International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Fraud Detection using Machine Learning Algorithms on Financial Transaction Data

Prachi [(1)], Asst. Prof. Narender Kumar [(2)]

Student [(1)], Professor [(2)]
Dept: Department of Computer Science
Hemvati Nandan Bahuguna Garhwal University
University: Doon University Dehradun
Email: prachirawal48@gmail.com

ABSTRACT:

Fraud detection is a critical challenge in the financial industry, as fraudulent activities can result in significant financial losses and damage to a company's reputation. Traditional rule-based systems for fraud detection often struggle to keep up with the rapidly evolving techniques used by fraudsters. To overcome these limitations, machine learning algorithms have gained popularity for their ability to detect fraudulent patterns in financial transaction data.

In this study, we propose a fraud detection system that leverages machine learning algorithms on financial transaction data. The goal is to develop a robust and accurate model that can identify fraudulent transactions while minimizing false positives. The system follows a two-step process: feature engineering and model training.

First, we perform extensive feature engineering to extract meaningful features from the transaction data. These features include transaction amount, location, time of day, and historical customer behavior. Additionally, we incorporate external data sources, such as IP geolocation databases and blacklists, to enhance the fraud detection capabilities.

Keywords: Fraud detection, machine learning, financial transactions, supervised learning, unsupervised learning, logistic deterioration, conclusion trees, random forests, support vector machines, neural networks, clustering algorithms, feature engineering, evaluation metrics, ensemble methods, incongruity recognition, feature selection.

1. Introduction:

Fraud detection using machine learning algorithms on financial transaction data has become an increasingly important field due to the rising prevalence of fraudulent activities in the digital age. Machine learning techniques offer powerful tools to identify patterns and anomalies in large volumes of transactional data, enabling organizations to proactively detect and prevent fraudulent activities.

Fraud detection involves the identification and prevention of fraudulent behavior within financial systems, such as credit card transactions, insurance claims, or online banking activities. Traditional rule-based methods and instruction booklet reviews are often insufficient to keep up with the sophistication and scale of modern fraud. Machine learning techniques can significantly enhance fraud recognition by automatically learning patterns and adapting to evolving deceptive strategies.

1.1. Common Machine Learning Algorithms for Fraud Detection:

There are several common machine learning algorithms used for fraud detection. These algorithms leverage various techniques to identify fraudulent patterns and anomalies in large datasets. Here are some of the commonly used algorithms:

1. Logistic Regression: Logistic deterioration is a binary classification algorithm that is often used in fraud recognition. It models the relationship between the input variables and the probability of a fraudulent event occurring. Logistic regression is straightforward, interpretable, and computationally well-organized.

2. Decision Trees: Decision trees are versatile algorithms that can be used for both organization and regression tasks. They generate a tree-like model of decisions and their possible consequences. Decision trees are often used in deception detection as they can capture complex associations and identify important facial appearance for distinguishing fraudulent behavior.

3. Random Forests: Random forests are an ensemble learning method that combines multiple decision trees to make predictions. They generate a set of decision trees and aggregate their outputs to make a final prediction. Random forests are effective for fraud detection as they can handle large datasets, reduce overfitting, and provide feature importance rankings.

4. Gradient Boosting Methods: Gradient boosting methods, such as Gradient Boosted Trees (GBT) and XGBoost, are powerful algorithms that build a strong predictive model by combining weak individual models. They iteratively train new models that focus on the misclassified examples from previous models. Gradient boosting methods are widely used in fraud detection due to their high accuracy and ability to handle imbalanced datasets.

5. Support Vector Machines (SVM): SVM is a popular algorithm used for both categorization and deterioration tasks. It constructs a hyper plane or set of hyper planes that maximize the separation between classes. SVM is effectual in fraud detection when dealing with high-dimensional data or when the pronouncement border line is non-linear and composite.

1.2 Challenges and Considerations:

While machine learning algorithms offer promising solutions for fraud detection, several challenges and considerations need to be addressed:

1. Data Volume and Complexity: Financial institutions generate vast amounts of data from various sources, including transactions, customer profiles, and external data feeds. Analyzing this data to identify fraudulent patterns and anomalies can be a complex task. The sheer volume of data and its diverse nature require sophisticated data processing and analytical techniques.

2. Real-Time Detection: Fraudsters are becoming increasingly sophisticated, and their methods are evolving rapidly. Financial institutions need to detect and respond to fraud in real-time to minimize potential losses. Real-time detection requires efficient data processing systems and advanced algorithms capable of analyzing and identifying fraudulent activities within milliseconds.

3. Data Quality and Integration: Effective fraud detection relies on accurate and reliable data. However, financial data can be inconsistent, incomplete, or contain errors. Data quality issues can significantly impact the accuracy and effectiveness of fraud detection models. Integrating data from multiple sources and systems also poses challenges, as different data formats and structures need to be harmonized.

4. False Positives and Negatives: Balancing the detection of genuine fraud cases while minimizing false positives (legitimate transactions flagged as fraudulent) and false negatives (fraudulent transactions classified as legitimate) is crucial. High false-positive rates can lead to unnecessary customer inconvenience and operational costs, while high false-negative rates can result in financial losses. Striking the right balance requires continuous model refinement and fine-tuning.

5. Adaptive Fraud Techniques: Fraudsters continuously adapt their techniques to circumvent detection mechanisms. They employ sophisticated tactics such as account takeover, synthetic identities, and insider collusion, making it challenging to detect fraudulent activities. Financial institutions must stay abreast of emerging fraud trends and update their detection systems and models accordingly.

