Reverse Diffusion Process: Where the model learns to remove the noise step by step, reconstructing data samples similar to the original dataset [24].

Compared to Generative Adversarial Networks (GANs), Diffusion Models offer better training stability and lower risk of mode collapse, making them a promising alternative for fraud detection systems. By incorporating synthetic fraudulent data, fraud detection models can improve accuracy in identifying fraudulent transactions, leading to higher precision and recall rates [23].

2.5 Summary

This chapter has provided an overview of credit card fraud detection, highlighting challenges such as class imbalance and methods to handle it. We explored resampling techniques like oversampling, undersampling, and hybrid methods, along with anomaly detection methods such as supervised learning, unsupervised learning, semi-supervised

learning, and deep learning. Additionally, we introduced Generative AI as a data augmentation technique, focusing on models like GANs, VAEs, and Diffusion Models. These methods improve fraud detection by balancing datasets and strengthening machine learning model performance.

Chapter 3: Related Work

The following sections categorize existing research based on their approach to fraud detection: anomaly detection-based approach and anomaly detection with generative AI-based approach. Additionally, the studies will also be summarized in tables, highlighting key aspects such as the type of model used, the method for handling class imbalance, datasets, and evaluation metrics (accuracy, precision, recall, and F1-score).

3.1 Anomaly Detection-based approach research

In this section, we explore studies that focus on anomaly detection techniques for fraud detection, highlighting different models used and their effectiveness in identifying fraudulent transactions.

Ebrahim et al.[25] addressed the class imbalance problem in credit card fraud detection using Fourier Transform-based spectral analysis to analyze transaction frequency patterns. They used a European dataset containing 284,807 transactions, with 492 identified as fraudulent. Their approach improved anomaly detection and reduced false positives. Several models were compared, including Logistic Regression, Decision Tree, and K-Nearest Neighbors (KNN). Support Vector Machine (SVM) classification model outperformed them, achieving 92% precision, 89% recall, 90% F1-score, 90% accuracy.

Mienye and Jere [26] studied deep learning techniques for credit card fraud detection. Their approach was applied to the European Credit Card Dataset [27]. Their study evaluated various classification algorithms, including CNNs, RNNs, LSTMs, gated recurrent unit (GRU), and Transformer-based models like BERT, finding that GRU achieved the best performance across the various metrics, achieving 0.99 accuracy, 0.79 precision, 0.99 specificity, 0.94 recall, and 0.82 F1-score. Meanwhile, Simple RNN and CNN seem to have the lowest performance compared to the other models.

Ali et al. [28] the European Credit Card Fraud Dataset [27], preprocessing steps such as outlier detection and removal were applied to the oversampled dataset. To handle the class imbalance, the researchers employed oversampling techniques

combined with outlier detection. Also, four machine learning algorithms were implemented: K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM), and Logistic Regression (LR). Among these, the Random Forest (RF) model achieved the best performance, with an accuracy of 99.96%, precision of 0.96, recall of 0.80, and an F1-score of 0.88, demonstrating its effectiveness in accurately detecting fraudulent transactions while minimizing false positives and false negatives.

Abdul Salam et al. [8] propose a federated learning model for credit card fraud detection using data balancing techniques such as SMOTE, AdaSyn, and RUS on Kaggle's dataset. Random Forest achieved the highest performance after resampling, with 99.99% accuracy, 99.98% precision, 100% recall, and an F1-score of 99.99%. The federated learning approach using PyTorch outperformed TensorFlow Federated in terms of accuracy but required more computational time.

Auru et al. [29] addressed the class imbalance problem in credit card fraud detection by developing an ensemble machine learning model that integrates multiple classifiers to enhance detection accuracy. Specifically, they combined Random Forest, Gradient Boosting, and CatBoost classifiers using a Stacking ensemble technique, where a meta-classifier (Logistic Regression) refines the final predictions. Their results based on Kaggle Credit card fraud showed that the stacked model achieved an accuracy of 96%, precision of 98% for fraudulent transactions, precision of 96% for non-fraudulent transactions, recall of 94%, and F1-score of 96%, outperforming traditional individual classifiers. To address data imbalance, Random Under-Sampling was applied to equalize the number of fraudulent and non-fraudulent transactions, ensuring unbiased predictions. The study demonstrates that using ensemble learning not only improves detection performance but also enhances adaptability to evolving fraud patterns.

Singh et al. [30] conducted a study on credit card fraud detection using Isolation Forest (IF) and Local Outlier Factor (LOF). The algorithms were applied to the European dataset (284,807 transactions, 492 fraudulent) [27] and the German dataset [31] (1,000 transactions, 300 fraudulent), with SMOTE, Random Undersampling, AllKNN, and SMOTE-ENN used for data balancing. The results showed that IF achieved 99.81% accuracy, 63% precision, 53% recall, and 54% F1-score on the European dataset, while LOF achieved 70.60% accuracy, 35% precision, 50% recall, and 41% F1-score on the German dataset when combined with Random

Undersampling. The study highlights the importance of data balancing techniques in improving fraud detection and reducing false positives and negatives.

Afriyie et al. [32] a supervised machine learning approach for credit card fraud detection by comparing the performance of Decision Tree, Random Forest, and Logistic Regression algorithms in classifying transactions as fraudulent or legitimate. Their study utilized a Kaggle dataset of simulated credit card transactions [43], which included both fraudulent and genuine records and covered transactions from January 1, 2020, to December 31, 2020. The study used undersampling to balance the number of fraudulent and non-fraudulent transactions to address the problem of class imbalance and make sure the models were not skewed toward the majority class. With an accuracy of 96%, precision of 9%, recall of 97%, and F1-score of 17%, the Random Forest model performed better than the other algorithms that were evaluated. According to the findings, the Random Forest model was the most efficient in identifying fraudulent credit card transactions, offering a notable enhancement over methods utilizing Decision Trees and Logistic Regression.

Sulaiman et al. [33] a hybrid fraud detection approach that combines Artificial Neural Networks (ANN) and Federated Learning (FL) to improve credit card fraud detection while maintaining data privacy. The model uses FL, which enables banks to train models locally without sharing raw data, to address issues like data imbalance and privacy concerns. The system generates synthetic fraudulent transactions using the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance and improve the model's capacity to identify uncommon fraud cases. Credit card datasets collected in real time were used to test the model. The proposed approach outperformed standard models such as Logistic Regression, SVM, and Random Forest, with precision of 95%, recall of 93%, and F1-score of 94%. The findings show that the hybrid ANN-FL model considerably enhances fraud detection accuracy while keeping.

Alamri et al. [34]used the Kaggle credit card dataset to address the issue of class imbalance in fraud detection through various sampling techniques. Oversampling methods such as SMOTE and Borderline-SMOTE, undersampling approaches like Tomek Links and Cluster Centroid, as well as hybrid methods, are employed to balance the data. Several machine learning algorithms including Decision Trees, Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), deep learning models such as

Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs), and ensemble methods like Random Forest and XGBoost are evaluated for their effectiveness in detecting fraudulent transactions. The study concludes that hybrid techniques, such as SMOTE combined with Edited Nearest Neighbors (SMOTE+ENN), Deep SMOTE, and cost-sensitive learning, significantly enhance accuracy and reduce false positive rates. Among the models tested, XGBoost combined with ADASYN achieved the best performance, reaching a fraud detection accuracy of 99%.

lleberi et al.[35] proposed a machine learning-based credit card fraud detection model using the Genetic Algorithm (GA) for feature selection. Their approach leverages Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), Artificial Neural Network (ANN), and Naive Bayes (NB) classifiers. For the European dataset, the Random Forest (RF) model achieved the best performance with 99.98% accuracy, 95.34% precision, 72.56% recall, and 82.41% F1-score. On the synthetic dataset, the Decision Tree (DT) achieved 100% accuracy, 99.51% precision, 99.71% recall, and 99.61% F1-score, demonstrating the superiority of the proposed method in detecting financial fraud.

