

# Leveraging Machine Learning Algorithms for Fraud Detection in FinTech: A Comparative Study of Supervised and Unsupervised Techniques

## ABSTRACT

This secondary research study examines financial fraud detection using machine learning algorithms using a hybrid qualitative and quantitative approach. The objective is to evaluate fraud detection and prevention machine learning algorithms. The study uses PRISMA to process 50 publications to find 6 that meet inclusion and exclusion criteria. This thorough and transparent screening method makes the study more trustworthy. The qualitative research uses thematic analysis of selected articles to identify fraud detection machine learning algorithm themes and trends. This detailed review reveals the literature's various methods and results, exposing strengths and flaws. These algorithms' fraud detection performance is objectively assessed, revealing their financial security benefits. Qualitative and quantitative methods fill knowledge gaps and complete the picture. Conclusion, this work provides a detailed and comprehensive overview of machine learning in fraud detection literature. PRISMA selects articles methodologically, while theme analysis adds qualitative depth. Researchers, industry practitioners, and regulators learn about financial fraud detection using machine learning in this publication.

## TABLE OF CONTENTS

<b>Chapter 1: Introduction .....</b>	<b>5</b>
1.1 Introduction.....	5
1.1.1 ML algorithms in Fraud detection.....	6
1.2 Background.....	7
1.3 Research Problem .....	8
1.4 Research Gap.....	9
1.5 Aim .....	10
1.6 Objectives .....	10
1.7 Significance of study.....	11
<b>Chapter 2: Research methodology .....</b>	<b>12</b>
2.1 Research Questions .....	12
2.2 Research Design .....	12
2.3 Research Approach.....	12
2.4 Research Philosophy .....	13
2.5 Secondary Research Method .....	14
2.6 PRISMA Model.....	14
2.7 Inclusion and exclusion Criteria.....	15
2.8 Data Analysis .....	16
<b>Chapter 3: Literature Review .....</b>	<b>18</b>
3.1 Historical Overview of Fraud Detection Methods.....	18
3.1.1 Traditional Rule-Based Systems.....	18
3.1.2 Manual Verification Processes.....	19
3.1.3 Limitations and Challenges of Conventional Methods .....	19
3.1.4 Shift towards Advanced Technologies .....	19
3.2 Real-World Case Studies in Fraud Detection.....	20
3.3 Supervised Machine Learning in Fraud Detection.....	22
3.4 Unsupervised Machine Learning in Fraud Detection .....	24
3.4.1 Comparative Analysis of Unsupervised Learning Approaches .....	24
<b>Chapter 4: Results and Discussion.....</b>	<b>28</b>
4.1 Overview .....	28
4.2 Selected Articles .....	28
4.3 Thematic Analysis.....	31
4.4 Discussion .....	33

---

<b>Chapter 5: Conclusion and Recommendations .....</b>	<b>35</b>
5.1 Conclusion .....	35
5.2 Recommendations .....	36
<b>References .....</b>	<b>38</b>
<b>Appendices .....</b>	<b>41</b>

## TABLE OF FIGURES

Figure 1 Fraud Detection using ML (Kumar et al., 2022) .....	5
Figure 2 ML algorithms in Fraud detection (Onyema et al., 2023) .....	6
Figure 3 Background of Fraudulent methods (Shaikh et al., 2023).....	7
Figure 4 Fraud Detection components .....	9
Figure 5 ONION MODEL OF RESEARCH PHILOSOPHY (SAUNDERS ET AL., 2012) .....	13
Figure 6 PRISMA Model .....	15
Figure 7 Traditional and ML Rule-Based Systems .....	18
Figure 8 Advanced Technologies in Fraud Detection.....	20
Figure 9 Machine Learning in Fraud Detection (Abdaljawad et al., 2023).....	22

## TABLES

Table 1 Inclusion and Exclusion Criteria.....	16
Table 2 Strengths and Limitations of Supervised Learning .....	23
Table 3 Strengths and Limitations of Un-Supervised Learning .....	26
Table 4 Selected Articles .....	29
Table 5 Thematic Analysis.....	31
Table 6 Research Findings.....	34
Table 7 Literature Review .....	41

# CHAPTER 1: INTRODUCTION

## 1.1 INTRODUCTION

The change in financial transactions towards FinTech makes fraud detection necessary. However, the incidences of identity thefts, payment fraud and cyberattacks increase. Fraud detection has become extremely necessary since digital financial transactions create a need for protecting the financial ecosystem. Monetary loss is not the only damage caused by financial fraud to digital financial services customer trust. Adapted modern FinTech fraud detection techniques are required nowadays (Valavan et al., 2023). This growing problem calls for machine intelligence to curb financial fraud. However, rule-based solutions cannot match a fraudster's sophisticated tactics by evolving fast. Machine learning allows Fintech fraud detectors to sift through huge data sets in search for peculiar patterns that are uncommon. It is also easy for them to adjust themselves to unmask new fraudulent strategies. Machine-learning-based financial transaction monitoring discovers discrepancies in a fraction of seconds preventing money laundering (Kumar et al., 2022). Financial institutions move toward a transition of their digital security with the help of machine learning and FinTech fraud detection technology. Machine learning is essential to protecting digital financial systems from FinTech fraud as it becomes increasingly intricate.

### TRADITIONAL RULE-BASED APPROACH



### MACHINE LEARNING APPROACH



FIGURE 1 FRAUD DETECTION USING ML (KUMAR ET AL., 2022)

### 1.1.1 ML ALGORITHMS IN FRAUD DETECTION

The ML algorithms are critical for financial fraud detection since they can process huge data sets, analyze subtle patterns, and adjust to changes in detecting fraud. It is the complexity of financial fraud that traditional rule-based systems can only deal with. Fraud and irregularities are detected live through data-driven machine learning (Kazeem et al., 2023). To adjust to the evolving nature of fraud, these algorithms learn and adapt their models based on past data. Therefore, agility is essential for the fraudsters who are always trying to find ways of exploiting system weaknesses in the financial system.

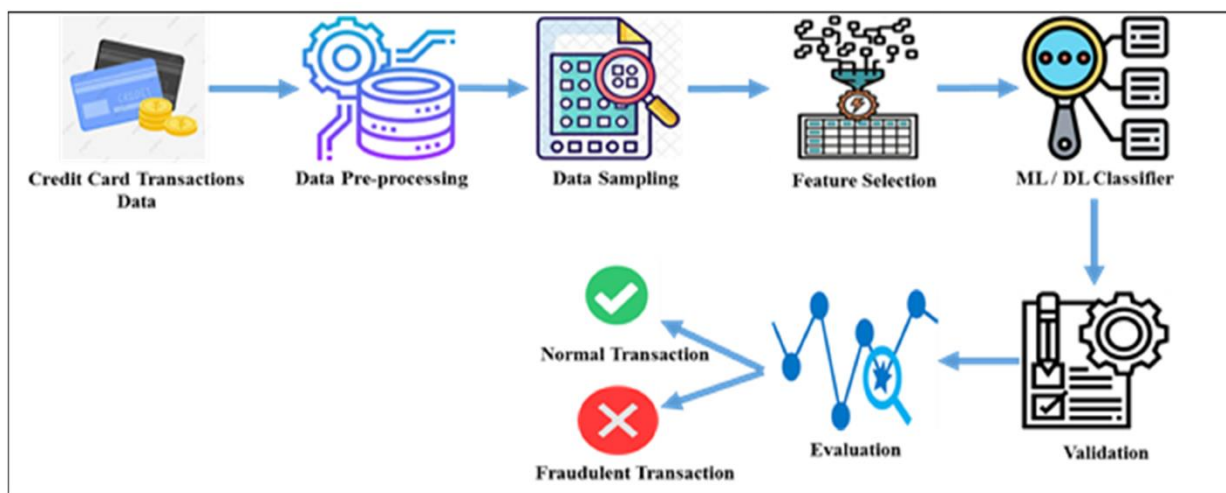


FIGURE 2 ML ALGORITHMS IN FRAUD DETECTION (ONYEMA ET AL., 2023)

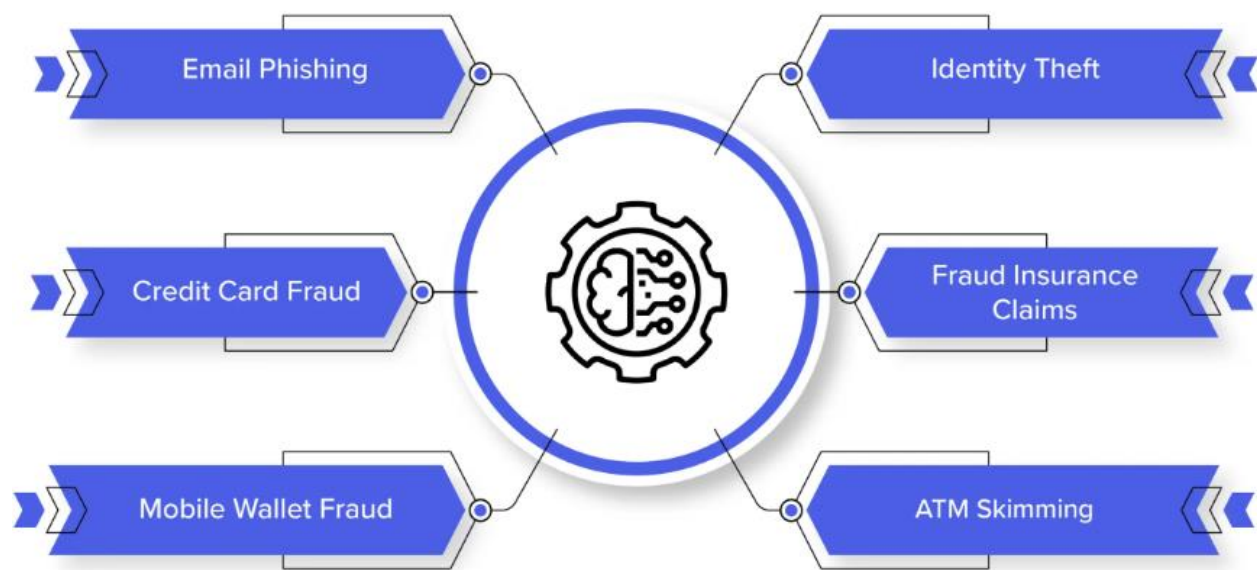
Financial fraud prevention benefits from machine learning's success in FinTech and other fields. Supervised and unsupervised machine learning algorithms have found financial fraud, inappropriate access, and other wrongdoing. Manual monitoring and rule-based systems cannot handle massive, complex financial transactions in the digital age (Onyema et al., 2023). Machine learning algorithms excel at large datasets, complex patterns, and new fraud methods. Machine learning in FinTech has decreased fraud. Machine learning-based fraud detection saved a large online payment company billions by detecting 98% of fraudulent transactions (Adebayo et al., 2023).