2. Research Methods

Fraud detection using machine learning algorithms on financial transaction data is a challenging and important problem. To conduct research in this area, you can follow a systematic approach that involves several key steps. Here's an outline of the research methods you can employ:

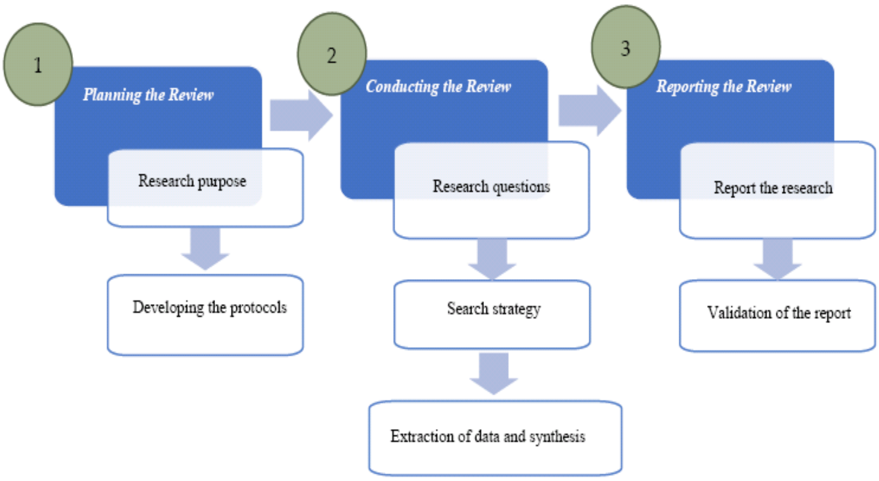

Figure 1. SLR Stages

2.1 Critical areas of research and development

Financial fraud recognition is a critical area of research and development, aimed at identifying and preventing fraudulent activities in various financial systems. more than a few research methods and techniques are in employment to detect monetary fraud, including:

1. Rule-based methods: Rule-based systems use predefined rules and patterns to recognize suspicious activities. These rules are developed base on professional knowledge and experience in detecting widespread fraud patterns. For example, if a transaction exceeds a certain threshold or if multiple transactions are conducted from dissimilar locations within a short period, it may raise misgiving.

2. Anomaly detection: Anomaly recognition technique aim to identify unusual or abnormal patterns in financial data. These methods use arithmetical models to establishment normal behavior and detect movement away from it. Unusual patterns may point toward potential fraud. For instance, if a customer suddenly exhibits significantly different spending behavior compared to their historical data or characteristic spending patterns of comparable customers, it might suggest deceptive activity.

3. Machine learning (ML) approaches: ML algorithms are increasingly being used in monetary fraud recognition due to their ability to gain knowledge of from data and identify complex patterns. Supervised learning algorithms can be trained using chronological data labeled as fraudulent or non-fraudulent communication to classify new transactions. Unsupervised learning algorithms, such as clustering or irregularity uncovering algorithms, can identify patterns or outliers that may point toward fraudulent behavior.

4. Data mining and pattern recognition: Data mining techniques engross extracting useful in sequence and patterns from large datasets. By analyzing immeasurable amounts of transactional data, patterns and correlations associated with fraudulent activities can be acknowledged. Pattern recognition algorithms can then be applied to new information to detect comparable patterns and flag potential scheme.

5. Social network analysis: Social network psychotherapy focuses on the relationships and connections sandwiched between individuals or entities concerned in financial transactions. By analyzing the network organization and transactional behavior, suspicious relationships or patterns can be recognized. For example, detecting networks of colluding individuals or identifying extraordinary transactional flows surrounded by a set of connections can indicate deception.

2.2.3. Study collection criterion

After applying the quest terms in the below digital libraries, a total number of 356 papers were discovered from all quest databases in which 72 imperceptible blanks were caught on and screened from the excavated documents. After screening counterparts, we kept up with the choosing operation exercising the 129 documents that remained. Authors design addition and rejection bars in the searching course to distinguish the most applicable documents. authors defend these probations following the provisions of the grade estimate morals in disposition to guaranty the class of the culled documents as hands down. We pay thecross- checking system to determine whether the named papers match these conditions in order to guarantee the credibility of the consequences. After applting all the below measures and the step of quality assessment criteria, 74 studies were ultimately attained, which are related to the disquisition questions. Tables 2 and 3 display the addition criteria and grade appraisal singly.

Table 2. Exclusion and inclusion criteria

S/N	Exclusion	Inclusion
1	Papers that servant’s central of engrossment on fiscal delusive deals	The papers that are directed from 2010 to 2021.
2	Papers that are in the conformation of objectifications, short documents, bills, and handwriting branches.	Papers that don't appertain to the use of machine literacy/ data mining styles.
3	Studies that don't bring up their donation appraisement criteria	papers that concentrate on fiscal deception discovery and applied Machine Learning styles.
4	Inquests that weren't printed in the English vocabulary dispatch.	A peer- oversaw do exploration number of jotting. Probations were conducted in English no further than.

Table 3. Quality Assessment

ID	Quality Assessment
1	Is the principle of the practice clear?
2	Are the ways easily formulated and demonstrated?
3	Are the offered ways easily accessible and enforced?
4	Is the experimental modus operandi easily depicted?
5	Does the practice fabricate benefactions to the SLR?
6	Are developmental trials easily couched?
7	Are the account procedures easily articulated?
8	Are the induction and prospects command really phrased?

2.3. Data Extraction and Synthesis

The data nativity proceeding involves looking figures to treasure data from the nominated documents for the SLR(2). holding the facts in the data nativity conformation, the prescribed disquisition problems for the SLR can be returned. The data pulled in the data geniture stand is sported in Table 4.

Table 4. Form of Information Extraction

Search Method	Information Extracted	Purpose of the Extraction
Manual Search	The block of financial flimflam manipulated in the literacy.	RQ1
	The modus operandi second- aspect for the fraud recognition.	RQ2
	The intention of the reverie	RQ1, RQ2
	The assessment criteria used to address which fashion	RQ3
	unborn bearing, trends, and breaches in the daydreaming	RQ4
	Conclusion of the daydreaming	RQ1, RQ2, and RQ4
Automatic Search	Title of the daydreaming	
	Study explanation and Meta- Analysis	

In the first string of Table 4, the quest ground plan used to concentrate the data for the data geniture including involuntary and manual birth is prescribed. The alternate and third queues are the order of the data pulled and the ambition of lodging the facts, singly. This information is analyzed to fabricate it easier for grouping similar studies together in terms of the type of fraudulent exertion addressed, the ways used for the fraud discovery, and the value judgment standards used for the evidence complex as well as the disquisition openings and future imperative.

3. Meta-Analysis and Research Method

This district presents the quest results attained from the no voluntary stage of the review proceeding, which involves concluding the actionable examinations to be considered in this SLR study. We first present the description of the reviewed studies in this SLR and similarly subsequently answer each of the disquisition interrogatives prescribed in the quarter.