Feng Gao et al.[36] used the Kaggle dataset and proposed ConNet, a deep semi-supervised anomaly detection model designed to handle highly imbalanced datasets. The preprocessing includes normalization and feature extraction to optimize deep learning performance. To handle class imbalance, contrastive learning is applied leveraging a few labeled anomalies and many unlabeled samples. The model is built using CNNs with contrastive loss for better feature representation. The model performance was evaluated with an ROC-AUC of 0.9740 and an PR-AUC of 0.7066, indicating it is effectiveness in detecting fraudulent transactions.

Kalid et al. [37] proposed a Multiple Classifier System (MCS) to handle class imbalance and overlapping class distributions in credit card fraud detection. For anomaly detection, they implemented a sequential classification approach, where C4.5 was used as the first classifier to classify out the majority class samples, followed by Naïve Bayes as the second classifier to classify out the minority class samples. Their

approach was applied to the Credit Card Fraud (CCF) and Credit Card Default Payment (CCDP)[38]datasets. achieving an accuracy of 99.9%, a recall of 87.2%, and True negative rate of 100% for the CCF dataset and an accuracy of 93%, a recall of 84%, and False negative rate of 95.5% for the CCDP dataset, outperforming traditional single-classifier method.

Gupta et al. [39] used the Kaggle dataset [25] and applied the Isolation Forest method, which is an unsupervised technique for anomaly detection implemented through H2O.ai. Due to the highly imbalanced nature of the dataset, the study did not use explicit resampling techniques. The performance of the model was assessed using the Area Under the Precision-Recall Curve (AUCPR), which achieved a score of 98.72%, alongside the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These results show that the Isolation Forest method works well for finding fraudulent transactions, making it a useful approach for fraud detection.

Rezapour [40] explored unsupervised anomaly detection to identify fraudulent transactions in credit card fraud detection. Their approach involved preprocessing the dataset with undersampling, following it by applying three anomaly detection techniques: One-Class SVM, Deep Autoencoder, and Robust Mahalanobis Distance. Their approach was evaluated on a real-world credit card transaction dataset. There was no comparison made across these three models in this research but based on the evaluation matrix we noticed the autoencoder performed the best with Accuracy = 89.0%, Precision = 90.7%, Recall = 87.2%, F1-score = 88.9% and the worse method was the Robust Mahalanobis Distance.

Ounacer et al. [14] explored the use of the Isolation Forest algorithm as an unsupervised anomaly detection method for credit card fraud. The study aims to detect fraudulent transactions in real time by comparing Isolation Forest to Local Outlier Factor (LOF), One-Class SVM (OCSVM), and K-Means. By using unsupervised learning without labeled data, the model successfully managed class imbalance using the Kaggle Credit Card Fraud Detection [25] which included 284,807 transactions and 492 fraud incidents. Isolation Forest beat the other approaches, with 97% accuracy, 85% precision, 89% recall, 87% F1-score, and 91% AUC. The outcomes demonstrated how well Isolation Forest performed in identifying fraudulent transactions while producing fewer false positives. The study suggests using Apache Spark in future

projects to increase real-time fraud detection capabilities.

Table 1 Presents a summary of previous studies on anomaly detection in Credit Card Fraud Detection:

Table 1: Summary of previous studies on anomaly Detection for Credit Card Fraud Detection

Ref	Year	Algorithms	How to handle Class Imbalance	Dataset	Evaluation Matrix
[25]	2025	**Support Vector Machine (SVM) classification model**, Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree	Oversampling, Undersampling, Correlation Matrix Analysis	Kaggle Credit card fraud [27]	Precision: 92%, Recall: 89%, FI-score: 90%, Accuracy: 90%
[26]	2024	GRU (Gated Recurrent Unit)	-	Kaggle Credit card fraud [27]	Accuracy = 99.9%, precision = 0.79, specificity = 0.99, recall = 0.94, and F1-score = 0.82
[28]	2024	K-Nearest Neighbors (KNN), **Random Forest (RF),** Support Vector Machine (SVM), Logistic Regression (LR)	Outlier detection techniques and removal of outliers from the oversampled dataset.	Kaggle Credit card fraud [27]	Accuracy = 99.96%, precision = 0.96, recall = 0.80, and F1-score = 0.88
[8]	2024	**Random Forest**, Decision Tree, KNN, Logistic Regression, CNN	SMOTE, AdaSyn, RUS, Hybrid Resampling	Kaggle Credit card fraud [27]	Accuracy: 99.99%, Precision: 99.98%, Recall: 100%, F1-score: 99.99%
[29]	2024	**Ensemble Model (Random Forest+ Gradient Boosting+ CatBoost)**	Random Under-Sampling	Kaggle Credit card fraud [27]	Accuracy 96%, Precision 98% for fraud, 96 % for non-fraud, Recall 94%, F1-score 96%
[30]	2024	**Isolation Forest (IF)** **Local Outlier Factor (LOF)**	SMOTE, Random Undersampling, AlIKNN, SMOTE-ENN	Kaggle Credit card fraud [27]	Accuracy:99.81%, Precision: 63%, Recall: 53%, F1-Score:54%

				German Credit Card Fraud dataset [31]	Accuracy:70.60%,Precision: 35%, Recall: 50%, F1-Score:41%
[32]	2023	Decision Tree (DT), **Random Forest (RF),** Logistic Regression (LR)	Undersampling	Kaggle Credit card fraud [47]	Accuracy of (96%), Precision of(9%), Recall of (97%), F1-score of (17%)
[33]	2022	**Artificial Neural Networks (ANN)**, Federated Learning (FL)**,** Synthetic Minority Oversampling Technique (SMOTE), Random Forest (RF), Support Vector Machine (SVM)	Synthetic Minority Oversampling Technique (SMOTE), Adaptive Synthetic Sampling (ADASYN) and RAMO Boost, Ensemble Classifiers	Real-World Dataset	Precision (95%), Recall (93%), F1-score (94%)
[34]		Decision Trees, Random Forest (RF), Logistic Regression (LR),NB , K-Nearest Neighbors (KNN), SVM, CNN, Long Short-Term Memory (LSTM), (DNN) **(XGBoost)**, **ADASYN** Adaptive Boosting (AdaBoost), Bagging Local Outlier Factor (LOF), Isolation Forest	oversampling, undersampling, and hybrid methods	Kaggle Credit card fraud [27]	Accuracy: 0.9900, Precision: 0.9500, Recall: 0.9200, F1-score: 0.9350
[35]	2022	**Decision Tree (DT)**, **Random Forest (RF)**, Logistic Regression (LR), Artificial Neural Network (ANN), Naïve Bayes (NB)	Genetic Algorithm (GA) for Feature Selection, Synthetic Minority Oversampling Technique (SMOTE)	Kaggle Credit card fraud [27] Synthetic Credit Card Fraud Dataset [41]	European: Accuracy: 99.98%, Precision: 95.34%, Recall: 72.56%, F1-Score: 82.41% Synthetic: Accuracy: 100% Precision: 99.51% Recall: 99.71%, F1-Score: 99.71%
[36]	2021	**CNNs**	contrastive learning is applied, leveraging a few labeled anomalies and many unlabeled samples	Kaggle Credit card fraud [27]	Performance evaluated AUC-ROC= 0.9740 and AUC-PR= 0.7066.
[37]	2020	**Multiple Classifier System (Random Forest with C4.5 followed by Naïve Bayes)**	Multiple Classifier System	Kaggle Credit card fraud [27] Credit Card Default Payment [38]	Accuracy: 99.9%, a recall of (0.872), and True negative rate of (1.00) accuracy of 93%, a recall of 84%, and True negative rate of 95.5%

[39]	2020	**Isolation Forest**	-	Kaggle Credit card fraud [27]	AUCPR=98.72%
[40]	2019	One-Class SVM, Deep Autoencoder, and Robust Mahalanobis Distance	UnderSamping	Real-World DataSet	Accuracy = 89.0%, Precision = 90.7%, Recall = 87.2%, F1-score = 88.9%
[14]	2018	**Isolation Forest (iForest),** Local Outlier Factor (LOF), One-Class Support Vector Machine (OCSVM), K-Means Clustering	unsupervised learning	Kaggle Credit card fraud [27]	accuracy of 97% , precision of 85%, recall of 89%, F1-score 87%.