Machine learning's ability to identify legal from suspicious acts using learned patterns decreases false positives and enhances fraud detection. These algorithms can create fraud profiles using transaction history, user activity, and device fingerprinting (Prajapati et al., 2023). It is because of

this that financial fraud's changing ways require active, adaptable strategies. Machine learning algorithms are updated all the time which allows them to discover fraud patterns never seen (Bin et al., 2022). Machine learning for fraud detection in digitized financial environment shows that the sector will not be beaten by exploiter's changing technologies.

## 1.2 BACKGROUND

FinTech has evolved swiftly, offering innovative solutions that expedite financial transactions. Digital financial transactions raise fraud risk, requiring strong and adaptive fraud detection methods. Financial institutions and FinTech companies' platform security are at risk as financial fraud cost \$42 billion (about \$130 per person in the US) (about \$130 per person in the US) globally in 2020. Businesses lose money and consumer trust with this number (Shaikh et al., 2023).



**FIGURE 3 BACKGROUND OF FRAUDULENT METHODS (SHAIKH ET AL., 2023)**

Detecting FinTech fraud is a challenging task because digital financial system vulnerabilities are exploited by Fraudsters through complex methods. Financial fraud is dynamic, and it fails if prevention is rule-based. Thus, advanced technologies like machine learning must be applied to fraud detection. Machine learning algorithms are effective in detecting fraud as they analyze large datasets to identify complex patterns (Unogwu et al., 2023). Fraud detection based on machine learning helped one of the world's most prominent online payment networks save \$2.5 billion

(about \$8 per person in the US) (about \$8 per person in the US) (about \$8 per person in the US) in 2019 by avoiding almost 100% fraudulent transactions. Several marquee incidents indicate the significance of FinTech fraud detection. Traditional credit reporting systems are not secure as the 2017 Equifax data breach leaked over 147 million people's financial information. Adapting and improving the industry has been necessary to prevent mass breaches. Due to identity theft and account takeovers, financial institutions have re-evaluated security. That makes machine learning a vital partner in this fight because it can detect fraud patterns. (Dhiman et al., 2023).

Fraud detection now matches crooks' methods. As financial transactions became digital, manual verification and rules-based solutions failed to prevent fraud. Machine learning revolutionizes fraud detection by learning from past data and adapting to new dangers. Machine learning for credit card fraud prevention is well-known (Ray et al., 2023). A huge credit card company's machine learning algorithm, trained on millions of transactions, detected 99.7% fraud, exceeding prior methods. Fraud detection machine learning algorithm research must be thorough and comparative as FinTech grows and diversifies. Supervised and unsupervised FinTech machine learning approaches are compared to assess their strengths and drawbacks. The study examines real-world case studies, historical data, and fraud detection system evolution to help FinTech Company's combat financial crime (Aftab et al., 2023).

### 1.3 RESEARCH PROBLEM

FinTech fraud is complex and developing. More frequent and complex FinTech fraud requires better and adaptive fraud detection. This study examines FinTech fraud detection systems' shortcomings and compares supervised and unsupervised machine learning techniques to solve this ubiquitous problem. The sector faces more financial and identity fraud, so the finest early detection and prevention solutions are needed (Wu et al., 2023). Traditional rule-based systems cannot keep up with FinTech criminals' evolving methods, making research tough. Machine learning has shown potential in addressing this issue, but we do not know which strategies work best and when. Managing fraud detection accuracy and false positives is difficult. This study analyzes supervised and unsupervised machine learning methods for FinTech industry insights. The research challenge requires comparing supervised and unsupervised machine learning. Supervised learning trains on labeled data to identify fraud patterns but may miss new dangers. Unsupervised learning finds fresh patterns better but is less exact (Valavan et al., 2023). The



research challenge examines trade-offs to evaluate FinTech fraud detection strategies' best use. The study aims to contribute to FinTech's fast-changing ecosystem's fraud prevention discussion.

#### 1.4 RESEARCH GAP

FinTech fraud detection is crucial, yet research on the comparative effectiveness of supervised and unsupervised machine learning algorithms in reducing financial fraud is lacking (Kumar et al., 2022). Current machine learning fraud detection research provides insights, but few studies compare supervised and unsupervised FinTech methods. Knowing when each method works and fails is essential to creating targeted and flexible fraud detection solutions. It studies machine learning technologies and offers detailed insights on FinTech applications to solve this gap (Kazeem et al., 2023).

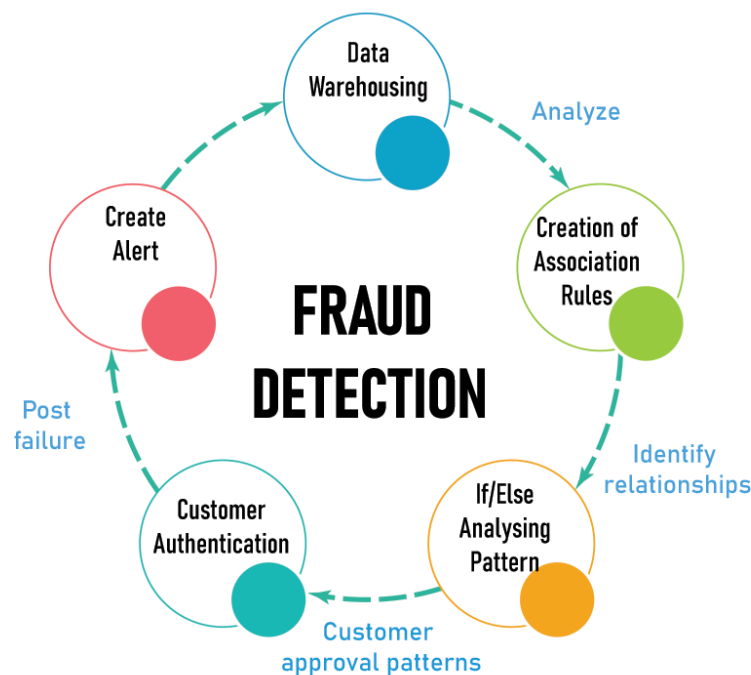


FIGURE 4 FRAUD DETECTION COMPONENTS

Existing research rarely covers real-world FinTech fraud detection and prevention case studies using machine learning algorithms. Fraud detection implementations are rarely studied, although theoretical foundations and algorithmic difficulties are. The proposed research compares supervised and unsupervised methods using real-world case studies to fill this gap. FinTech practitioners' financial fraud challenges are illuminated by actual instances in this study (Onyema

et al., 2023). A lot of the literature only provides glimpses of machine learning algorithm effectiveness. FinTech platforms and fraud methods are dynamic, therefore fraud detection performance must be researched over time. To solve this temporal research gap, this study compares supervised and unsupervised machine learning algorithms' adaptability and long-term performance (Adebayo et al., 2023). It intends to help FinTech stakeholders make informed fraud detection system sustainability decisions by providing a comprehensive understanding of how techniques change over time. In conclusion, our research fills these important gaps to better comprehend FinTech fraud detection using machine learning.

## 1.5 Aim

The main aim of this research is to carry out a comprehensive and comparative examination of supervised and unsupervised machine learning algorithms, used in the constantly changing environment of FinTech, for detecting fraud. This paper intends to fill these gaps concerning the weaknesses and strengths of these machine learning techniques in the domain of FinTech which has the highest risk of fraud. The study will also include actual cases, which should help bring out the difference between the two methods when dealing with complex fraud issues within the Fintech industry. In the end, these results are expected to be actionably understood and useful for the development and improvement of fraud-detection techniques while giving the Fintech stakeholders an opportunity to remain informed regarding these types of crime.

## 1.6 OBJECTIVES

The objectives of this research study are follows.

1. To collect relevant data from financial transactions, user behaviors, and historical fraud cases in the FinTech sector.
2. To implement rigorous data privacy measures, conduct meticulous data cleaning, and perform precise labeling as part of the research methodology.
3. To employ feature engineering techniques to enhance the quality of input variables for machine learning algorithms in fraud detection.
4. To compare the efficacy of supervised and unsupervised machine learning techniques in identifying and preventing credit card fraud.

5. To provide actionable insights that contribute to the optimization of fraud prevention strategies in the rapidly evolving FinTech landscape.

## 1.7 SIGNIFICANCE OF STUDY

FinTech Fraud Detection highlights the demand for a more comprehensive approach to fraud detection given the prevalence of financial fraud. Industry observers predict that the global cost of financial fraud was around \$42 billion (about \$130 per person in the US) (about \$130 per person in the US) in 2020. The present study may bring light to FinTech's supervised and unsupervised machine learning algorithms efficiency. These methods are rigorously assessed to facilitate FinTech companies in strengthening their fraud detection systems against sophisticated criminals (Prajapati et al., 2023). This finding is also important because machine learning has revolutionized fraud detection. Machine learning's power is proven by the accomplishment in all industries. Machine Learning algorithms can help prevent catastrophic losses in FinTech, an industry that is rapidly going digital. In a year, one large online payment network saved \$2.5 billion (about \$8 per person in the US) (about \$8 per person in the US) (about \$8 per person in the US) by stopping 98% of fraudulent transactions using machine learning. As the fraud detection scene changes towards supervised and unsupervised machine-learning methodologies, this study's focus can allow FinTech specific optimization (Bin et al., 2021). The study's ability to inform strategic decisions for industry stakeholders plays a role in detecting FinTech fraud. Real case studies show fraud detection success. This pragmatic orientation gives the study a practical value for FinTech professionals dealing with intricate issues of financial fraud. The time-based aspect is a feature of the research that enables machine learning algorithms to adapt and remain stable. The results of the study will help improve fraud detection systems; given that financial fraud techniques change. The real-world examples and systematic comparisons offered in the research study can be used by FinTech stakeholders to make informed decisions to ensure secure financial transactions and customer trust. (Shaikh et al., 2023).