3.1. Study Description

The composition of papers bonding to fiscal fraud detection using Machine Learning paths from 2011 to 2022 is flashed in Figure 2, which provides a chronological summary of the got out papers used in this reexamination. The graph illustrates how exploration in this field has shown a promoting trend in recent occasions, especially since 2014 when the number of papers got out began to roll out significantly. furthest documents harnessed in this reconsideration were released after 2014. Within the study period, 2019 endured the topmost number of papers(11) got out in this area, succeeded by 20 blanks in 2019 and 2018. It can be adhered that there was a lower rate of applicable publications in 2014 as just four blanks were examined in that time. numeral 3 shows colorful newspaper cognomens and the work of practicable papers that were harnessed in this reconsideration.

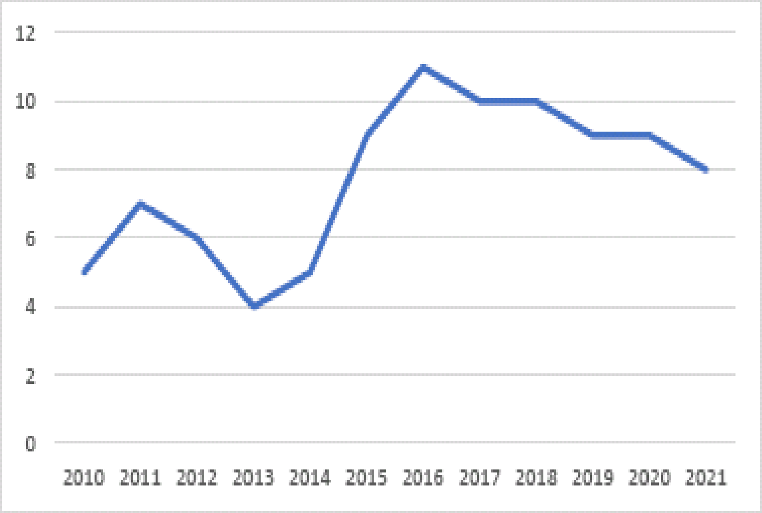

Figure No 2. Year of Summer Article

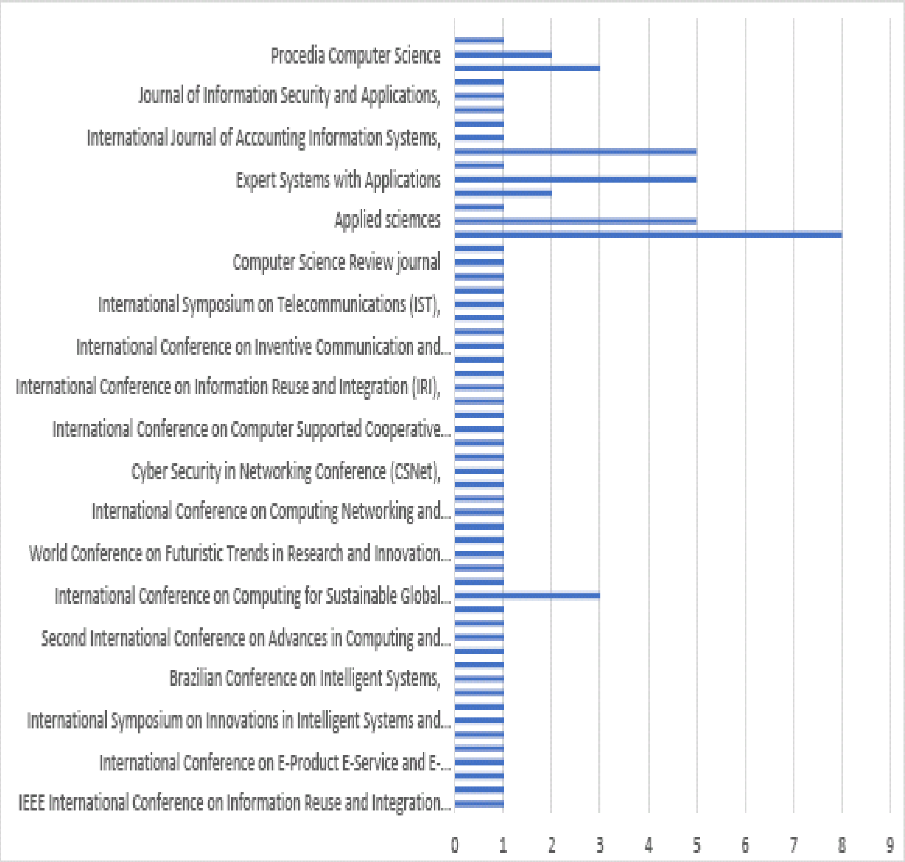

Figure No 3. Article per year journals

3.2. Result Combination

This district presents the aftereffects of the data amalgamation to maneuver the exploration interrogatives grounded on the named blanks. Ergo, in this district, the exploration interrogatives allowed for the SLR will exist answered.

3.2.1. RQ1: What are the dissimilar Categories of Fraudulent behavior That Are Addressed Using Machine Learning technique?

Machine learning (ML) techniques are widely used to address various categories of fraudulent activities. Some of the key categories of fraud that ML techniques can help combat include:

1. Credit Card Fraud: ML models can be trained to detect fraudulent credit card transactions by analyzing patterns and anomalies in transaction data. ML algorithms can learn from historical data to identify suspicious activities such as unusual purchase amounts, atypical transaction locations, or inconsistent spending patterns.

2. Insurance Fraud: ML techniques can be in employment to detect fraudulent insurance claims. ML models can analyze large amounts of data related to claims, policyholders, and historical fraudulent patterns to identify suspicious claims. This can include identifying abnormal claim patterns, discrepancies in reported information, or links between individuals involved in multiple fraudulent claims.

3. Banking and Financial Fraud: ML algorithms can be utilized to detect various forms of banking and financial fraud, such as account takeover, money laundering, and identity theft. ML models can analyze customer behavior, transactional data, and account activity to detect anomalies and flag suspicious activities that deviate from normal usage patterns.