3.2 Anomaly Detection with Generative AI Research

In this section, we explore studies that combine anomaly detection with generative AI to enhance fraud detection. The reviewed research examines how synthetic data generation improves model performance and helps address data imbalance.

Alshameri and Xia [42] proposed a CNN-based Variational Autoencoder (VAE) for credit card fraud detection using an anomaly detection approach. The algorithm was applied to the European dataset [27],They relied on anomaly detection to address the class imbalance problem without modifying the dataset distribution. They trained two VAE models: VAE-normal, which learned patterns from legitimate transactions, and VAE-anomaly, which focused on fraudulent transactions. They then developed the VAE-anomaly rescaled model, in which Fraud detection was established on reconstruction error, where fraudulent cases showed higher deviations. The model achieved a precision of 93%, recall of 92%, and F1-score of 92%, outperforming traditional classifiers such as Logistic Regression, SVM, and Random Forest.

Mienye at al. [43] used two datasets: the European credit card fraud dataset[27] and the Brazilian Credit Card Dataset [44]. To handle the class imbalance issue, the researchers employed stratified k-fold cross-validation. The proposed algorithms included hybrid models combining Generative Adversarial Networks (GAN) with LSTM, GRU, and Simple RNN architectures. The evaluation metrics on the European dataset demonstrated strong performance, with the GAN-GRU model achieving perfect precision (1.0), an F-measure of 0.996, sensitivity of 0.992, and specificity of 1.0. On

the Brazilian dataset, the GAN-LSTM model achieved a precision of 0.988, an F-measure of 0.953, sensitivity of 0.920, and specificity of 0.965, showcasing the effectiveness of the hybrid approach in credit card fraud detection.

Selvarajan et al. [45] used adversarial networks and generative artificial intelligence (GAI) to identify fraud from dynamic content policies. The Adversarial Network is used to distinguish between authentic and fraudulent information, while the GAI creates content patterns. 97% accuracy, precision 98%, recall 94%, and F1-score 96% were obtained by testing the model under four different sets of conditions where it must generate large patterns with fewer dynamic fluctuations. These results were superior to those of the conventional method, which had 77% accuracy. With improved space complexity, stability, and time complexity and fewer computational resources, the model detects fraud in real time using pattern recognition and Generative Adversarial Networks (GANs). Adaptive learning, dynamic scaling, and pattern dimension normalization are some of the methods that increase its capability.

Ding et al. [46] addressed the class imbalance problem in credit card fraud detection by using Generative AI, specifically an improved Variational Autoencoder Generative Adversarial Network (VAEGAN), to generate synthetic fraud transactions, this augmented dataset was used to train XGBoost, Logistic Regression, Decision Trees, Random Forest, and Neural Network. The algorithms were applied to the European dataset [27], Their results showed that XGBoost trained on the improved VAEGAN-augmented dataset achieved the highest accuracy 97.8%, precision 91.97%, recall 88.4%, and F1-score 88.4% compared to traditional resampling methods.

Ghaleb et al. [47] uses Kaggle dataset and examines different algorithms for detecting credit card fraud., focusing on handling class imbalance using Generative Adversarial Networks (GANs) and ensemble learning. The ESMOTE-GAN technique, which combines SMOTE and GAN models, achieved the highest accuracy, improving fraud detection by 3.2% with a 0% false alarm rate.

Jiang et al.[48] used two datasets: the European credit card fraud dataset[27] and the IEEE-CIS Fraud Detection dataset(private). To handle the class imbalance issue, the researchers employed an anomaly detection approach, training the model exclusively on normal transactions and treating fraudulent transactions as anomalies. The proposed

algorithm, UAAD-FDNet, combines an autoencoder with Generative Adversarial Networks (GAN) and incorporates a Feature Attention mechanism to enhance feature learning. The evaluation metrics demonstrated strong performance, with the European dataset achieving a precision of 0.98, recall of 0.76, and an F1-score of 0.85, while the IEEE-CIS dataset achieved a precision of 0.93, recall of 0.63, and an F1-score of 0.75.

Cheah et al. [49] used the European credit card fraud dataset[27]. To handle the class imbalance issue, the researchers employed several techniques, including SMOTE, GAN, SMOTified-GAN, SMOTE+GAN, and GANified-SMOTE. They implemented Feed-forward Neural Networks (FNN), Convolutional Neural Networks (CNN), and a hybrid FNN+CNN model for fraud detection. The evaluation metrics revealed that the GANified-SMOTE technique combined with the hybrid FNN+CNN model achieved the best results, with a precision of 0.94, recall of 0.85, and an F1-score of 0.89, demonstrating its effectiveness in accurately identifying fraudulent transactions.

Strelcenia et al. [50]considered the class imbalance problem of credit card fraud detection using Generative Adversarial Networks (GANs), i.e., CTAB-uGAN and WGAN-GP, to generate synthetic fraudulent transactions. The data was augmented and later used for training deep learning models to enhance classification performance. Their results demonstrated that GAN-RF presented the highest accuracy (99.83%), SDG-GAN reported a recall of 80.90% and precision of 98.63%, while CTAB-GAN reported an F1-score of 27.4%. The study showed that GAN-based augmentation was outperforming traditional oversampling methods like SMOTE in tackling imbalanced datasets, improving fraud detection effectiveness.

Fiore et al [4] uses the European credit card fraud dataset[27]. To address the class imbalance issue, the researchers employed Generative Adversarial Networks (GANs) to generate synthetic samples of the minority class, which were then merged with the original training data to create a more balanced dataset. The proposed algorithm combined GANs with a deep neural network-based classifier for fraud detection. The evaluation metrics demonstrated strong performance, with an accuracy of 0.99, precision of 0.93, F-measure of 0.82, sensitivity of 0.73, and specificity of 0.99, highlighting the effectiveness of the hybrid approach in improving classification performance for credit card fraud detection.