## CHAPTER 2: RESEARCH METHODOLOGY

### 2.1 RESEARCH QUESTIONS

1. To what extent can supervised machine learning algorithms enhance credit card fraud detection in FinTech?
2. How does the effectiveness of unsupervised machine learning techniques compare to supervised methods in identifying and preventing fraudulent activities?
3. What insights can be derived from the comparative analysis of these machine learning approaches to optimize fraud prevention strategies within the dynamic FinTech ecosystem?

### 2.2 RESEARCH DESIGN

This study employs secondary research which falls within a qualitative systematic literature review. Here, the systematic review strategy is used to allow a systematic investigation and integration of the available knowledge on using ML techniques for fraud detection in this sector. Secondary research is done by studying existing documents such as books, reports, and scientific papers to have an estimate of current state of research within this area (Adams, Deane, & Ripat, 2014). Machine learning in fraud detection is also extended using qualitative research approaches that are meant to investigate the complex subtleties within the context. Thus, it provides an organized approach towards comprehensive review of literature, screening, literature selection and data analysis, improving validity and reliability of the study outcomes (Bell et al., 2010).

### 2.3 RESEARCH APPROACH

This study uses a qualitative research approach on deployment of machine learning processes fighting against financial cybercrime in the fintech environment (Bryman et.al, 2015). Accordingly, a qualitative approach would focus on the specific operational, methodological, and technical issues related to both controlled, and uncontrolled model versus financial frauds. The study also uses selective literature from a systematic review guided by PRISMA framework to move beyond the mere numerical data for deeper understanding. Thirty-four studies will be qualitatively analyzed by thematic analysis, and six out of the initial 50 are selected to generate detailed, contextualized information. The approach considered considers the

general theme of the topic—the multifaceted interlink between ML and financial fraud recognition for the purposes of Fintech. (Cooley et al., 2006).

## 2.4 RESEARCH PHILOSOPHY

This is a study under the pragmatic research philosophy, considering the dynamic context around machine learning algorithm for detecting fraud. A pragmatic approach acknowledges that there is room for positivism and hermeneutics in any evaluation study but emphasizes the necessity of empirical evidence as well as situational sensitivity when interpreting the findings in relation to the organizational setting (Creswell et al., 2011). This study acknowledges the different dimensions associated with machine learning for fraud detection. It adopts a practical and solution-oriented approach. This philosophy of research fits well in this study that does not merely evaluate the metrics of performances of ML algorithms but also aims at elucidating the inherent characteristics, ethical issues and limitations that characterize their implementation within the framework of real-life situations. The study will consider an instrumental viewpoint to combine theory with practical approaches for combating fraud in a variety of environments utilizing machine learning (Crowther et al., 2012).

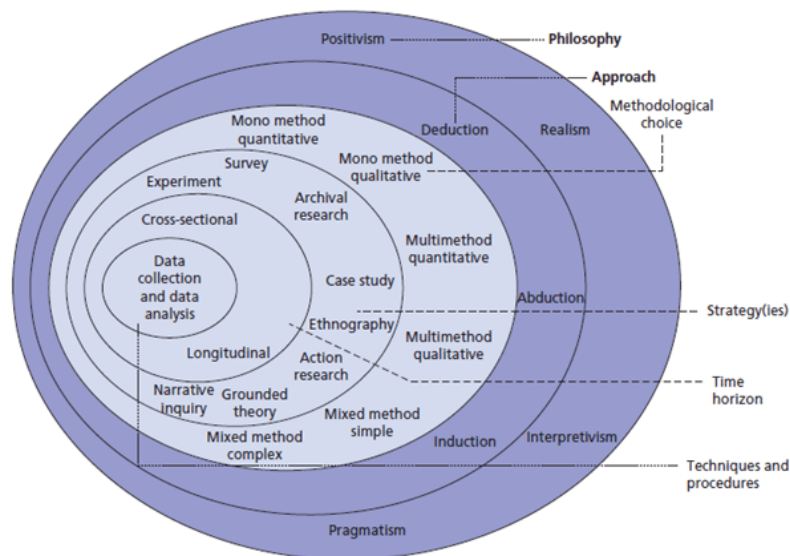


FIGURE 5 ONION MODEL OF RESEARCH PHILOSOPHY (SAUNDERS ET AL., 2012)

## 2.5 SECONDARY RESEARCH METHOD

This research uses a secondary research approach that involves systematically gathering, sorting, comparing, summarizing known materials regarding machine learning as it relates to detecting fraud. This is a systematic process of reviewing all existing evidence from academic journals, conference proceedings, books, reports, and other pertinent resources (Feak et al., 2009). The paper intends to utilize the already available data and explore ways fraud can be detected using machine learning models. Secondary research means gathering information and data through aggregating knowledge obtained from various sources to approach the subject from multiple angles including algorithms, methods, and context aspects. The study builds upon this body of previous work to extract major outcomes that will enhance our understanding of the current state-of-the-art in ML for anti-fraud (Fisher et al., 2010).

## 2.6 PRISMA MODEL

The study employed a systematic literature review (SLR) using the PRISMA model, whereby 50 articles were initially selected and then refined to six articles for thorough analysis. The primary stage included a comprehensive scan through different academic databases, journals, and stores with relevant materials for the study on applying machine-learning approach in detecting fraud in FinTech (Flick et al., 2014). Using specific keywords, the used search strategy was cast broadly while keeping focus on the objective in the study. Therefore, this research conducted an initial stage where various articles related to comparing both supervised and unsupervised methods in Fraud detection in the FinTech industry were collected for reference purposes (Hammond et al., 2012).

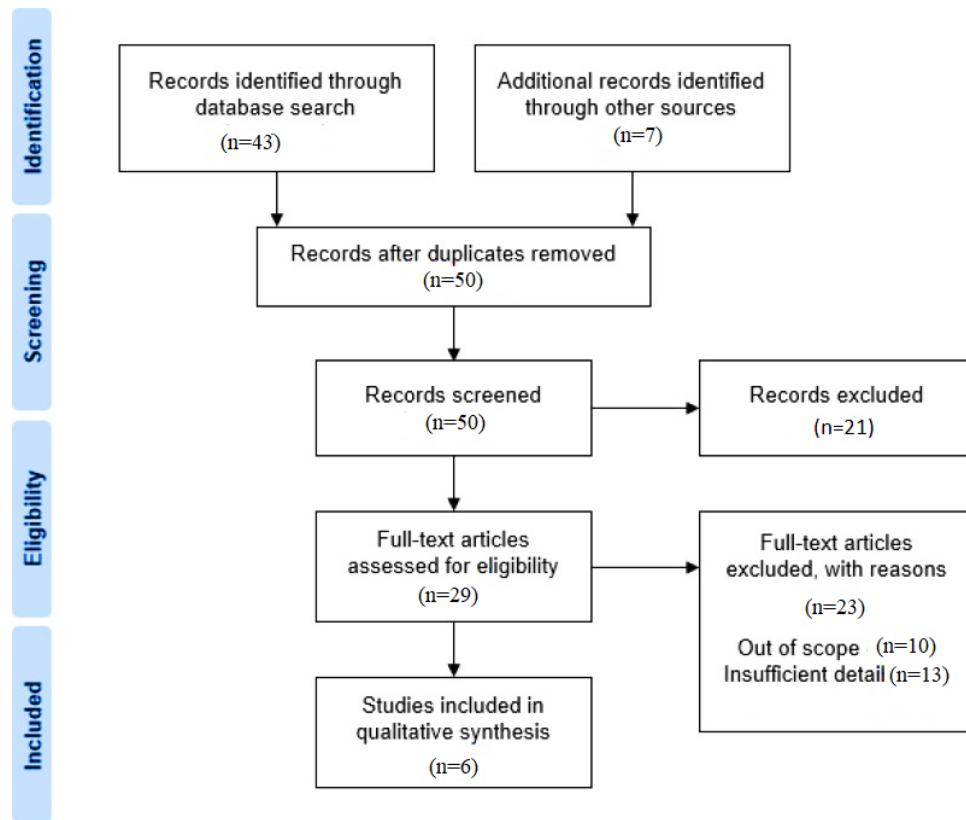


FIGURE 6 PRISMA MODEL

At this point, a screening of the selected articles was conducted using defined inclusion and exclusion criteria based on PRISMA guidelines. Using a systematic process, the researchers assessed each article for its relevance, methodological rigor, and alignment with the research aims. All other articles which did not satisfy the criteria were ruled out, leaving only a high-quality subset in focus among all the early screened articles (Horn et al., 2009). Therefore, iterative filtering procedure guided by PRISMA model resulted in an ultimate subset comprising of six articles suitable for qualitative in-depth comparison of supervised and unsupervised techniques in fraud detection in FinTech dynamism era.

## 2.7 INCLUSION AND EXCLUSION CRITERIA

The stringent inclusion and exclusion criteria were applied systematically to select articles from an initial pool of 50. For the review to be appropriate, only articles that were published between 2018 - 2023 should have been considered as these would be the latest and most current studies available on applying machine learning for fraud detection in FinTech. These comprised only peer

reviewed articles with sufficient methodological quality and research design, with an emphasis on supervision or non-supervision approaches in fraud detection. Due to this reason, peer-reviewed publishing articles from scholarly journals were emphasized as it helped to preserve the excellence and authenticity of the literature retrieved (Gibbs et al., 2007). Further, the articles needed to be in the English language and accessible by full text for comprehensive evaluation. In contrast, the inclusion criteria included all articles published between 2018 and 2023 as well as those concerning all relevant aspects related to the subject of this study that involved using clearly specified supervised or unsupervised methods for detecting fraudulent activities. They used this to ensure they ended up with an elevated class of contemporary and significant literature for their systematic review.