4. Online Retail Fraud: ML techniques can help combat fraudulent activities in online retail, such as fake product reviews, account takeovers, or fraudulent transactions. ML models can analyze customer data, browsing behavior, and purchase history to identify anomalies or suspicious patterns that may indicate fraudulent activity.

5. Healthcare Fraud: ML algorithms can be applied to identify fraudulent activities in the healthcare sector, including billing fraud, prescription fraud, and identity theft. ML models can analyze medical records, billing data, and patterns of healthcare provider behavior to flag suspicious claims or irregularities.

6. Social Media and Online Advertising Fraud: ML techniques can be used to detect fraud in the realm of online advertising and social media platforms. ML models can analyze user behavior, engagement metrics, and content patterns to identify fake accounts, click fraud, or suspicious advertising activities.

3.2.2. RQ2: What Are the Machine Based Techniques for Financial deception Detection Employed in the writing?

Machine learning techniques play a crucial role in detecting and preventing financial fraud. Here are some commonly employed machine learning-based techniques for financial fraud detection:

1. Anomaly Detection: Anomaly detection algorithms are used to identify unusual patterns or outliers in financial transactions. Machine learning models can be trained to recognize deviations from normal behavior, such as unexpected transaction amounts, unusual spending patterns, or abnormal user activities. Techniques like clustering, nearest neighbor, and autoencoders are commonly used for anomaly detection.

2. Supervised Learning: Supervised learning algorithms are trained on labeled datasets to classify transactions as fraudulent or legitimate. These algorithms learn patterns and features from historical data, such as transaction amounts, merchant categories, location, time, and user behavior, to make predictions on new transactions. Popular algorithms include logistic regression, decision trees, random forests, and support vector machines.

3. Unsupervised Learning: Unsupervised learning algorithms are used to determine patterns and structures in data exclusive of pre-existing labels. They can help identify clusters or groups of transactions that exhibit similar characteristics, which can be positive for fraud detection. Techniques like clustering (k-means, DBSCAN), association rule learning, and principal component analysis (PCA) are commonly employed.

4. Neural Networks: Deep learning models, such as neural networks, are increasingly used for fraud detection due to their capability to learn compound patterns and features. Recurrent neural networks (RNNs) and convolution neural networks (CNNs) can investigate sequential and structured data to capture sequential dependencies and detect deceptive activities.

5. Natural Language Processing (NLP): NLP techniques can be applied to analyze textual data, such as operation descriptions, customer notes, or online road and rail network, to identify fraudulent content or patterns. Response analysis, topic modeling, and named entity thanks are some NLP technique employed for fraud detection.

3.2.3. RQ3: What Are the presentation Evaluation Metrics second-hand for Financial Fraud Detection Using mechanism erudition Methods

Financial fraud detection using contraption learning methods involves the application of various performance assessment metrics to assess the effectiveness of the models. These metrics help evaluate the model's ability to accurately detect deceptive activities and distinguish them from legitimate transactions. Some commonly used presentation evaluation metrics for financial scheme detection using mechanism learning method include:

1. Accuracy: Accuracy measures the taken as a whole correctness of the model's predictions by comparing the number of in the approved manner classified instances (both fraudulent and legitimate) to the total number of instances. However, accuracy alone may not be sufficient if the dataset is imbalanced, where the number of legitimate transactions far exceeds the number of fraudulent transactions.

2. Precision: Precision measures the proportion of correctly acknowledged fraudulent transactions out of all instances predicted as fraudulent. It focuses on minimizing false positives, which are legitimate transactions mistakenly identified as fraudulent. Higher precision indicates a lower rate of false positives.

3. Recall (Sensitivity or True Positive Rate): Recall measures the percentage of correctly identified fraudulent transactions out of all authentic fraudulent transactions. It focuses on minimizing false negatives, which are deceitful transactions that are incorrectly classified as legitimate. Higher recall indicates a lower rate of false negatives.

4. F1 Score: The F1 score is the harmonic mean of precision and recall and provides a balanced measure of a model's performance. It combines both precision and recall into a single metric, offering a comprehensive evaluation of the model's capability to identify fraudulent transactions while minimizing false positives and false negatives.

5. Specificity (True Negative Rate): Specificity procedures the proportion of correctly identified legitimate transactions out of all actual legitimate transactions. It focuses on minimizing false positives, similar to exactitude but specifically for legitimate transactions. Higher specificity indicates a lower rate of false positives for legitimate transactions.

3.2.4. RQ4 What Are the Gaps and expectations Research track in Machine-Learning-Based scheme Detection?

Machine-learning-based fraud detection has made significant advancements in recent years. However, several gaps and future research directions still exist in this field. Here are some of the key areas that warrant further investigation:

1. **Adversarial attacks:** Fraudsters constantly evolve their techniques to evade detection systems. Future research should focus on developing robust fraud detection models that can withstand adversarial attacks. This involves exploring techniques such as adversarial training, generative adversarial networks (GANs), or incorporating anomaly detection mechanisms to detect novel attack patterns.

2. **Interpretability and explain ability:** Machine learning models used for swindle uncovering often operate as black boxes, construction it challenging to understand the reasons behind their predictions. Future do research should concentrate on developing interpretable and explainable fraud uncovering models. This can help investigators and auditors understand the decision-making process of the models, identify potential biases, and gain insights into the features contributing to fraudulent activities.

3. **Imbalanced datasets:** Fraudulent activities are relatively rare compared to legitimate transactions, resulting in imbalanced datasets where positive (fraud) samples are scarce. Future research should focus on addressing the challenges posed by imbalanced data, such as using sampling techniques (undersampling, oversampling, or hybrid methods) or exploring ensemble methods and active learning approaches to improve fraud detection performance.

4. **Real-time detection:** Many fraud detection systems operate in batch mode, analyzing transactions retrospectively. However, real-time detection is crucial to prevent fraudulent activities promptly. Future research should aim to develop efficient and scalable algorithms that can handle large-scale data in real-time, enabling timely detection and prevention of fraudulent transactions.

5. **Cross-domain and cross-channel fraud detection:** Fraudsters often exploit multiple channels and domains simultaneously, making it essential to develop fraud detection models that can effectively detect and correlate fraud activities across different domains (e.g., online transactions, mobile banking, insurance claims) and channels (e.g., web, mobile apps, call centers).