Table 2 Presents a summary of previous studies on anomaly detection with generative AI in Credit Card Fraud Detection:

Table 2 : Summary of Anomaly Detection with Generative AI Approaches in Credit Card Fraud Detection

Ref	Year	Algorithms	How to handle Class Imbalance	Dataset	Evaluation Matrix
[42]	2024	**CNN-based Variational Autoencoder**, Logistic Regression, SVM, and Random Forest.	-	Kaggle Credit card fraud [27]	precision = 0.93, recall = 0.92, and F1-score = 0.92
[43]	2024	**GAN-LSTM**, **GAN-GRU**, GAN-Simple RNN, LSTM, GRU, Simple RNN.	stratified k-fold cross-validation	Kaggle Credit card fraud [27]	**(GAN-GRU)** precision=1, F-measure=0.996, Sensitivity=0.992, Specificity=1
				Brazilian Credit Card Dataset[44]	**(GAN-LSTM)** precision=0.988, F-measure=0.953, Sensitivity=0.920, Specificity=0.965
[45]	2024	**Pattern Recognition Techniques,** Dynamic Scaling and Adaptive Learning, Pattern Dimension Normalization, Optimization Algorithms	GANs, Pattern Dimension Normalization, Adaptive Learning, Dynamic Scaling	Real-World Dataset	Accuracy (97%), Precision (98%), Recall (94%), F1-Score (96%)

[46]	2023	**XGBoost**, Logistic Regression, Decision Trees, Random Forest, and Neural Network.	VAEGAN	Kaggle Credit card fraud [27]	Accuracy = 0.978, Precision = 0.9197, Recall = 0.884, F1-score = 0.884
[47]	2023	**ESMOTE-GAN** ,RF, SMOTE ,XGBoost ,KNN,SVM ,LR,ANN	ESMOTE-GAN technique, which combines SMOTE and GAN models	Kaggle Credit card fraud [27]	ESMOTE-GAN= improving fraud detection by 3.2% with a 0% false alarm rate.
[48]	2023	**UAAD-FDNet (Autoencoder + GAN)**	Anomaly detection, trained only on normal transactions, Feature Attention mechanism.	Kaggle Credit card fraud [27]	Precision =0.98, Recall= 0.76, F1-score= 0.85
				IEEE-CIS Fraud Detection (private)	Precision= 0.93, Recall= 0.63, F1-score= 0.75
[49]	2023	FNN, CNN, Hybrid **FNN+CNN**	SMOTE, GAN, SMOTified-GAN, SMOTE+GAN, **GANified**-SMOTE	Kaggle Credit card fraud [27]	Precision: 0.94, Recall: 0.85, F1-score: 0.89
[50]	2023	**Random Forest,** Naïve Bayes, C4.5 Decision Tree.	GANs, CTAB-GAN, WGAN-GP, SDG-GAN, OCAN, cWGAN, Duo-GAN, Tuned-GAN, XGBoost	Real-World Dataset	Accuracy (99.83%), recall of (80.90%), precision of (98.63%), F1-score of (27.4%)
[4]	2019	**Generative Adversarial Networks (GANs)with Deep Neural Networks (DNNs)**	GAN-generated synthetic samples	Kaggle Credit card fraud [27]	Accuracy=99%,Precision=0.93,Fmeasure=0.82,Sensitivity=0.73, Specificity=0.99.

3.3 Discussion

Data augmentation has been recognized as a useful method for addressing class imbalance in fraud detection. Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have been extensively used to generate synthetic fraudulent transactions, Research conducted by Ding et al. [46] and Strelcenia et al. [50] has shown that the application of Generative AI models, such as VAEGAN and GAN-based methods, enhances the performance of fraud detection by generating synthetic fraudulent transactions. Also, the study by Cheah et al. [49] shows that the GANified-SMOTE technique, combined with FNN+CNN, yields higher precision and recall than conventional oversampling methods. These findings suggest that data augmentation can significantly improve fraud detection models by providing more representative fraudulent transaction samples.

Regarding the type of approach used in fraud detection, supervised learning is the dominant methodology. Many studies, such as Ali et al. [28] and Mienye et al. [26], rely on supervised models like Random Forest, XGBoost, and Neural Networks to classify transactions and the most commonly used algorithm for credit card fraud detection is Random Forest. Many studies, such as Ali et al. [28], Auru et al. [29], and Abdul Salam et al. [8], have shown that Random Forest consistently achieves high accuracy, precision, recall, and F1-score in fraud detection tasks. However, some research, such as Rezapour [40] and Gupta et al. [39], explores unsupervised learning techniques like Isolation Forest and One-Class SVM, which do not require labeled fraud data. While unsupervised methods are useful for detecting unknown fraud patterns, supervised approaches remain the preferred choice due to their higher accuracy and reliability when labeled data is available.

The datasets used in fraud detection research vary between public and private datasets. The most commonly used public dataset is the European Credit Card Fraud Dataset from Kaggle [27], featured in 22 studies, including Ali et al. [28], Cheah et al. [49], and Ding et al. [46]. Other public datasets include the German Credit Card Fraud dataset [31], the Credit Card Default Payment [38], and the Brazilian Credit Card Dataset[44]. Additionally, some studies, such as Jiang et al. [48] and Alamri et al.[34], use private financial institution datasets and the IEEE-CIS Fraud Detection Dataset, which are often more realistic but not publicly accessible.

3.4 Summary

This chapter provides a review of previous research related to credit card fraud detection. It discusses studies that applied anomaly detection as well as those that combined anomaly detection with generative AI techniques. The chapter explains how researchers addressed the issue of class imbalance and worked to improve fraud detection models. It also summarizes the models, datasets, and evaluation metrics used across the reviewed studies.

Chapter 4: Methodology:

This chapter outlines the methodology that we used to address the problem of credit card fraud detection using a combination of anomaly detection techniques and Generative AI for data augmentation. The overall goal is to develop a system that can effectively detect fraudulent transactions, particularly in the context of class imbalance where fraudulent samples are very limited. Figure 7 presents the methodology.