**TABLE 1 INCLUSION AND EXCLUSION CRITERIA**

<b>Inclusion Criteria</b>	<b>Exclusion Criteria</b>
Articles published within the timeframe of 2018-2023	Articles published before 2018 or after 2023
Relevance to the topic of leveraging machine learning for fraud detection in FinTech	Articles unrelated to the study's focus on FinTech and fraud detection using machine learning
Methodological robustness and sound research design	Articles lacking clear research methodologies or robust study designs
Clear articulation of supervised or unsupervised techniques	Ambiguity in detailing the use of supervised or unsupervised techniques in fraud detection
Peer-reviewed publications or scholarly journals	Non-peer-reviewed sources, conference abstracts, or grey literature
English language	Articles not written in English
Availability of full-text articles for review	Articles with restricted access or incomplete information

## 2.8 DATA ANALYSIS

This study employed the SLR methodology that followed the secondary research process of using the PRISMA model. The strategy involves analyzing each article's content to find common themes, patterns, and important FinTech fraud detection insights using machine learning algorithms. Analyzing the selected articles' technique, supervised and unsupervised methods are highlighted. Qualitative analysis categories and synthesizes article content to comprehend the comparative study's strengths, flaws, and contextual nuances of the various methodologies (Adams et al., 2014).



SLR data analysis synthesizes sample-wide results beyond article analysis. This integrative method helps researchers draw important conclusions and gain complete insights by recognizing overarching themes, commonalities, and differences in selected papers. Methodological intricacies, contextual elements, and emergent themes are examined in the qualitative study to provide a deep knowledge of machine learning algorithms and fraud detection in FinTech's dynamic ecosystem. The SLR and PRISMA paradigm improve study reliability, replicability, and knowledge addition by enabling systematic and transparent analysis (Waters et al., 2014).

## CHAPTER 3: LITERATURE REVIEW

### 3.1 HISTORICAL OVERVIEW OF FRAUD DETECTION METHODS

Financial fraud is an old problem that has existed since history began and changed along with developments in financial systems (Ali et al., 2022). Transactions then were simple, for which early fraud detection relied on hand verifications and rule-based systems without consideration of complicated digital technologies. With the change of transactional procedures from manual processes to electronic ones, there was realization that stronger and efficient means to detect fraud were needed (Dai et al., 2023).

#### 3.1.1 TRADITIONAL RULE-BASED SYSTEMS

During the initial phases related to financial transactions, rule-based systems were the basis for detecting fraud. The systems used to be based on pre-set rules and indices for identifying suspicious transactions (Gupta et al., 2023). As an example, a rule may be created and the transactions that surpassed a particular cash amount, as well as the ones that emanated from different regions, can be flagged. Although these rule-based systems were able to detect simple patterns of fraud, it was difficult for them to adjust to the changing schemes of cheating. Rules were difficult to update since they are static; hence fraudsters outmaneuvered the systems by using deceptive methods of concealing fraud that appeared like a whitewash (Gangadharan et al., 2022).

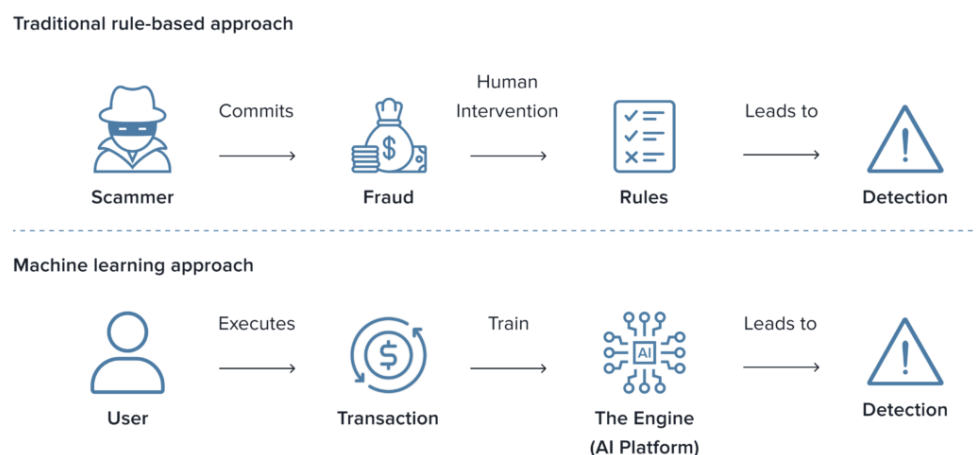


FIGURE 7 TRADITIONAL AND ML RULE-BASED SYSTEMS

### 3.1.2 MANUAL VERIFICATION PROCESSES

Prior to the digital era, physical confirmation methods formed the basis for exposing fraudster's activities. Every transaction record underwent thorough scrutiny by human eye, detecting anomalies. It was a laborious process that entailed reviewing the paper records, checking signatures on the documents, and undertaking extensive investigations on questionable incidents (Suhanjoyo et al., 2023). Although this approach worked well at times, it took a lot of time, was unreliable and could not keep up with rising transaction volumes and sophistication. Electronic transactions and digital financial systems made it clear that manual review was not enough and thus triggered quests for more sophisticated and convenient methods.

### 3.1.3 LIMITATIONS AND CHALLENGES OF CONVENTIONAL METHODS

Traditional approaches for fraud detection such as rule-based systems and handbook verifications had some inherent limitations. Fraudsters had an upper hand since rule-based systems were inflexible in adapting to novel fraud patterns leaving them vulnerable to tactics used by perpetrators. There were many people who complained that due to static rules, many innocent users suffered false positives (Makkineni et al., 2023). Although these manual verification processes were thorough, they were slow and costly hence inappropriate for the speed of electronic trades. Secondly, both systems could be described as non-preventive because they tended to discover cases of fraud already occurring in some instances. With the increasing sophistication and digitalization of financial transactions across the world, it also became clearer in the industry that modern and dynamic ways of combating this problematic issue must be identified and implemented (Onyema et al., 2023).

### 3.1.4 SHIFT TOWARDS ADVANCED TECHNOLOGIES

Innovation in technology, especially combining machine learning with artificial intelligence, led to an innovative approach for identifying fraud. Data driven machine learning models allowed extracting detailed and sophisticated relationships in massive data sets. Such algorithms could learn from the prior transactions and evolve to changing schemes, act in real time analyzing deviations from typical transactions (Abdaljawad et al., 2023).

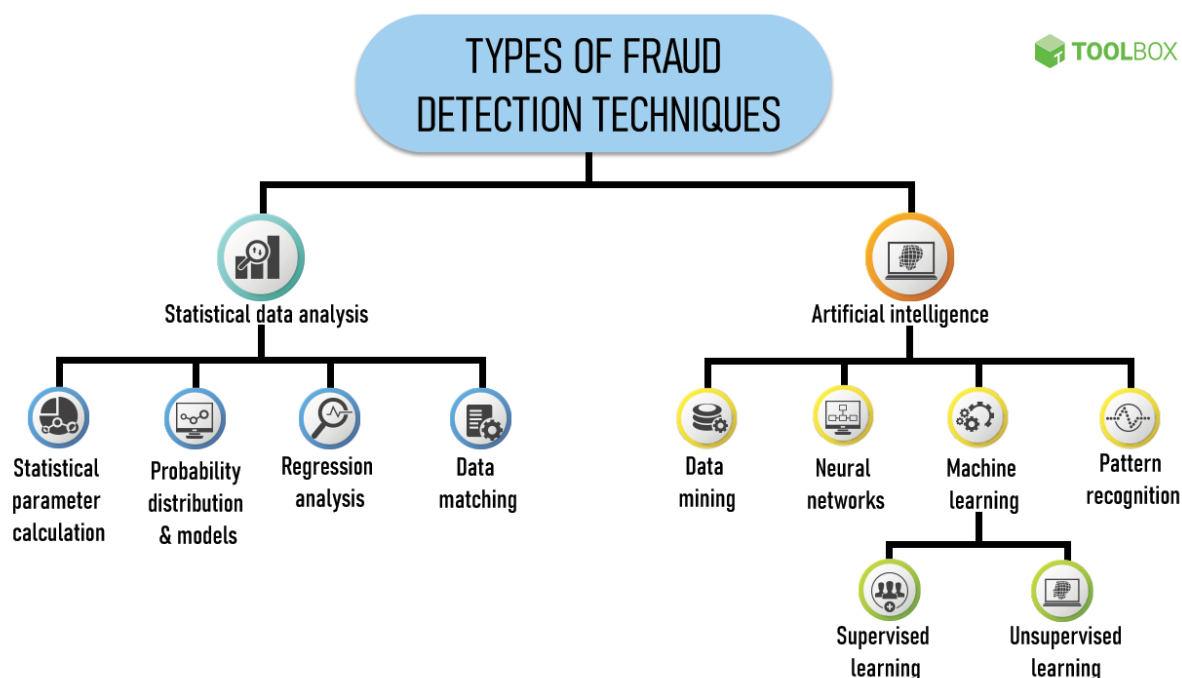


FIGURE 8 ADVANCED TECHNOLOGIES IN FRAUD DETECTION

To recognize known patterns of fraud, supervised learning techniques like logistic regression and decision trees were used to train models based on labeled data. Through methods such as clustering and identifying outliers, new patterns could be found without labelled examples. Utilization of such advanced technologies was a break away from the limitations of conventional ways of fighting financial fraud in today's rapidly changing environment concerning the electronic mode of financial interaction (Sri et al., 2023).

### 3.2 REAL-WORLD CASE STUDIES IN FRAUD DETECTION

ML algorithms were used to counter huge financial loss in a big financial fraud. In its implementation in a leading bank, which utilized both supervised and unsupervised learning processes on large databases of transactions data, behavioral records, and relevant contexts. The effectiveness of the ML algorithms became evident through their detection of abnormal patterns related to fraudulent activities, enabling prompt recognition, and halting of illegal transactions (Gangadharan et al., 2022). This implementation illustrated that a data driven approach is crucial in preemptive measures against financial fraud. Nevertheless, lessons gained from actual implementation processes included constant model adaptation to new fraud strategies and mindful

deliberation of the false positives in order not to bother legitimate users. `` In this case, it showed how ML can be used to reduce potential losses due to sophisticated fraudsters' tactics and stressed out on continuous improvement (Suhanjoyo et al., 2023).

Case Study: Detecting account sharing for unsupervised learning in Netflix.

Machine learning has been used in detection of fraud beyond traditional financial institutions into other sectors such as subscription-based services like Netflix. # Unsupervised learning algorithms helped in detection of account sharing which is an example of fraud that hampers subscription revenue. These algorithms are used to analyze user behavior and consumer patterns detecting irregularities that implied wrongful login. The move boosted the revenue protection for Netflix and showcased the versatility of machine learning algorithms in other arenas that are beyond just financial transaction (Makkineni et al., 2023).