4. Methodology

The experimenters are trying to develop a fraud discovery technology that uses machine literacy and deep literacy ways to determine whether the online deals are real or fake grounded on the sale databases. still, the discovery of online sale fraud is getting further and more delicate as illegal payments come closer to licit bones

4.1 Dataset

We collected the dataset called “fraudulent sale” from Kaggle Depository for fraud rummage sale discovery. The dataset consists of 6362620 records with 10 features. The mean value of all deals is 144972 USD while the largest sale recorded in this data set quantities to 1981430 USD. still, as you might be guessing right now grounded on the mean and outside, the distribution of the financial value of all deals is heavily right- slanted. The vast maturities of deals are fairly small and only a bitsy bit of deals comes indeed close to the outside.

The features of the fraudulent sale dataset are shown in Table 1.

Table 5: shows the features in the dataset

Features	**Description**
tread	Maps a unit of time in the authentic world. In this case 1 step is 1 hour of time. Total steps 744 (30 days simulation).
variety	redeem, CASH-OUT, withdrawal, compensation and convey
quantity	Amount of the convention in cramped currency.
numeric	purchaser who happening the business
Old balance Org	initial sense of balance sooner than the business deal
New balance Orig	new sense of balance after the operation
Name Dest	customer who is the beneficiary of the business
Old balance Dest	Initial balance recipient before the business. Note that in attendance is not in sequence for clientele that establish with M (Merchants).
New balance Dest	New sense of balance recipient after the contract. Note that there is not in sequence for customers that institute with M (Merchants).
Is Fraud	This is the communication made by the deceptive agents inside the simulation. In this specific dataset the fraudulent performance of the agent’s aims to profit by taking be in charge of or customers balance sheet and try to unoccupied the resources by transferring to another explanation and then cashing out of the system.
Is Flagged Fraud	The business model aims to be in charge of colossal transfers from one account to another and decoration illegal attempts. An illegal attempt in this dataset is an attempt to transport more than 200.000 in a on its have possession of business deal.

4.2 Dataset Analysis

Model efficiency is measured using product presentation criteria comparable as delicacy, recall, perfection and F1- Score.

The objective point is Fraud which is a twice over point with 0(not fraud) and 1(is fraud). There are 6084104 on-fraudulent deals (98.870) and 6485 fraudulent deals (0.129).

As anticipated, paramount deals are non deceptive. The following mental picture underline this significant discrepancy (see Figure 1).

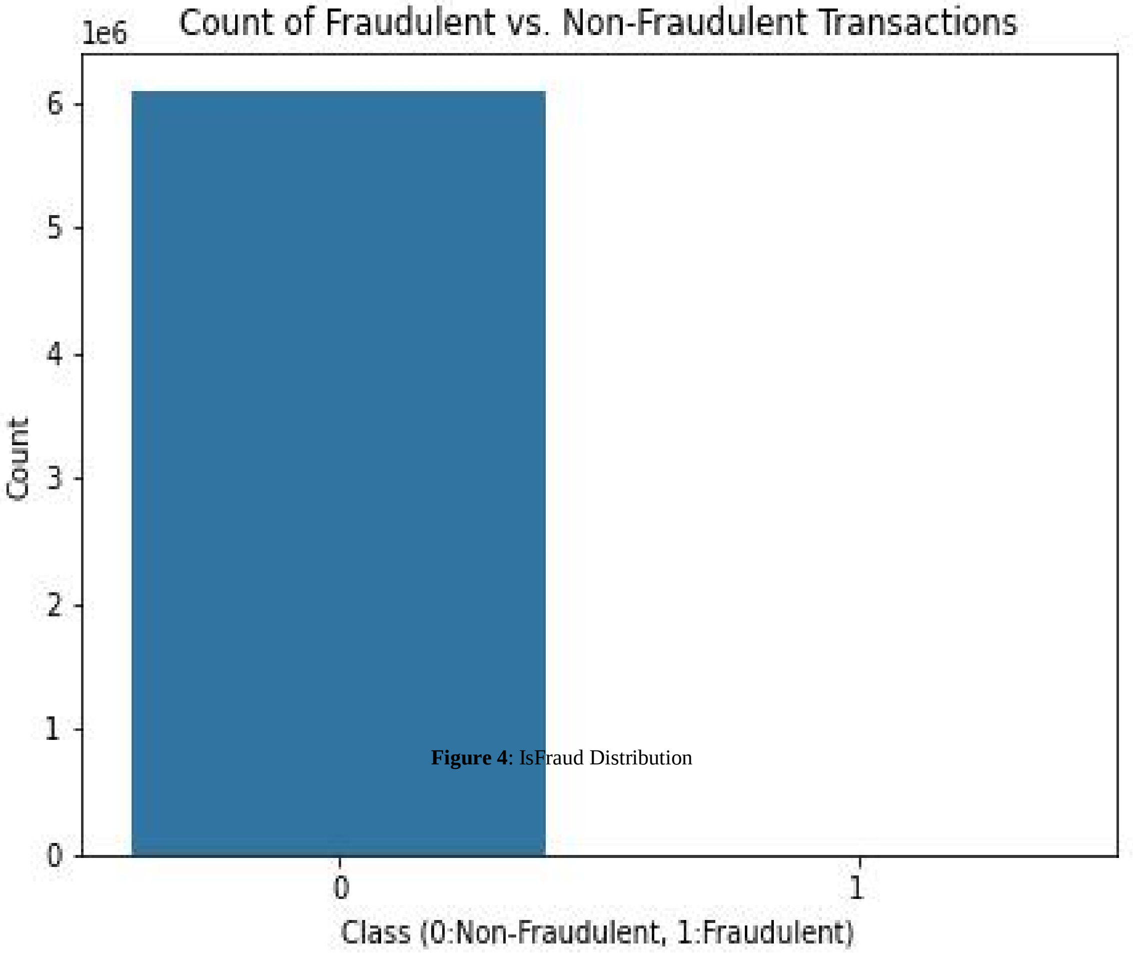

Figure 4: IsFraud Distribution

Finally, it would be motivating to know if there are any earth-shattering correlations between our predictors, more than increasingly with regard to our class inconsistent (isFraud). One of the good numbers visually appealing ways to bring to a close that is by using a headman.