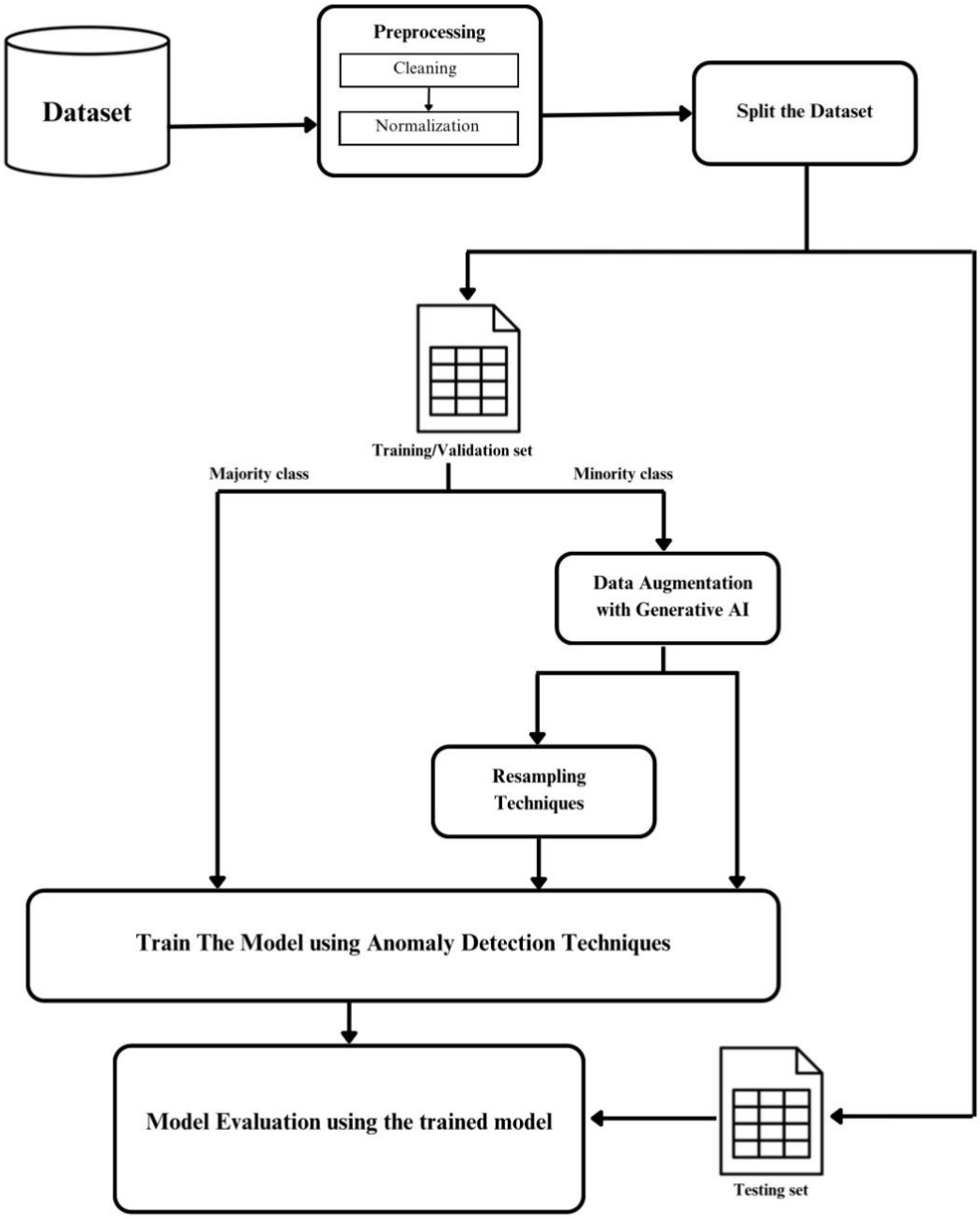


Figure 7 Methodology

4.1 Preprocessing

In our project, the preprocessing stage was important to prepare the data before training. It included cleaning the data by removing missing, repeated, or incorrect values. After that, normalization was applied to make sure all numerical features are on the same scale. These steps helped improve the quality of the data and made it easier for the model to learn correctly.

4.2 Data Division

Before balancing the dataset using both Generative Adversarial Networks (GANs) and a hybrid resampling method (SMOTE + GAN), we divided the final dataset into 80% for training and validation and 20% for testing. The model was trained on a balanced dataset and evaluated on data that better represents real-world fraud patterns. This helped improve the reliability of the evaluation metrics and reduced the impact of the original class imbalance.

4.3 Data Augmentation with Generative AI

For generating examples from fraudulent transactions, a VAE-GAN framework is chosen. The VAE-GAN hybrid model is effective because it combines the benefits of both VAE's diversity and GAN's fidelity .

VAEs are good at compressing data into a latent space, which leads to many possible variations for each feature as mentioned in chapter 2 section 2.4.2. This enables the creation of new data examples. However, sometimes the outputs lack sharpness and detail, which are characteristics of high-quality data. As a result, examples generated by VAE alone may appear unclear. On the other hand, GANs produce high-quality, realistic examples because of their Discriminator as mentioned in chapter 2 section 2.4.1 GANs can generate sharp, high-definition new data through adversarial training. However, they face challenges such as training instability and mode collapse which is a situation where the variety of the data produced is too limited [54].

By using a hybrid approach, such as VAE-GAN Figure 8, we can combine the strengths of both models.

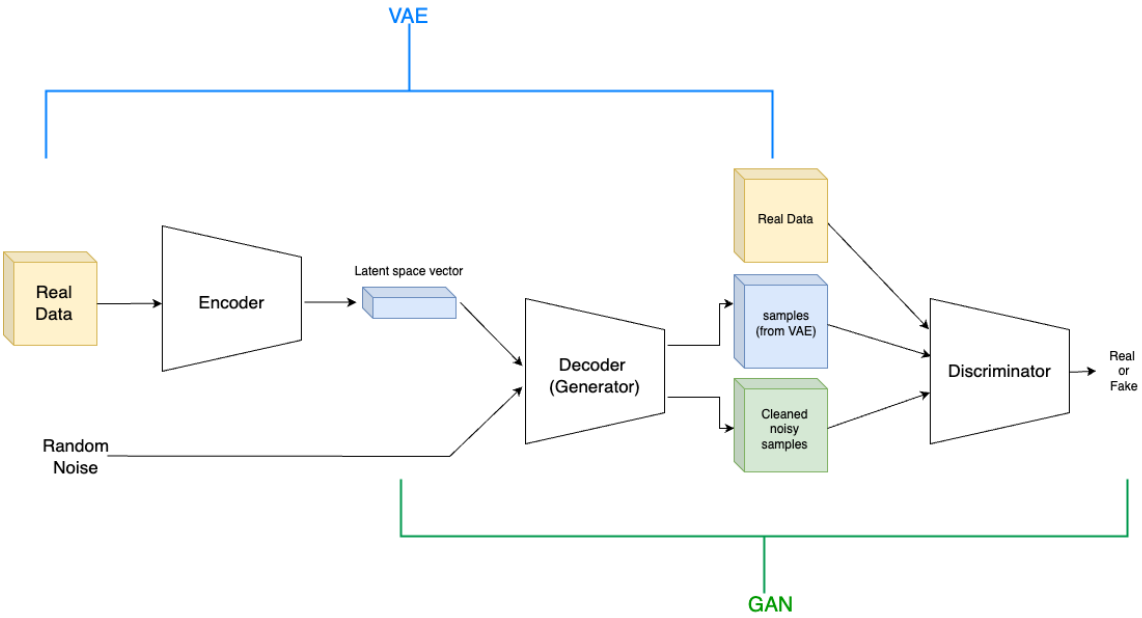


Figure 8 VAE-GAN architecture

4.4 Resampling Technique

After applying generative AI for data augmentation, we further addressed the imbalance between fraud and normal transactions using the Synthetic Minority Oversampling Technigue (SMOTE). SMOTE creates new fraud samples by generating points between existing ones. This helps improve the quality and variety of the fraud class, making it better represented in the dataset. It also helps reduce the model's bias toward the majority class and increases its ability to identify rare fraud cases. To evaluate the effectiveness of this approach, we will compare two methods: one that uses only Generative AI and another that combines Generative AI with SMOTE resampling, evaluating which method achieves better fraud detection performance.

4.5 Model Training Using Anomaly detection techniques

For model training, we will apply two techniques: a supervised learning algorithm, such as Random Forest, and an unsupervised learning algorithm, such as One-Class Support Vector Machine (OC-SVM). We will evaluate their performance to identify each method's strengths and weaknesses in detecting fraud.

4.5.1 Random forest

The Random Forest algorithm is an ensemble learning method that builds several decision trees during the training process and determines the final class based on the majority vote from these trees[55]. This approach is especially beneficial for classification tasks involving large and complex datasets, as it helps mitigate overfitting and enhances predictive accuracy[55].

In this research, Random Forest is used as part of the anomaly detection framework to classify transactions as fraudulent or non-fraudulent. The algorithm is trained on the augmented training dataset, allowing it to learn patterns that differentiate between normal and anomalous behaviors. Each decision tree in the forest is trained on a random subset of the data and features, which promotes model diversity and robustness.

To better understand how Random Forest works, Figure 9 below provides a visual representation of the process. It shows how multiple decision trees are built from randomly sampled subsets of the training data, each making an independent prediction. These predictions are then combined using majority voting to produce the final classification.