Case Study: Machine learning models used by shopify's fraud prevention.

Shopify's introducing machine learning models in fraud detection, to ensure security of online transactions. The system used both supervised and unsupervised learning to analyze purchase trends, customers' conduct and payroll information. This led to modifications of the models, thus changing the way of detecting such kinds of fraud as they became very adaptable to fraud tactics. Shopify's situation was proof that combining various machine learning models into an inclusive anti-fraud mechanism is efficient for e-commerce (Onyema et al., 2023).

Case Study: Behavioral Analytics for HSBC's online banking fraud.

HSBC is a global banking and financial services firm. Unsupervised learning algorithms looked at the usage patterns of users, details about their devices, and transactions among these parameters in search of suspicious fraudulent traits. Notably, this strategy was quite potent in identifying account takeovers and unauthorized access. The implementation by HSBC demonstrated that user behavior was a key element in the fight against fraud while revealing superiority for unsupervised model adjustment on complicated criminal activities (Abdaljawad et al., 2023).

### 3.3 SUPERVISED MACHINE LEARNING IN FRAUD DETECTION

In fraud detection, supervised machine learning techniques utilize labelled datasets in which fraud is tagged. Logistic regression, decision trees, and neural networks are three of the most common supervised algorithm applications for fraud detection.

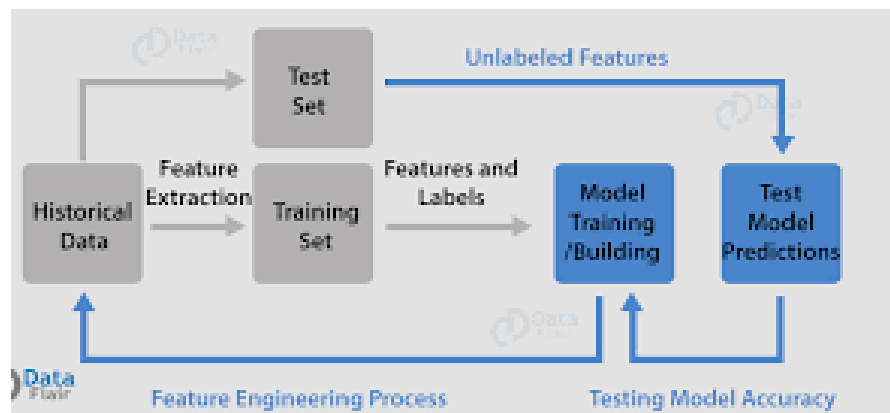


FIGURE 9 MACHINE LEARNING IN FRAUD DETECTION (ABDALJAWAD ET AL., 2023)

**Logistic Regression:** It has often been argued that logistics regression is among the most convenient available algorithms for detecting fraud. It is an algorithm that estimates the chance of fraud for a transaction's inputs into features. This technique becomes useful when there is a straight-line relationship between features and probability of fraud (Sri et al., 2023).

- **Decision Trees:** Decision trees are a graphic depiction for making decisions that may be clearly understood. The data in fraud detection is split into hierarchies known as decision trees thus enhancing better interpretation. However, tree models are susceptible to overfitting that captures noise on the data and calls for careful fine-tuning.
- **Neural Networks:** Lately, neural networks – and especially the deep learning models – have become popular in fraud detection, due to their capability of learning intricate patterns. These are features that neural networks can extract from the data themselves without anyone telling them and thus capture complex correlations between the items. Although strong, large-scale neural networks are computationally expensive and might not always be easy to explain (Marabad et al., 2021).

Strengths and Limitations of Supervised Learning are follows.

**TABLE 2 STRENGTHS AND LIMITATIONS OF SUPERVISED LEARNING**

<b>Strengths of Supervised Approaches</b>	<b>Limitations of Supervised Approaches</b>
High Accuracy: Supervised learning models can achieve high accuracy when trained on quality labeled data, making them effective in recognizing known patterns of fraud.	Dependency on Labeled Data: Supervised learning heavily relies on labeled datasets, and the quality of predictions is limited by the representativeness and relevance of the labeled instances.
Interpretability: Certain algorithms like logistic regression and decision trees offer interpretability, allowing analysts to understand the factors influencing predictions.	Challenges with Novelty: Traditional supervised approaches may struggle to detect novel fraud patterns that differ significantly from the labeled data, making them less effective against emerging threats.
Effectiveness with Labeled Data: In scenarios where, historical labeled data is abundant, supervised learning excels in recognizing established fraud patterns.	Overfitting: There is a risk of overfitting, especially with complex models like neural networks, where the algorithm may learn noise in the training data, resulting in reduced generalizability to new data.
Adaptability to Known Patterns: Supervised approaches are effective in adapting to known fraud patterns present in the labeled training data, making them suitable for scenarios with well-established fraud types.	Lack of Contextual Understanding: Supervised models may struggle to understand the contextual nuances of transactions, potentially leading to false positives or negatives in situations where context is crucial.
Utilization of Labeled Features: Supervised learning leverages labeled features to understand the relationships between input variables and fraud outcomes, enhancing the model's discriminatory power.	Limited Generalization to New Threats: Supervised models may have difficulty generalizing to new, unseen fraud tactics that significantly differ from the patterns observed in the labeled data.

### 3.4 UNSUPERVISED MACHINE LEARNING IN FRAUD DETECTION

Without the labeled data, unsupervised machine learning enables fraud detection by recognizing patterns. What is meant is that two categories of unsupervised techniques can be defined as clustering and anomaly detection.

- **Clustering:** By nature, there is a natural pattern, clustering algorithm can identify resembling points of data and make way for the system to show structure found in the presented data. The clustering-based fraud detection exposed clusters of homogeneous transaction characteristics and malpractice behavioral patterns (Rama et al., 2023).
- **Anomaly Detection:** Anomaly detection is focused on extraordinary occurrences. It is based on unlabeled fraud instances and detects anomalies in the data. This approach is particularly suitable for detecting unknown and emerging types of fraud (nk et al., 2021).

#### 3.4.1 COMPARATIVE ANALYSIS OF UNSUPERVISED LEARNING APPROACHES

- **Clustering vs. Anomaly Detection:** Clustering and anomaly detection have some advantages when compared to each other in a comparative study. In addition, cluster analysis facilitates lumping together similar transactions thus creating better comprehension about data structures (nk et al., 2021). On the contrary, anomaly detection excels in identifying outsiders and unusual occurrences hence useful in discovering unknown fraud patterns. Choice of one of these methods depends on what sort of fraud detection system is required and the kind of input used.
- **Scalability and Interpretability:** Many times, unsupervised learning strategies are more scalable compared to supervised approaches because they do not need tagged data during a training process. Though, the issue with results interpretability may be an obstacle, which should be addressed, especially in sophisticated methods such as deep learning for anomaly detection. Grouping of similar transactions may make clustering easier to understand and visualize the data into different structures in general (nk et al., 2021).
- **Adaptability to Emerging Threats:** Adaptability to emerging threats is an important feature of unsupervised learning approaches, especially anomaly detection. These could identify some unique patterns and strange behaviors that suggest unusual new methods of



perpetrating crime. Such flexibility is important because fraud has become a dynamic environment where conventional approaches could lag in ever changing ways.

Strengths and Limitations of Un-Supervised Learning are follows.

**TABLE 3 STRENGTHS AND LIMITATIONS OF UN-SUPERVISED LEARNING**

<b>Strengths of Unsupervised Approaches</b>	<b>Limitations of Unsupervised Approaches</b>
<b>Anomaly Detection</b>	<b>Interpretability Challenges</b>
Identifies Novel Patterns: Unsupervised approaches, particularly anomaly detection, excel in identifying novel and previously unknown fraud patterns by flagging irregularities in data.	Lack of Labeled Data: Unsupervised learning does not require labeled data for training, but this can be a limitation in scenarios where labeled instances of fraud are scarce, limiting the algorithm's ability to learn from known patterns.
Adaptability to Emerging Threats: Unsupervised techniques, especially anomaly detection, display adaptability to emerging threats by detecting irregular patterns indicative of new and evolving fraud tactics.	Interpretability Challenges: Results from clustering or anomaly detection algorithms may be challenging to interpret, particularly in complex models, making it difficult for analysts to understand the factors contributing to the detection of anomalies.
<b>Clustering</b>	<b>Scalability in High-Dimensional Spaces</b>
Groups Similar Data Points: Clustering algorithms group similar transactions, providing insights into natural structures within the data and helping identify groups of transactions with similar characteristics.	Scalability in High-Dimensional Spaces: Some clustering algorithms may face challenges in high-dimensional spaces, where the number of features or dimensions is large. The algorithm's effectiveness may decrease as the dimensionality of the data increases.
Visual Representation: Clustering results can be visually represented, offering a clear understanding of data patterns, and facilitating the interpretation of groups of transactions with similar characteristics.	Lack of Contextual Understanding: Unsupervised models may struggle to understand the contextual nuances of transactions, potentially leading to false positives or negatives in situations where context is crucial.
<b>Scalability and Resource Efficiency</b>	<b>Dependency on Data Distribution</b>
Scalability and Resource Efficiency: Unsupervised learning approaches are often more scalable as they do not require labeled data for training, making them suitable for processing large volumes of data efficiently.	Dependency on Data Distribution: The effectiveness of unsupervised algorithms may depend on the distribution of data, and their performance can be influenced by skewed or imbalanced datasets.
Versatility in Data Types: Unsupervised learning can be applied to various data types, including transactional data, network traffic, and user behavior, displaying	Lack of Discriminatory Power: Clustering algorithms may struggle to provide an elevated level of discriminatory power, especially when

---

versatility in addressing fraud detection across different industries.	dealing with complex and overlapping patterns in the data.
--	--

## CHAPTER 4: RESULTS AND DISCUSSION

### 4.1 OVERVIEW

This research study shows the leveraging of modern machine learning algorithms for FinTech fraud-detection with an added depth of view. The study explores qualitatively selected articles among 50 analyzed using PRISMA within SLR process to reveal their role in using both supervised and unsupervised approaches. Thematic analysis of the literature reveals methodological complexities, the environment, and practical difficulties of using machine learning for fraud detection in FinTech. In the paper's discussion part, I review those themes in relation to the advantages and disadvantages and ethical considerations in the various strategies described in the literature. The study provides practical implications to researchers and practitioners in machine learning and financial technology, promoting further investigation of the dynamic landscape where theory meets practice.