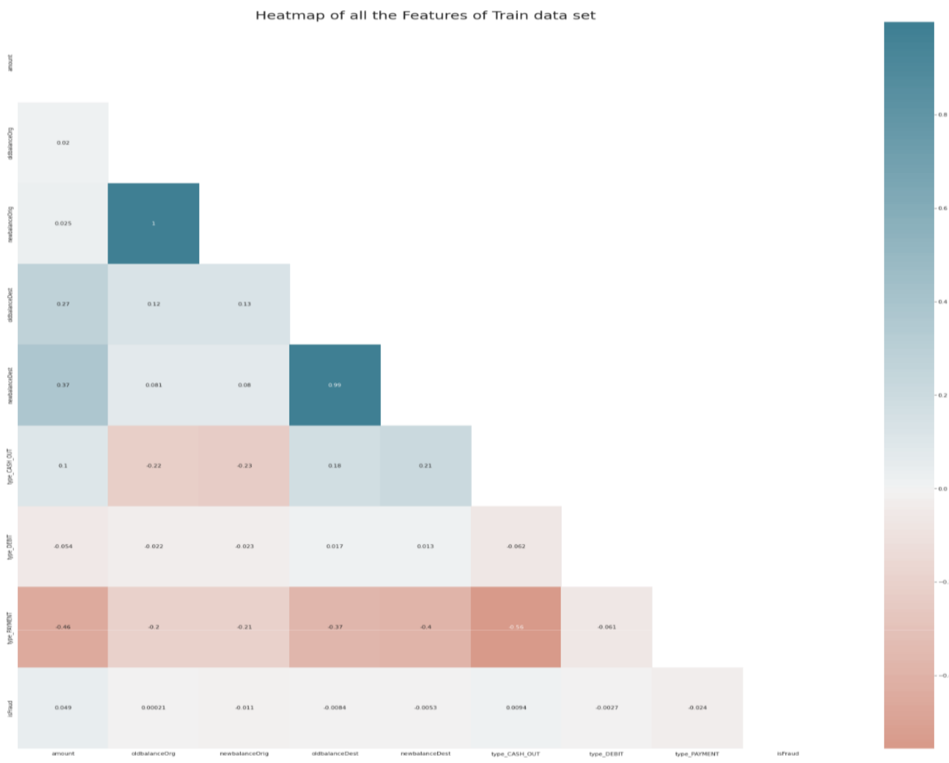

Figure 5: HeapMap Showing Dataset

As can be seen in outline 2, some of our predictors do feel to be linked with the is deception variable. However, there feel to be moderately momentous correlations for analogous variables. This can presumably be accredited to the factor huge class imbalance might disfigure the significance of certain correlations with felicitations to our class variable. Similarly we have checked the unconstructive correlation and positive association with class isFraud as an be seen in Figure 3 and Figure 4. It twisted out that the facial appearance with positive relationship are amount,, And the negative relationship are new balance starting point, new balance Debt,, and category DEBIT.