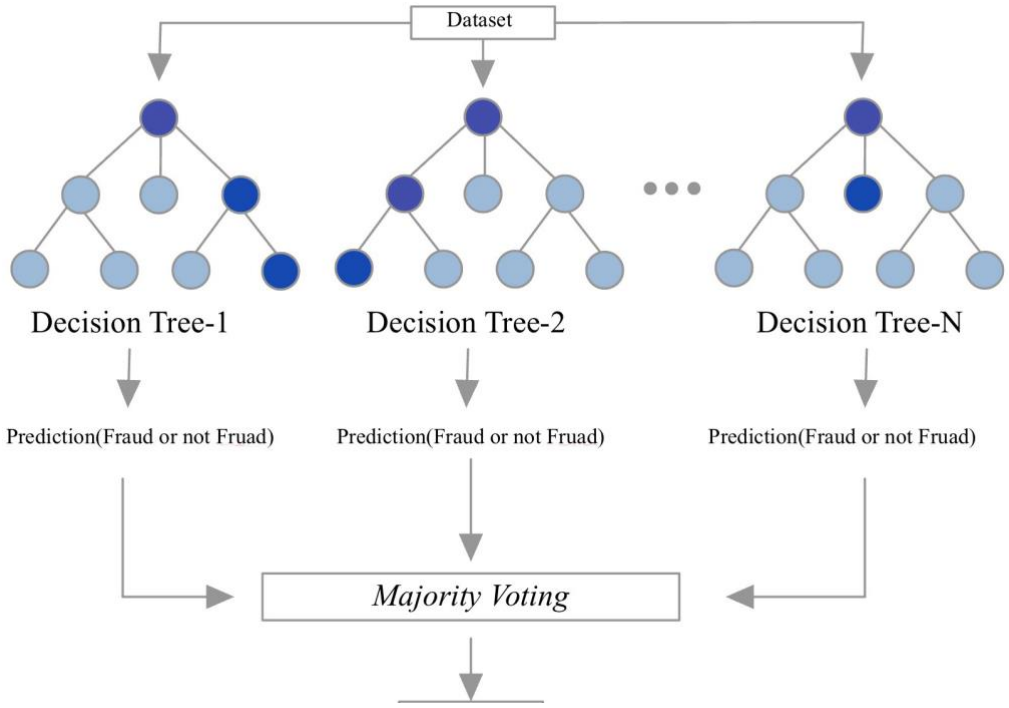


Figure 9 Random Forest Classifier

4.5.2 One-class Support Vector Machine

One-class SVM (OC-SVM) algorithm is one of the unsupervised machine learning. It is a commonly used approach for detecting anomalies. One-class SVMs separate a high- dimensional feature space into two regions: one for normal instances and one for outliers. This is done by constructing an optimal hyperplane that maximizes the margin between the normal instances and the origin. In this way, One-Class SVMs learn to encapsulate the characteristics of normal data instances and provide a measure of their separability.

We will use the One-Class Support Vector Machine (SVM) algorithm for anomaly detection. One-Class SVM is particularly effective in scenarios where the data predominantly comprises normal observations, and anomalies are rare or not well-defined. The algorithm works by learning a decision function that captures the distribution of the normal data and identifies any instances that significantly deviate from this learned boundary as anomalies. By leveraging the One-Class SVM,our goal is to build a robust model capable of accurately distinguishing between normal and anomalous patterns within the dataset, thereby improving the reliability and effectiveness of the anomaly detection process.

4.6 Model Evaluation

The trained models are evaluated on the testing set using standard metrics: accuracy, precision, recall, and F1-score. Due to the imbalanced nature of the dataset, we focus primarily on the F1-score, as it provides a better balance between precision and recall, reflecting the model's ability to correctly detect fraud without being biased by the majority class.

4.7 Summary

This chapter outlines the methodology we will employ, beginning with the preprocessing of the dataset. We will then divide it into training/validation and testing sets. Following this, we will apply data augmentation with generative AI to the minority class, along with resampling techniques. We will train the entire dataset using anomaly detection models, specifically random forest and one-class SVM. Finally, we will evaluate the model using evaluation metrics.

Chapter 5: Experimental Design

This chapter describes the experimental setup for improving credit card fraud detection. It covers the dataset used, the evaluation metrics applied to measure model performance, the hypotheses formulated for testing, and the tools and libraries utilized to build and evaluate the models.

5.1 Dataset Overview

In our project, we decided to use the European Credit Card Fraud Detection Dataset, which we found on Kaggle[54]. This dataset was a good choice because it comes from real-world credit card transactions and is often used to test fraud detection models. One of the main challenges in this dataset is that it is highly imbalanced, meaning there are very few fraud cases compared to normal ones.

The dataset has 284,807 transactions, and only 492 of them are labeled as fraud, which is around 0.17%. This imbalance makes it difficult for models to detect fraud correctly, and that is why we focused on solving this issue using Generative Adversarial Networks (GANs). Each transaction includes 30 features. Most of them were transformed anonymous features using a technique called Principal Component Analysis (PCA) to hide personal and sensitive information. These features are named V1 to V28, and their meanings are not known. Even though we do not know exactly what they represent, they are still useful for training machine learning models.

As shown in Table 3, The dataset also includes two original features: Time and Amount. The Time feature indicates how many seconds have passed since the first transaction in the dataset, while Amount shows the value of the transaction in euros. In addition to these, there is the Class feature, which is the target variable we aim to predict. It is a binary label where a value of 0 represents a normal (non-fraudulent) transaction, and a value of 1 represents a fraudulent transaction

Table 3 Dataset Features

Feature	Type	Description
Time	Numerical	Time in seconds since the first transaction
Amount	Numerical	Value of the transaction in euros
V1 to V28	Numerical	PCA-transformed anonymous features
Class	Categorical	Target label: 0 = normal, 1 = fraud

5.2 Evaluation Metrix

To evaluate the performance of the proposed models in detecting credit card fraud, we are focusing on: Accuracy, Precision, Recall, and F1-score. These metrics provide a comprehensive understanding of how well the models perform, particularly in the context of highly imbalanced datasets where the F1-score provides a balance between Precision and Recall and is more informative than Accuracy alone[56].

Table 4 represents the confusion matrix, which helps us understand how well the model is performing. It compares the actual labels (fraud or legitimate) with the model's predictions and is used to calculate important metrics like accuracy, precision, recall, and F1-score[56].

Table 4 confusion matrix

	Predicted: Fraud	Predicted: Legitimate
Actual: Fraud	True Positive (TP)	False Negative (FN)

Actual: Legitimate	False Positive (FP)	True Negative (FN)

Metric Definitions:

- **Accuracy**:

  Measures the overall correctness of the model.

$$Accuracy = \frac{TP + TN}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

- **Precision**:

  Measures how many of the predicted fraud cases were actually fraud[56].

$$Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}}$$

- **Recall (Sensitivity)**:

  Measures how many actual fraud cases were correctly identified.

$$Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}}$$

- **F1-Score**:

  The harmonic mean of precision and recall. It is the most useful metric when dealing with class imbalance.

$$F1 - Score \ = 2 \text{ x } \frac{precision \ x \ Recall}{\mathrm{Precision} + \mathrm{Recall}}$$

5.3 Hypothesis

These hypotheses guide the project and help in evaluating the proposed methodology:

- **Hypothesis 1:** Incorporating generative AI techniques into credit card fraud detection systems will significantly improve detection performance—measured by precision, recall, and F1-score—through synthetic data augmentation.

- **Hypothesis 2:** Using hybrid generative models like VAE-GAN will produce more realistic and diverse synthetic fraud samples compared to using VAE or GAN alone.
- **Hypothesis 3:** Combining synthetic samples with classical models like Random Forest and One-Class SVM will result in more accurate and reliable fraud detection in imbalanced datasets.

These hypotheses will be tested through experiments using real-world datasets, and their results will help evaluate the success of our proposed approach.

5.4 Simulation Tool

We will use the Python programming language to build and test our fraud detection models. Python is widely used in machine learning because it is easy to use and has many useful libraries. For building generative models like VAE, GAN, and VAE-GAN, we plan to use TensorFlow and Keras. For data preprocessing and evaluation, we intend to use libraries such as Scikit-learn, Pandas, and NumPy. We may also use Matplotlib and Seaborn to visualize our results.