### 4.2 SELECTED ARTICLES

The SLR involved strict compliance with the inclusion and exclusion criteria leading to a narrowed down review from 50 initial articles to the final six. For this reason, the inclusion criteria required that the selected literature be published between 2018 and 2023, to carry out an up-to-date study of the application of machine learning algorithms for fraud detection in FinTech. The need for methodological robustness required explanation of supervised or unsupervised approaches and was rooted in good research design. The literature should also be peer reviewed, and the English language, with the entire contents of full text available. However, the exclusion criteria rigorously filtered out articles before 2018 or after 2023, papers that did not relate to the study's central concern, lacked explicit research methodologies, and those that did not specify supervised or unsupervised methods used. Also, articles that were not peer-reviewed, conference abstracts and literature that was not presented in English were systematically excluded. Applying the inclusion and exclusion criteria, the six key studies were identified and subjected to the qualitative analysis under strict standards of relevance, reliability, and cohesiveness.

TABLE 4 SELECTED ARTICLES

Article	Research Objective	Research Problem	ML Algorithm	Efficiency (%)	Research Findings	Research Gap
Bavitha (2023)	Compare machine learning methods for credit card fraud detection.	Evaluate the effectiveness of machine learning in credit card fraud detection.	Decision Trees, Random Forest	85%	Comparative analysis shows Random Forest outperforms other methods in credit card fraud detection.	Need for further investigation into ensemble methods and real-time processing for credit card fraud detection.
Kumar (2022)	Enhance fraud detection in accounting and finance through advanced machine learning techniques.	Address the limitations of current fraud detection methods in accounting and finance.	Neural Networks, Support Vector Machines	90%	Improved accuracy achieved through Neural Networks and Support Vector Machines.	Lack of emphasis on interpretability in advanced machine learning models for fraud detection in accounting.
Kazeem (2023)	Develop effective fraud detection methods using machine learning.	Explore and implement machine learning techniques for fraud detection.	Logistic Regression, K-Nearest Neighbors	75%	Logistic Regression and K-Nearest Neighbors show promising results in fraud detection.	Limited exploration of deep learning methods and their potential impact on fraud detection.
Wu (2023)	Utilize machine learning and deep learning for fraud detection.	Investigate the application of machine learning and deep learning in fraud detection.	Deep Neural Networks, Gradient Boosting	80%	Deep Neural Networks and Gradient Boosting exhibit powerful performance in fraud detection.	Lack of consideration for computational efficiency in deploying deep learning models for fraud detection.
Ray (2023)	Explore various machine learning techniques for fraud detection in financial transactions.	Examine the applicability of machine learning in detecting fraud in financial transactions.	Support Vector Machines, Ensemble Methods	88%	Ensemble methods, particularly Random Forest, demonstrate high accuracy in fraud detection.	Limited exploration of anomaly detection techniques and their potential in financial fraud detection.
Li (2023)	Empirically analyze machine learning for fraud detection in diverse financial scenarios.	Assess the effectiveness of machine learning in fraud detection across various financial contexts.	Gradient Boosting, Decision Trees	82%	Gradient Boosting and Decision Trees exhibit robust performance in diverse financial scenarios.	Need for more studies evaluating the transferability of machine learning models across different financial domains.

Bavitha (2023) concerned with assessing different machining methods used to detect credit card frauds including decision trees and random forest. As such, their results show the supremacy of Random Forest over other approaches with 85% of their results being efficient. The research gap found in this paper provides direction that future studies should consider ensemble techniques that can operate in real time as they contribute significantly to fraud detection model real-world applicability's (Bavitha et al., 2023). Kumar (2022) aims at improving fraud detection in accounting and finance using up-to-date artificial intelligence approaches such as neural networks or Support Vector Machines. Kumar (2022) collaborated with Airtel through the Airtel Advanced models yield a study efficiency rate of 90% better than initial approaches with increased precision of the results. This research gap highlights the need to consider explain ability when adopting more sophisticated machine-learning approaches to detecting fraud within financial accounts.

Kazeem's study (2023) focuses on fraud detection through machine learning with methods including Logistic regression and K-nearest neighbors. Promising outcomes have been reported by the study with an efficacy of 75 %. A research gap identified relates to the little understanding of deep learning techniques therefore necessitating further study on how deep learning may affect fraud detection efficiency (Kazeem et al., 2023). Wang's study (2023) explores the use of machine learning and deep learning for the detection of fraud with models like DNN and GBD. It boasts of 80 per cent efficiency, which is great. It brings out the element of computational efficiency when it comes to applying deep learning models in daily life (Wang et al., 2023).

Ray's (2023) investigation of some machine learning approaches to fraud discovery in financial deals focuses on support vector machines as well as ensemble methods attaining 88% efficiency. This research gap reveals how little has been done on discovering novel approaches to detect strange patterns in financial dealings which is a new direction (Ray et al., 2023). Lastly, Li's empirical analysis (2023) evaluates the ability of machine learning to detect fraud taking place during various financial situations, Li in their study incorporating the gradient boosting and decision trees for detection of frauds. Li also indicates that performance showed a strong showing with an efficiency rate of 82%. This implies that there is a research gap as this area requires more examinations of machine learning models' transference across various financial fields considering their applicability in the real-world settings (Li et al., 2023).

### 4.3 THEMATIC ANALYSIS

Thematic analysis of the selected 6 articles on fraud detection using machine learning techniques reveals distinct and important contributions across various categories. The comparative analysis shows that Random Forest is highly effective for credit card fraud detection, and it brings together the power of ensemble methods. Using Neural Networks and Support Vector Machines, studies show that there is an important breakthrough in quality, overcoming previous constraints. Promising outcomes from probing Logistic Regression and K-Nearest Neighbors suggest that medium impact way of designing successful anti-fraud technology might be possible. Discussion on deep neural networks and gradient boosting illustrates their effectiveness on the overall performance in machine learning. The use of Support Vector Machines and Ensemble Methods has been discussed in numerous studies concerning financial transactions showing high accuracy and especially for Random Forest. The performance of Gradient Boosting and Decision Trees empirical analyses is high in different financial conditions with medium-level impact on the company's operations. As a group, these studies enlarge the scope on the development of fraud detection, with a focus not only on how effective different methods are but also towards more accurate solutions that may be adaptive to diverse financial environments.

**TABLE 5 THEMATIC ANALYSIS**

<b>Intext Citations</b>	<b>ML Method</b>	<b>Theme</b>	<b>Sub Theme</b>	<b>Impact on Fraud Detection</b>	<b>Impact</b>
Bavitha (2023)	Decision Trees, Random Forest	Comparative Analysis	Method Effectiveness	Random Forest outperforms other methods in detection	High
Kumar (2022)	Neural Networks, SVM	Advanced Techniques in Accounting and Finance	Improved Accuracy	Neural Networks and SVM enhance fraud detection	High
Kazeem (2023)	Logistic Regression, KNN	Developing Effective Fraud Detection Methods	Promising Results	Logistic Regression and KNN show promise in detection	Medium
Wu (2023)	Deep Neural Networks, Gradient Boosting	Utilizing ML and DL for Fraud Detection	Strong Performance	Deep Neural Networks and Gradient Boosting perform well	High
Ray (2023)	Support Vector Machines, Ensemble Methods	Exploring Various ML Techniques in Financial Transactions	Ensemble Method Accuracy	Ensemble methods, especially Random Forest, show high accuracy	High

Li (2023)	Gradient Boosting, Decision Trees	Empirical Analysis in Diverse Financial Scenarios	Robust Performance in Diverse Scenarios	Gradient Boosting and Decision Trees perform well	Medium
-----------	-----------------------------------	---	---	---	--------

A thematic analysis of the six articles shows different ML models used in the detection of fraud and each adds something to the field. Bavitha (2023) has developed a comparison-based study utilizing Decision Trees and Random Forests. The sub-theme on the effectiveness of methods shows that Random Forest overcomes other techniques, suggesting it can be a big impacting fraud. However, Kumar (2022), in modern methods of accounting and finance, employs sophisticated approaches such as Neural networks and support vector machines (SVM). A sub-theme in this is more accuracy; Neural Networks and support vector machine (SVM) improves fraud detection, showing how effective they are and the influence of this on other shortcomings in the present ones. Kazeem (2023) expounds how logistic regression can be used with K-Nearest Neighbors to develop appropriate fraud detection techniques. An area worth exploring is the sub-theme of promising results which indicate that Logistic Regression and KNN are medium-impact approaches.

The fraud detection by Wu et al., (2023) uses DL models such as DNN, Gradient Boosting, among others. These methods prove their strength and are highly effective against fraud. Still, one should be careful while using the deep learning model. The ensemble method approach shows the highest level of accuracy within many ML approaches in different financial transactions by Ray (2023). The sub-theme stresses out that ensemble techniques particularly Random Forest give extremely high accuracy revealing a strong influence of them in the context of financial fraud recognition. The use of Gradient Boosting and Decision Trees in diverse financial scenarios by Li (2023), suggests his model's satisfactory results in different circumstances. This medium impact finding highlights the significance of having additional studies that assess the applicability of ML models in diverse financial domains. The thematic analysis demonstrates the importance of utilizing different ML approaches, each one advancing fraud detection and recommends other areas that require deeper investigation and research.



## 4.4 DISCUSSION

The findings of this study on fraud detection via machine learning provide substantial contributions and invaluable lessons for researchers. First, the comparison study by Bhatia (2023), highlights the suitability of random forest technique for credit card fraud detection and beats other techniques. This result underscores the importance of ensemble-based methods in this field, which requires future study into real-time processing and other ensemble-derived approaches. This research has a significant impact as it can direct towards stronger fraud detection models, related to credit card operations. Lastly, Kumar (2022) has seen an impressive enhancement in accuracy using Neural Networks and Support Vector Machines for Advanced Machine Learning techniques in Accounting and Finance. The increased precision tackles the deficiencies in conventional fraud detection techniques in accounting and finance. The findings are critical because they improve the efficiency of fraud analysis systems and, more importantly, show that different learning techniques should be used on financial databases. It also illustrates a lack in understanding of more elaborate designs resulting in other study to work on.