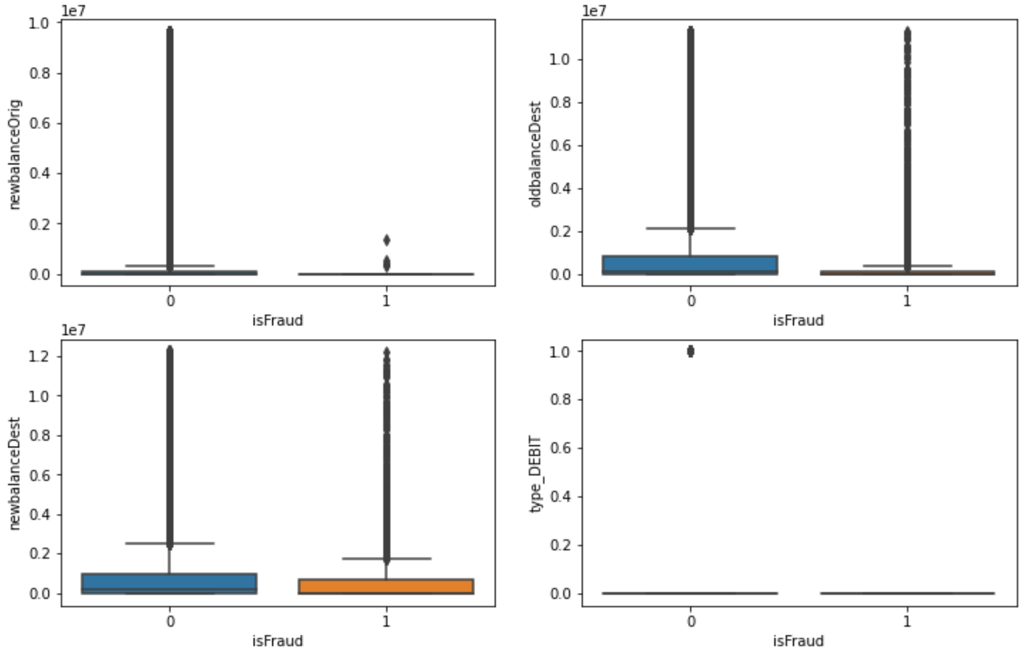

Figure 6:IsFraud Negative Correlation

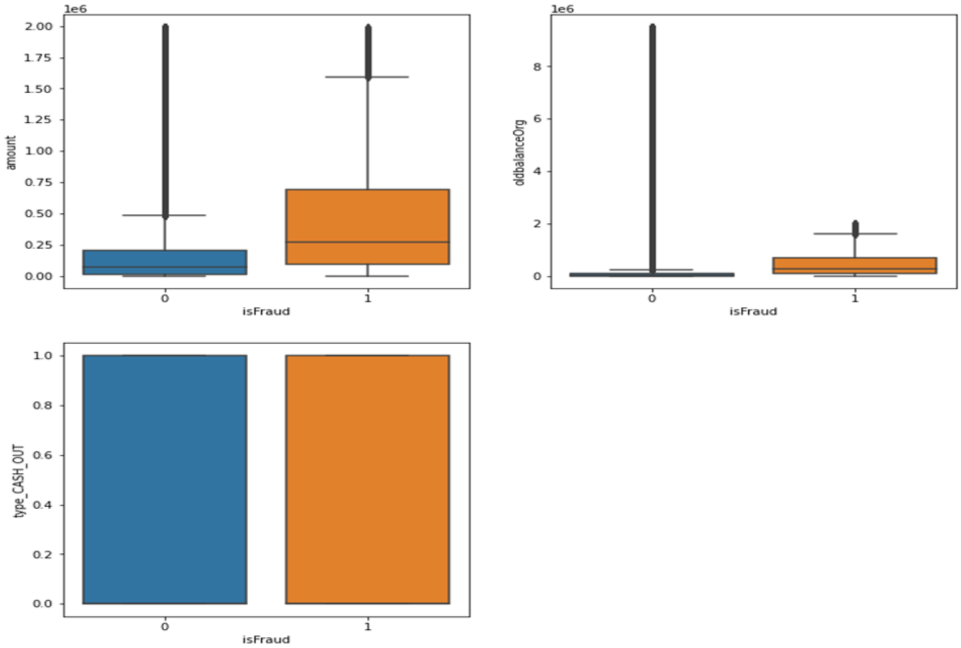

Figure 7: IsFraud Positive Correlation

4.3 First Experiment:

We've used the dataset as is (unstable). We've conclude the dataset into three datasets training, validating and testing. The rate of the splitting was (60 x 20 x 20). We've qualified and tested each model and recorded the results (delicacy, Precision, Recall, F1- score and time considered necessary for the training process in seconds) as can be seen in Table 2.

Table 6: Unbalance Dataset

Model Name	**Accuracy**	**Precision**	**Recall**	**F1_score**	**Time in Sec**
Decision Tree	98.97%	88.22%	84.44%	86.29%	35
MLP Regressor	98.94%	98.94%	98.94%	98.93%	859
Random Forest Classifier	98.97%	98.96%	98.97%	98.96%	1098
Complement NB	54.05%	0.24%	98.62%	0.47%	5
MLP Classifier	98.95%	96.34%	55.41%	70.36%	766
Gaussian NB	58.93%	0.26%	98.62%	0.53%	6
Bernoulli NB	98.89%	100.00%	0.15%	0.30%	5
LGBM Classifier	98.92%	64.80%	55.64%	59.87%	34
Ada Boost Classifier	98.93%	90.67%	41.65%	57.08%	435
K. Neighbors Classifier	98.96%	91.03%	67.14%	77.28%	914
Logistic Regression	98.91%	94.50%	14.21%	24.71%	34
Bagging Classifier	98.97%	93.78%	81.58%	87.25%	251
Deep Learning	98.956%	98.898%	60.752%	75.268%	280

This dataset is tyrannically imbalanced (utmost of the deals arena-fraud). So the algorithms are much more likely to classify new compliances to the middle age class and high delicacy will not tell us anything. To address the predicament of imbalanced dataset we can use under slice and oversampling data approach ways. Oversampling increase the number of nonage class members in the training set. The advantage of oversampling is that no in sequence from the original education set is lost unlike in under slice, as all compliances from the nonage and maturity classes are kept. On the other hand, it's prone to over befitting. There's a type of oversampling called SMOTE (Synthetic nonage Oversampling fashion), which we're disappearing to use to make our dataset balanced. It creates synthetic points from the nonage class.

Also we should not use weakness score as a metric with imbalanced datasets (will be in general high and deceiving), rather we be supposed to use f1-score, perfection/ recall achieve and confusion matrix

Recall of fraud cases (perceptivity) summarizes true positive rate(True positive/ True positive False Negative)- how numerous cases we got truthful out of all the positive bones

Recall ofnon-fraud (particularity) summarize true unconstructive rate(True negative/ True negative False positive)- how numerous cases we got acceptable out of all the negative bones

Precision of fraud cases (True positive/ True positive False positive) summarizes the delicacy of fraud cases detected- out of all prognosticated as fraud, how abundant are correct

Precision ofnon-fraud cases (True negative/ True negative False negative) summarizes the delicacy of non-fraud cases detected- out of all prognosticated Anson-fraud, how numerous are correct

F1- score is the harmonious mean of bring to mind and perfection.

4.4 Experiment Second:

We've evenhanded the dataset using SMOTE approach. The class summit (is Fraud) now is balanced (50 for fraud sale and 50 for non-fraud sale) as in Figure 5. Also we've resolve the dataset for three datasets training, validating and testing as in the first trial. The rate of splitting was (60 x 20 x 20). We've trained and experienced each representation and recorded the results (delicacy, Precision, Recall, F1- score and moment in time needed for the training process in seconds) as can be seen in Table 3.

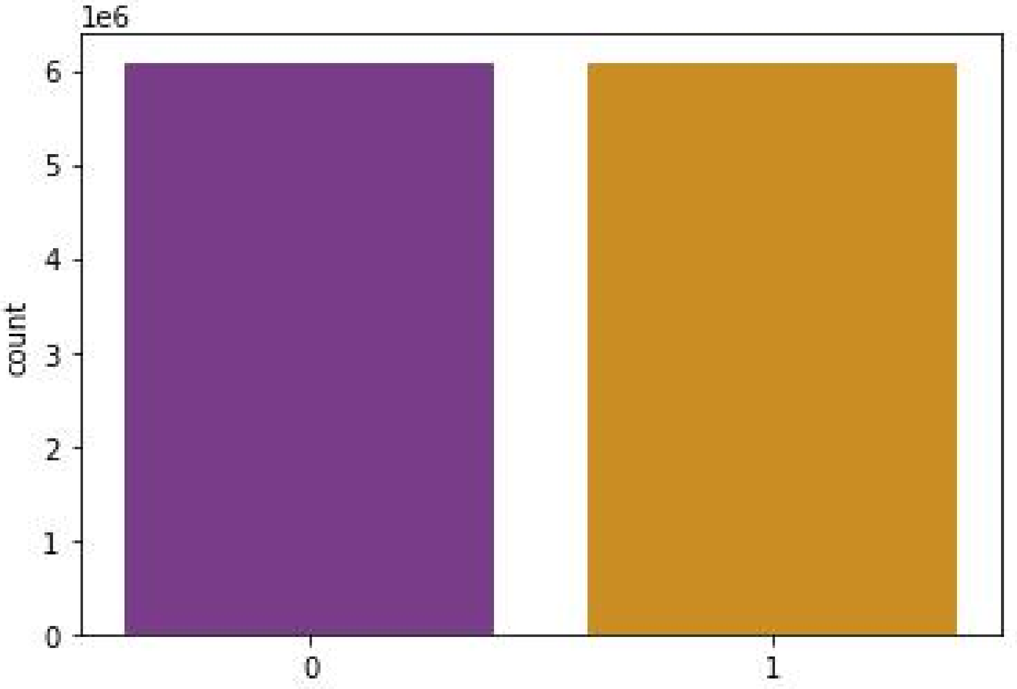

Figure 8: Features isFraud Balance

Table 7: shows the result of the 12 models with SMOTE (balanced dataset)