We chose Python because it offers many tools and libraries that help us handle complex models and large datasets more easily. As for the development environment, we might use Google Colab as our development environment because it gives us free GPU access and makes it easier for the team to work together.

5.5 Summary

This chapter presented the experimental setup for evaluating credit card fraud detection using generative AI. We used the European dataset from Kaggle, which is

highly imbalanced. We will evaluate the model performance using accuracy, precision, recall, and F1-score, with a focus on F1-score due to class imbalance. Three hypotheses were proposed that we plan to use, especially regarding the impact of VAE-GAN and classical models like Random Forest. Python and libraries such as TensorFlow, Keras, and Scikit-learn were used, and experiments were conducted in Google Colab for collaborative development and GPU support.

Chapter 6: Conclusion

The growing use of credit cards and digital transactions has led to a significant rise in fraudulent activities. In response, our project aimed to enhance credit card fraud detection by leveraging advanced machine learning techniques, specifically Generative

Adversarial Networks (GANs) and Variational Autoencoders (VAEs). During the preprocessing phase, we focused on data cleaning, which involved removing missing, duplicate, or incorrect entries, and applied normalization to ensure consistency across numerical features. These steps were essential to improve the overall data quality for model training.

A major challenge in fraud detection is the severe class imbalance, as legitimate transactions vastly outnumber fraudulent ones. To mitigate this issue, we implemented a hybrid resampling approach that combines SMOTE and GANs, thereby enhancing the diversity and realism of synthetic fraud samples and improving the model's sensitivity to rare fraudulent cases. Through our literature review, we found that integrating generative AI with traditional machine learning models, such as Random Forests, significantly enhances performance metrics including accuracy, recall, and F1-score compared to standard oversampling methods. Moreover, unsupervised learning techniques One-Class SVM have shown effectiveness in identifying previously unseen fraud patterns; however, supervised learning methods generally outperform when labeled fraud data is available.

In this project, we utilized a VAE-GAN hybrid model, combining the Variational Autoencoder's ability to generate diverse samples with the GAN's strength in producing highly realistic data, to synthesize high-quality fraud samples and achieve better dataset balance. Furthermore, we explored various anomaly detection techniques and compared their performance against both traditional fraud detection models and AI-driven approaches, aiming to identify the most effective method for fraud detection.

Locally, this work assists banks and financial institutions in improving the reliability of fraud detection systems, minimizing financial losses, and enhancing customer trust. More broadly, it demonstrates how AI and synthetic data generation can address real-world challenges in cybersecurity and financial security. Given the changing strategies employed by fraudsters and the increasing subtlety of anomalous behavior, it is crucial to continue advancing techniques that can effectively model complex data distributions, incorporating both sophisticated resampling strategies and powerful anomaly detection methods.

References

[1] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection,” ACM Comput Surv, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.

[2] R. Bin Sulaiman, V. Schetinin, and P. Sant, “Review of Machine Learning Approach on Credit Card Fraud Detection,” Human-Centric Intelligent Systems, vol. 2, no. 1–2, pp. 55–68, Jun. 2022, doi: 10.1007/s44230-022-00004-0.

[3] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective,” Expert Syst Appl, vol. 41, no. 10, pp. 4915–4928, Aug. 2014, doi: 10.1016/j.eswa.2014.02.026.

[4] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” Inf Sci (N Y), vol. 479, pp. 448–455, Apr. 2019, doi: 10.1016/j.ins.2017.12.030.

[5] K. Chaudhary, J. Yadav, and B. Mallick, “A review of fraud detection techniques: Credit card,” Int J Comput Appl, vol. 45, no. 1, pp. 39–44, 2012.

[6] R. Bin Sulaiman, V. Schetinin, and P. Sant, “Review of Machine Learning Approach on Credit Card Fraud Detection,” Human-Centric Intelligent Systems, vol. 2, no. 1–2, pp. 55–68, Jun. 2022, doi: 10.1007/s44230-022-00004-0.

[7] U. Machine Learning Group, “Credit Card Fraud Detection,” Kaggle. Accessed: Feb. 22, 2025. [Online]. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

[8] M. Abdul Salam, K. M. Fouad, D. L. Elbably, and S. M. Elsayed, “Federated learning model for credit card fraud detection with data balancing techniques,” Neural Comput Appl, vol. 36, no. 11, pp. 6231–6256, Apr. 2024, doi: 10.1007/s00521-023-09410-2.

[9] C. Huyen, Designing Machine Learning Systems, 1st ed. Sebastopol, CA, USA: O’Reilly Media, 2022.

[10] Y. SUN, A. K. C. WONG, and M. S. KAMEL, “CLASSIFICATION OF IMBALANCED DATA: A REVIEW,” Intern J Pattern Recognit Artif Intell, vol. 23, no. 04, pp. 687–719, Jun. 2009, doi: 10.1142/S0218001409007326.

[11] A. Singh, R. K. Ranjan, and A. Tiwari, “Credit Card Fraud Detection under Extreme Imbalanced Data: A Comparative Study of Data-level Algorithms,” Journal of

Experimental & Theoretical Artificial Intelligence, vol. 34, no. 4, pp. 571–598, Jul. 2022, doi: 10.1080/0952813X.2021.1907795.

[12] T. K. Dang, T. C. Tran, L. M. Tuan, and M. V. Tiep, “Machine Learning Based on Resampling Approaches and Deep Reinforcement Learning for Credit Card Fraud Detection Systems,” Applied Sciences, vol. 11, no. 21, p. 10004, Oct. 2021, doi: 10.3390/app112110004.

[13] R. Mohammed, J. Rawashdeh, and M. Abdullah, “Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results,” in 2020 11th International Conference on Information and Communication Systems (ICICS), IEEE, Apr. 2020, pp. 243–248. doi: 10.1109/ICICS49469.2020.239556.

[14] C. Seiffert, J. Van Hulse, and T. M. Khoshgoftaar, “Hybrid sampling for imbalanced data,” Integr Comput Aided Eng, vol. 16, no. 3, pp. 193–210, Jun. 2009.

[15] S. Ounacer, H. Ait, E. Bour, Y. Oubrahim, M. Y. Ghoumari, and M. Azzouazi, “Using Isolation Forest in anomaly detection: the case of credit card transactions,” vol. 6, no. 2, pp. 394–400, 2018, [Online]. Available: http://pen.ius.edu.ba

[16] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation Forest,” in 2008 Eighth IEEE International Conference on Data Mining, IEEE, Dec. 2008, pp. 413–422. doi: 10.1109/ICDM.2008.17.

[17] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, 1st ed. Cambridge, MA, USA: MIT Press, 2016.

[18] L. Ruff et al., “Deep One-Class Classification,” in Proceedings of the 35th International Conference on Machine Learning (ICML 2018), PMLR, 2018, pp. 4393–4402.

[19] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” Inf Sci (N Y), vol. 479, pp. 448–455, Apr. 2019, doi: 10.1016/j.ins.2017.12.030.

[20] I. J. Goodfellow et al., “Generative Adversarial Networks,” Jun. 2014.

[21] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” Inf Sci (N Y), vol. 479, pp. 448–455, Apr. 2019, doi: 10.1016/j.ins.2017.12.030.