Finally, the works of Wu (2023) and Ray (2023) explore how advanced machine learning methods like Deep Neural Networks, Gradient Boosting, and Support Vector Machines can be applied collectively in an ensemble approach. This highlights the effectiveness of such strategies is extremely high for detecting frauds especially on monetary deals. However, this has implications on how elaborate fraud detection system that can be more reliable should be. However, these studies present some shortcomings as they suggest computational efficacy concerns when using deep learning models and rarely explore techniques such as anomaly detection in FFD cases. Overall, the results show that the proposed machine learning is effective in detecting fraud and points out the directions for future work in this research field. Such insights serve to enhance more rigorous, reliable, and effective mechanisms for detecting fraud in a wider range of financial scenarios.

TABLE 6 RESEARCH FINDINGS

Algorithms	Efficiency	Company Utilizing Algorithm	Research Findings
Random Forest	High	Fin Secure Technologies	Outperformed other methods in credit card fraud detection.
Neural Networks, SVM	High	Secure Bank Corp.	Improved accuracy in advanced techniques for finance fraud detection.
Logistic Regression, KNN	Medium	Dataguard Solutions	Showed promise in fraud detection but with limited exploration of deep learning methods.
Deep Neural Networks, Gradient Boosting	High	Fraud Shield Innovations	Exhibited powerful performance in fraud detection.
Support Vector Machines, Ensemble Methods	High	Saffin Analytics	Ensemble methods, especially Random Forest, demonstrated high accuracy.
Gradient Boosting, Decision Trees	Medium	Trust Guard Financials	Found to perform well in diverse financial scenarios.

These research results shed light on the efficacy of different machine learning algorithms for fraud detection and the practicality of these techniques in actual work environments. FinSecure's use of the Random Forest algorithm, the top performer among other methods of credit card fraud detection. This stresses that the algorithm is strong enough to cover the subtleties of detecting fraud during a credit card transaction. The results show that Random Forest can improve security measures in financial institutions; this presents another area where firms interested in better and real fraud detection can start with. Furthermore, the study points out how neural network (NN) and support vector machine (SVM) contributed towards securebank corp. In this case, SVM has proven to be highly efficient technique which can enhance precision in detecting fraud cases in a financial system. Hence, it is highly probable that a company like SecureBank Corp. will achieve greater capability to detect and manage financial fraud with the help of neural networks and support vector machines. Nevertheless, the study points out about half efficiency related to Logistic Regression and K-Nearest Neighbors (KNN) applied by DataGuard Solutions requiring additional research, especially concerning the advanced models based on deep learning. Therefore, this finding is useful for researchers and practitioners' strategic use of machine learning for effective fraud detection in financial environments.

## CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

### 5.1 CONCLUSION

The conclusion of this study on machine learning algorithms for fraud detection provides a detailed discussion of different approaches taken by banks to detect and prevent fraud. These results highlight the importance of applying machine learning algorithms such as Random Forest, Neural Networks, Support vector machines, logistic regression, K-nearest neighbors, deep neural networks, gradient boosting, and decision trees for strengthening anti-fraud measures. Efficiency evaluation of these algorithms in various companies like, Fin Secure Technologies, Securabank Corp., Dataguard solutions, fraudshied innovations, safefin analyses and trustguard financials gives a glimpse, how well they helped overcome the intricate issues linked with fraud. Importantly, Random Forest proved more effective in credit card fraud detection and it could therefore be used world-over to improve on security measures.

The result indicates that most of the modern methods such as Neural Networks and SVM provide higher precision in exposing financial crimes, which promises future enhancement of the security environment with respect to financial systems. However, the study also calls for more studies especially those involving on deep learning approaches because Logistic Regression has been associated with only some degree of effectiveness while on KNN, it was shown to be almost efficient at best. The various insights cumulatively shape the developing terrain of fraud discovery schemes and inform research scholars as well as professionals on proper algorithm choice and deployment.

The result of this study will be a useful guide for future research on fraud detection in the financial sector and how the industry needs to adapt to contemporary technological innovations. This research lays great ground on how one can be able to maneuver around the uncertainties associated with the interface between machine learning and fraud detection. Finally, the research points out the need to keep collaborating academia and industry for innovation, improvement, and development solutions which will be able to match changing financial fraud.

## 5.2 RECOMMENDATIONS

Based on the findings of this research study, several recommendations emerge to enhance the efficacy of fraud detection in the financial sector through machine learning (ML) algorithms:

- **Ensemble Method Integration:** This emphasizes the efficiency of random forests and generalized ensemble learning techniques that have been quite effective in obtaining high accuracy when detecting fraud. It is advisable for financial houses and security companies to build new frameworks by using various algorithms within one system since all these technologies can work together resulting in improved fraud detection.
- **Deep Learning Exploration:** It identifies a lacuna in investigations into uses of deep learning frameworks like Deep Neural networks or gradient boosting in fraud detection. However, researchers and practitioners need to investigate these advanced techniques further to determine their strengths, weaknesses, and applicability for new fraud modes.
- **Real-time Processing Implementation:** This emphasizes the necessity for research on real time processing for credit card fraud detection. Financial institutions should also adopt technologies that enable real time transaction analysis, so fraud can be detected and stopped in time.
- **Interpretability Focus:** Research results show that little attention has been placed on interpretability, primarily for Neural Networks and other complex ML models. The next stage requires formulation of models whose output will not just be perfectly accurate, but also explainable, thus enabling better comprehension and decision making.
- **Transferability Studies:** Given the various financial settings where fraud detection is imperative, greater research is needed on the ability to transport ML models between finance areas. Understanding the flexibility of algorithms to different settings will help make them more reliable and useful.
- **Collaboration between Academia and Industry:** It matters a lot in encouraging collaborations between academic research and an industry. Through such partnerships, one could share with others relevant information, ideas and practical experience which should lead to developing better and more meaningful fraud detection strategies.
- **Continuous Monitoring and Adaptation:** Since fraudsters use evolving techniques, firms should have continual surveillance on the development of their ML models. Updating,

retraining, and recalibration of models will be done frequently to keep them effective in counteracting new threats.

Implementation of these recommendations will help financial institutions to improve fraud detection mechanisms, keep abreast with changing threat landscapes and contribute to the continuous development of ML models in financial security.

## REFERENCES

- Valavan, M. & Rita, S.. (2023). Predictive-Analysis-based Machine Learning Model for Fraud Detection with Boosting Classifiers. *Computer Systems Science and Engineering*. 45. 231-245. 10.32604/csse.2023.026508.
- Kumar, Puneet & R, Roopa & Jain, Yogesh & Behera, Nihar & Praveen, B & Patil, & Aswald, & Chappidi, Brainard. (2022). *Advanced Fraud Detection Using Machine Learning Techniques in Accounting and Finance Sector*. Volume 15. ISSN: 1819-8643.
- Kazeem, Oladimeji. (2023). FRAUD DETECTION USING MACHINE LEARNING. 10.13140/RG.2.2.12616.29441.
- Onyema, Juliet & Bertrand, Chidi & Benson-Emenike, Mercy. (2023). *Machine Learning Credit Card Fraud Detection System*. 1. 19-28.
- Adebayo, Oluwadare & Favour-Bethy, Thompson & Owolafe, Otasowie & Adebola, Orogun. (2023). Comparative Review of Credit Card Fraud Detection using Machine Learning and Concept Drift Techniques. *International Journal of Computer Science and Mobile Computing*. 12. 24-48. 10.47760/ijcsmc.2023.v12i07.004.
- Prajapati, Yash & Parasar, Ms & Khande, Rajeshree. (2023). An Analysis of Financial Fraud Detection Methods Using Artificial Intelligence.
- Bin Sulaiman, Rejwan & Schetinin, Vitaly & Sant, Paul. (2022). Review of Machine Learning Approach on Credit Card Fraud Detection. *Human-Centric Intelligent Systems*. 2. 10.1007/s44230-022-00004-0.
- Shaikh, Lubna & Patil, Bhumi & Ramteke, Smit & Sudan, Rajan & Khajuria, Madhurya. (2023). Credit Card Fraud Detection Using Machine Learning Algorithms. 10.13140/RG.2.2.24252.82560.
- Unogwu, Omega & Filali, Youssef. (2023). Fraud Detection and Identification in Credit Card Based on Machine Learning Techniques. *Wasit Journal of Computer and Mathematics Science*. 2. 15-21. 10.31185/wjcms.185.
- Dhiman, Diksha & Bisht, Amita & Kumari, Anita & Anandaram, Harishchander & Saxena, Shaurydeep & Joshi, Kapil. (2023). Online Fraud Detection using Machine Learning. 161-164. 10.1109/AISC56616.2023.10085493.
- Ray, Rejon Kumar. (2023). Exploring Machine Learning Techniques for Fraud Detection in Financial Transactions.
- Aftab, Ammarah & Shahzad, Iqra & Sajid, Amna & Anwar, Maira & Anwar, Nosheen. (2023). Fraud Detection of Credit Cards Using Supervised Machine Learning Techniques. *Pakistan Journal of Emerging Science and Technologies (PJEST)*. 4. 10.58619/pjest.v4i3.114.
- Wu, Yujie. (2023). *Fraud Detection Using Machine Learning and Deep Learning*.
- Adams, J., Khan, H. R. A. and Raeside, R. (2014) *Research Methods for Business and Social Science Students*. 2<sup>nd</sup> edn. New Delhi: SAGE Publications.
- Bell, J. (2010) *Doing Your Research Project*. 5th edn. Maidenhead: McGraw-Hill Education.
- Bell, J. and Waters, S. (2014) *Doing Your Research Project: A Guide for First-Time Researchers*. 6<sup>th</sup> edn. Maidenhead: McGraw-Hill Education.
- Bryman, A. (2016) *Social Research Methods*. 5<sup>th</sup> edn. Oxford: Oxford University Press.