[22] R. Szeliski, “Computer Vision: Algorithms and Applications 2nd Edition,” 2021. [Online]. Available: https://szeliski.org/Book,

[23] T. Sattarov, M. Schreyer, and D. Borth, “FinDiff: Diffusion Models for Financial Tabular Data Generation,” Sep. 2023, [Online]. Available: http://arxiv.org/abs/2309.01472

[24] C. Yang, T. Wang, and X. Yan, “DDMT: Denoising Diffusion Mask Transformer Models for Multivariate Time Series Anomaly Detection,” Oct. 2023, [Online]. Available: http://arxiv.org/abs/2310.08800

[25] M. A. Ebrahim, W. Ghonim, and A. H. Abd El-Aziem, “IJEMS A Prototype for Credit Card Fraud Detection System,” 2025.

[26] I. D. Mienye and N. Jere, “Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions,” IEEE Access, vol. 12, pp. 96893–96910, 2024, doi: 10.1109/ACCESS.2024.3426955.

[27] U. Machine Learning Group, “Credit Card Fraud Detection,” Kaggle. Accessed: Feb. 22, 2025. [Online]. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

[28] A. H. Ali and A. Ali Hagag, “An enhanced AI-based model for financial fraud detection,” International Journal of Advanced and Applied Sciences, vol. 11, no. 10, pp. 114–121, Oct. 2024, doi: 10.21833/ijaas.2024.10.013.

[29] F. Auru, S. Adewumi, and V. I. Yemi-Peter, “An Ensemble Model for Credit Card Fraudulent Transactions Detection and Classification,” 2024. [Online]. Available: www.ftstjournal.com

[30] P. Singh, K. Singla, P. Piyush, and B. Chugh, “Anomaly Detection Classifiers for Detecting Credit Card Fraudulent Transactions.”

[31] UCI Machine Learning Repository, “German Credit Dataset.” Accessed: Mar. 03, 2025. [Online]. Available: https://www.kaggle.com/datasets/uciml/german-credit

[32] J. K. Afriyie et al., “A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions,” Decision Analytics Journal, vol. 6, Mar. 2023, doi: 10.1016/j.dajour.2023.100163.

[33] R. Bin Sulaiman, V. Schetinin, and P. Sant, “Review of Machine Learning Approach on Credit Card Fraud Detection,” Human-Centric Intelligent Systems, vol. 2, no. 1–2, pp. 55–68, Jun. 2022, doi: 10.1007/s44230-022-00004-0.

[34] M. Alamri and M. Ykhlef, “Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques,” Electronics (Basel), vol. 11, no. 23, p. 4003, Dec. 2022, doi: 10.3390/electronics11234003.

[35] E. Ileberi, Y. Sun, and Z. Wang, “A machine learning based credit card fraud detection using the GA algorithm for feature selection,” J Big Data, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[36] F. Gao, J. Li, R. Cheng, Y. Zhou, and Y. Ye, “ConNet: Deep Semi-Supervised Anomaly Detection Based on Sparse Positive Samples,” IEEE Access, vol. 9, pp. 67249–67258, 2021, doi: 10.1109/ACCESS.2021.3077014.

[37] S. N. Kalid, K.-H. Ng, G.-K. Tong, and K.-C. Khor, “A Multiple Classifiers System for Anomaly Detection in Credit Card Data With Unbalanced and Overlapped Classes,” IEEE Access, vol. 8, pp. 28210–28221, 2020, doi: 10.1109/ACCESS.2020.2972009.

[38] I.-C. Yeh, “Default of Credit Card Clients,” 2009, UCI Machine Learning Repository. Accessed: Feb. 22, 2025. [Online]. Available: https://doi.org/10.24432/C55S3H

[39] M. S. Gupta, S. Patel, S. Kumar, and G. Chauhan, “ANOMALY DETECTION IN CREDIT CARD TRANSACTIONS USING MACHINE LEARNING,” International Journal of Innovative Research in Computer Science & Technology, vol. 8, no. 3, May 2020, doi: 10.21276/ijircst.2020.8.3.5.

[40] M. Rezapour, “Anomaly Detection using Unsupervised Methods: Credit Card Fraud Case Study,” International Journal of Advanced Computer Science and Applications, vol. 10, no. 11, 2019, doi: 10.14569/IJACSA.2019.0101101.

[41] IBM Research, “Synthetic Credit Card Fraud Dataset.” Accessed: Mar. 03, 2025. [Online]. Available: https://ibm.ent.box.com/v/tabformer-data/folder/130747715605

[42] F. Alshameri and R. Xia, “An Evaluation of Variational Autoencoder in Credit Card Anomaly Detection,” Big Data Mining and Analytics, vol. 7, no. 3, pp. 718–729, Sep. 2024, doi: 10.26599/BDMA.2023.9020035.

[43] I. D. Mienye and T. G. Swart, “A Hybrid Deep Learning Approach with Generative Adversarial Network for Credit Card Fraud Detection,” Technologies (Basel), vol. 12, no. 10, Oct. 2024, doi: 10.3390/technologies12100186.

[44] “Default of Credit Card Clients Dataset.” Accessed: Mar. 04, 2025. [Online]. Available: https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset

[45] S. Selvarajan et al., “Generative artificial intelligence and adversarial network for fraud detections in current evolutional systems,” Expert Syst, Feb. 2024, doi: 10.1111/exsy.13740.

[46] Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang, “Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network,” IEEE Access, vol. 11, pp. 83680–83691, 2023, doi: 10.1109/ACCESS.2023.3302339.

[47] F. A. Ghaleb, F. Saeed, M. Al-Sarem, S. N. Qasem, and T. Al-Hadhrami, “Ensemble Synthesized Minority Oversampling-Based Generative Adversarial Networks and Random Forest Algorithm for Credit Card Fraud Detection,” IEEE Access, vol. 11, pp. 89694–89710, 2023, doi: 10.1109/ACCESS.2023.3306621.

[48] S. Jiang, R. Dong, J. Wang, and M. Xia, “Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network,” Systems, vol. 11, no. 6, Jun. 2023, doi: 10.3390/systems11060305.

[49] P. C. Y. Cheah, Y. Yang, and B. G. Lee, “Enhancing Financial Fraud Detection through Addressing Class Imbalance Using Hybrid SMOTE-GAN Techniques,” International Journal of Financial Studies, vol. 11, no. 3, Sep. 2023, doi: 10.3390/ijfs11030110.

[50] E. Strelcenia and S. Prakoonwit, “A Survey on GAN Techniques for Data Augmentation to Address the Imbalanced Data Issues in Credit Card Fraud Detection,” Mach Learn Knowl Extr, vol. 5, no. 1, pp. 304–329, Mar. 2023, doi: 10.3390/make5010019.

[51] M. Alamri and M. Ykhlef, “Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques,” Electronics (Basel), vol. 11, no. 23, p. 4003, Dec. 2022, doi: 10.3390/electronics11234003.

[52] M. S. Gupta, S. Patel, S. Kumar, and G. Chauhan, “ANOMALY DETECTION IN CREDIT CARD TRANSACTIONS USING MACHINE LEARNING,” International Journal of Innovative Research in Computer Science & Technology, vol. 8, no. 3, May 2020, doi: 10.21276/ijircst.2020.8.3.5.

[53] R. S. A and B. Sathish Babu, “Synthesizing Realistic Knee MRI Images: A VAE-GAN Approach for Enhanced Medical Data Augmentation,” 2024. [Online]. Available: www.ijacsa.thesai.org

[54] “kaggle dataset.”

[55] Belgiu, Mariana, and Lucian Drăguţ. "Random forest in remote sensing: A review of applications and future directions." ISPRS journal of photogrammetry and remote sensing 114 (2016): 24-31.
[56] Hossin, Mohammad, and Md Nasir Sulaiman. "A review on evaluation metrics for data classification evaluations." International journal of data mining & knowledge management process 5.2 (2015): 1.