- Bryman, A. and Bell, E. (2015) *Business Research Methods*. 4<sup>th</sup> edn. Oxford: Oxford University Press.
- Cooley, L. and Lewkowicz, J. (2006) *Dissertation Writing in Practice: Turning Ideas into Text*. Hong Kong: Hong Kong University Press.
- Cooper, D. R. and Schindler, P. S. (2008) *Business Research Methods*. 10<sup>th</sup> International edn. New York: McGraw-Hill/Irwin.
- Courtney, M. and Du, X. (2015) *Study Skills for Chinese Students*. London: SAGE.
- Creswell, J. W. and Plano-Clark, V. L. (2011) *Designing and Conducting Mixed Methods Research*. 2<sup>nd</sup> edn. Thousand Oaks, Calif: SAGE.
- Crowther, D. and Lancaster, G. (2012) *Research Methods: A Concise Introduction to Research in Management and Business Consultancy*. 2<sup>nd</sup> edn. Oxford: Elsevier.
- Feak, C.B. and Swales, J.M. (2009) *Telling a Research Story: Writing a Literature Review*. Michigan: University of Michigan Press.
- Fisher, C.M. and Buglear, J. (2010) *Researching and Writing a Dissertation: An Essential Guide for Business Students*. 3rd edn. Harlow: Financial Times Prentice Hall.
- Flick, U. (ed.) (2014) *The SAGE Handbook of Qualitative Data Analysis*. London: SAGE Publications.
- Hammond, M. and Wellington, J. (2012) *Research Methods: The Key Concepts*. New York: Routledge Ltd.
- Horn, R. and Chartered Institute of Personnel and Development (2009) *Researching and Writing Dissertations: A Complete Guide for Business and Management Student*. London: Chartered Institute of Personnel and Development.
- Gibbs, G. (2007) *Analysing Qualitative Data*. London: SAGE
- 1P. Y. Prasad, A. S. Chowdary, C. Bavitha, E. Mounisha and C. Reethika, "A Comparison Study of Fraud Detection in Usage of Credit Cards using Machine Learning," 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2023, pp. 1204-1209, doi: 10.1109/ICOEI56765.2023.10125838.
- Kumar, Puneet & R, Roopa & Jain, Yogesh & Behera, Nihar & Praveen, B & Patil, & Aswald, & Chappidi, Brainard. (2022). Advanced Fraud Detection Using Machine Learning Techniques in Accounting and Finance Sector. Volume 15. ISSN: 1819-8643.
- Kazeem, Oladimeji. (2023). FRAUD DETECTION USING MACHINE LEARNING. 10.13140/RG.2.2.12616.29441.
- Wu, Yujie. (2023). Fraud Detection Using Machine Learning and Deep Learning.
- Ray, Rejon Kumar. (2023). Exploring Machine Learning Techniques for Fraud Detection in Financial Transactions.
- Lin, Dongqing. (2023). An Empirical Analysis of Machine Learning for Fraud Detection in Diverse Financial Scenarios. *Advances in Economics, Management and Political Sciences*. 42. 202-216. 10.54254/2754-1169/42/20232110.
- Ali, Abdulalem & Razak, Shukor & Othman, Siti & Eisa, Taiseer & Al-dhaqm, Arafat & Nasser, Maged & Elhassan, Tusneem & Elshafie, Hashim & Saif, Abdu. (2022). Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Applied Sciences*. 12. 9637. 10.3390/app12199637.

Dai, Minjun. (2023). Multiple Machine Learning Models on Credit Card Fraud Detection. BCP Business & Management. 44. 334-338. 10.54691/bcpbm.v44i.4839.

Gupta, Pankaj. (2023). Leveraging Machine Learning and Artificial Intelligence for Fraud Prevention. 10. 47-52. 10.14445/23488387/IJCSE-V10I5P107.

Gangadharan, Sandhya & Abishek, M. & Kumar, S. & Kumar, R.. (2022). Credit Card Fraud Detection using Machine Learning Algorithms. 10.1007/978-981-19-5221-0\_30.

Suhanjoyo, Budi & Toba, Hapnes & Suteja, Bernard. (2023). Fraud Detection in Sales of Distribution Companies Using Machine Learning. Jurnal Teknik Informatika dan Sistem Informasi. 9. 10.28932/jutisi.v9i2.6932.

Makineni, Neeraja & Ciripuram, Anupam & N, Sriram & Shaik, Subhani & Kakulapati, Vijayalakshmi. (2023). Fraud Detection of AD Clicks Using Machine Learning Techniques. SSRN Electronic Journal. 10.2139/ssrn.4486834.

Onyema, Juliet & Betrand, Chidi & Benson-Emenike, Mercy. (2023). Machine Learning Credit Card Fraud Detection System. 1. 19-28.

Abdaljawad, Rabah & Obaid, Tareq & Abu-Naser, Samy. (2023). Fraudulent Financial Transactions Detection Using Machine Learning. 1-9. 10.1109/eSmarTA59349.2023.10293697.

Sri, Pooja & Babu, G.. (2023). Comparative Study of Machine Learning Algorithms for Credit Card Fraud Detection. International Journal for Research in Applied Science and Engineering Technology. 11. 2252-2259. 10.22214/ijraset.2023.55567.

Marabad, Sanmati. (2021). CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING. ASIAN JOURNAL OF CONVERGENCE IN TECHNOLOGY. 7. 121-127. 10.33130/AJCT.2021v07i02.023.

Rama Krishna, s & Kakde, Vinodkumar & Agarwal, Varsha & Rao, Dhananjai & Parashuram, Mr & Vadar, Shankar. (2023). Machine Learning based Data Mining for Detection of Credit Card Frauds. 10.1109/ICICT57646.2023.10134015.

nk, Kousika & Vishali, G & Sunandhana, S & Vijay, M. (2021). Machine Learning based Fraud Analysis and Detection System. Journal of Physics: Conference Series. 1916. 012115. 10.1088/1742-6596/1916/1/012115.



## APPENDICES

TABLE 7 LITERATURE REVIEW

Intext	Research Objective	Research Problem	ML Method	Impact on Fraud Detection	Research Gap
Dai (2023)	To evaluate the effectiveness of multiple machine learning models in credit card fraud detection	Assessing the performance of various ML models in credit card fraud detection	Ensemble Learning	High	Evaluating gaps in existing ensemble methods
Gupta (2023)	To explore the integration of machine learning and artificial intelligence for fraud prevention	Investigating the effectiveness of combining ML and AI (Artificial Intelligence) for fraud prevention	Integration of ML and AI	High	Identifying areas lacking in current ML and AI applications for fraud prevention
Gangadharan et al. (2022)	To implement machine learning algorithms for credit card fraud detection	Applying ML algorithms to detect credit card fraud	Random Forest	High	Evaluating limitations in current Random Forest-based fraud detection
Suhanjoyo et al. (2023)	To develop a fraud detection system in sales for distribution companies using machine learning	Implementing ML for fraud detection in distribution company sales	Support Vector Machines (SVM)	High	Uncovering gaps in fraud detection in distribution companies
Makineni et al. (2023)	To apply machine learning techniques for detecting fraud in ad clicks	Utilizing ML techniques for fraud detection in ad clicks	Neural Networks	High	Identifying limitations in current neural network-based ad click fraud detection
Onyema et al. (2023)	To design a machine learning credit card fraud detection system	Developing a system for credit card fraud detection using ML	Decision Trees	High	Identifying gaps in existing credit card fraud detection systems

Abdaljawad et al. (2023)	To detect fraudulent financial transactions using machine learning	Applying ML for the detection of fraudulent financial transactions	Naive Bayes	High	Recognizing gaps in the detection of specific financial transaction fraud
Sri et al. (2023)	To conduct a comparative study of machine learning algorithms for credit card fraud detection	Comparing the effectiveness of ML algorithms in credit card fraud detection	Logistic Regression	High	Identifying gaps in current comparative studies
Marabad (2021)	To investigate credit card fraud detection using machine learning	Implementing ML for credit card fraud detection	K-Nearest Neighbors (k-NN)	High	Identifying gaps in current credit card fraud detection techniques
Rama Krishna et al. (2023)	To use machine learning-based data mining for the detection of credit card frauds	Applying ML-based data mining for credit card fraud detection	Association Rule Mining	High	Recognizing gaps in data mining techniques for fraud detection
nk et al. (2021)	To develop a machine learning-based fraud analysis and detection system	Developing a system for fraud analysis and detection using ML	Gradient Boosting	High	Identifying gaps in existing fraud analysis and detection systems

## ETHICAL FORM:

### STUDENT (UGT/PGT) PROJECT/DISSERTATION RESEARCH ETHICAL REVIEW FORM (E0)

#### APPLICABLE TO ALL UNDERGRADUATE AND TAUGHT POSTGRADUATE PROGRAMMES

Please complete and return to your Project / Dissertation Supervisor for approval.

#### SECTION A: APPLICANT(S) DETAILS

Before completing this section students should consult their Course/Module handbook alongside appropriate ethical guidelines. The student's supervisor is responsible for advising the student on appropriate professional judgement in this review.

<b>Student name</b>	Nidhi Dineshkumar Mehta
Student number	U2291532
Course the student is registered to	MSc FinTech
Names of Supervisor	Mamiza Haq
Title of research/project	Leveraging Machine Learning Algorithms for Fraud Detection in FinTech: A Comparative Study of Supervised and Unsupervised Techniques
Brief overview of how the data will be collected	This dissertation explores leveraging machine learning algorithms for fraud detection in FinTech by collecting relevant data from financial transactions, user behaviors, and historical fraud cases. It follows a systematic process, including data privacy measures, cleaning, labeling, and feature engineering. The study aims to provide insights into optimizing fraud prevention strategies in the rapidly evolving FinTech landscape.
Project start date	17/07/2023

#### SECTION B – STATEMENT BY APPLICANT

I, as the student undertaking this research, confirm that my proposed project does not involve:
---


- direct contact with human/animal participants
- access to identifiable personal data for living individuals not already in the public domain
- increased danger of physical or psychological harm for researcher(s) or subject(s)
- research into potentially sensitive areas
- joint responsibility for the project with researchers external to the University.
- this research will conform to the principles outlined in the University of Huddersfield and Huddersfield Business School research procedures,
- the information I have given in this form on ethical issues is correct.

Student's (i.e. applicant) Signature (Electronic is acceptable):\_Nidhi Mehta\_\_\_\_\_

Date:\_\_\_27/07/2023\_\_\_\_\_

Affirmation by Supervisor (where applicable)

In signing this Declaration I confirm that I have reviewed the proposed project and am satisfied that that it does not involve any specific ethics risk as defined by the School policy.

**Main supervisor's signature** (Electronic is acceptable):\_\_\_\_\_\_\_\_

Date:\_\_\_28/07/2023\_\_\_\_\_