Model Name	Accuracy	Precision	Recall	F1_score	Time in Sec
Decision Tree Classifier	98.958%	98.945%	98.971%	98.958%	**67**
MLP Regressor	98.389%	98.390%	98.389%	98.389%	1244
Random Forest Classifier	98.950%	98.950%	98.950%	98.950%	1550
Complement NB	78.753%	70.222%	98.798%	82.438%	10
MLP Classifier	98.420%	98.269%	98.574%	98.421%	1798
Gaussian NB	78.331%	69.820%	98.756%	82.145%	11
Bernoulli NB	90.442%	84.089%	98.746%	91.251%	11
LGBM Classifier	98.710%	98.588%	98.833%	98.710%	89
AdaBoost Classifier	97.556%	95.853%	98.411%	97.598%	623
K Neighbors Classifier	98.681%	98.421%	98.943%	98.681%	1624
Logistic Regression	96.694%	94.452%	98.213%	96.774%	109
Bagging Classifier	98.962%	98.946%	98.978%	98.962%	414
Deep Learning	98.183%	98.305%	98.058%	98.183%	560

As can be seen from table 2 and table 3 the entourages

In the unbalanced dataset, the Recall and F1- Score is veritably low for utmost models.

The only two models with unhinged dataset and have veritably high Recall and F1- Score are MLP Regressor and Random Forest Classifier only.

In the balanced dataset, the Recall and F1- Score is veritably high for all models.

The top two loftiest models are Bagging Classifier and Decision Tree Classifier.

Some models can give high recall and F1- Score nevertheless of the dataset balanced or unstable.

After finishing the education of the 13 models we determined the elegant features in the dataset by conniving point worth using Bagging Classifier for the reason that it's the fashionable classifier as in Figure 5.

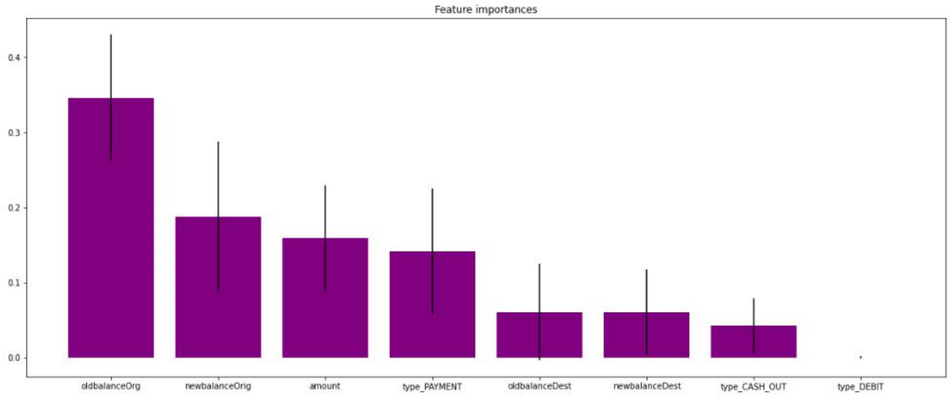

Figure 9: Feature Importance

5. Discussion

In this section, the contented of the SLR is stressed out, which includes all the rage fiscal fraud discovery and piece of equipment literacy ways used in unearthing styles. We distributed the findings of ML ways and the method type in this SLR grounded on their frequency of operation. From the review, it can be experiential that out of all ML approaches discovered in this study, the most popular bones from 2011 to 2021 are epitomized. As a consequence, we linked that the NB algorithm is the most popular move toward used for relating false breaking in in the fiscal sector followed by the SVM and ANN with 11, 10, and 10 papers, independently. This shows that NB, SVM, and ANN are the good number well-liked machine literacy ways used for the fiscal fraud unearthing grounded on the reviewed literature. Figure 4 and Figure 5 shows the frequency distribution of the equipment literacy ways used for deception discovery and the financial fraud types addressed.

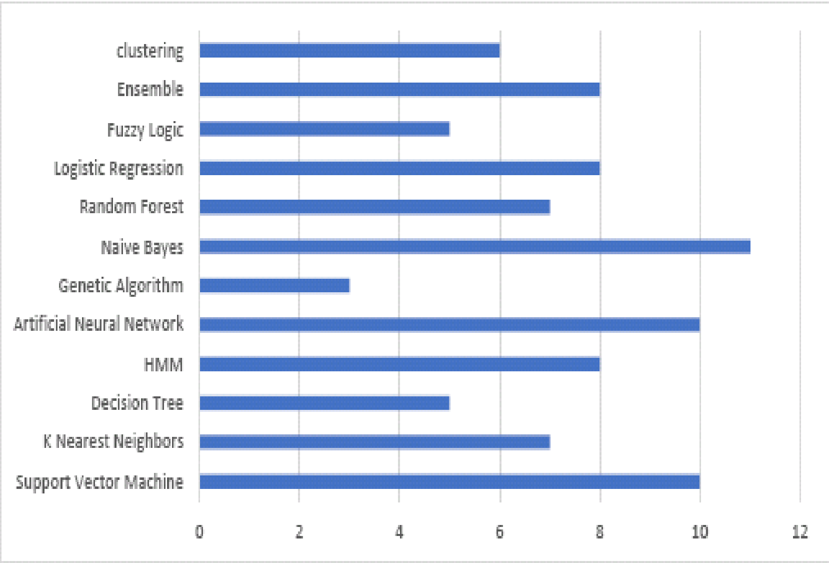

Figure 10. Frequency of fraud Detection

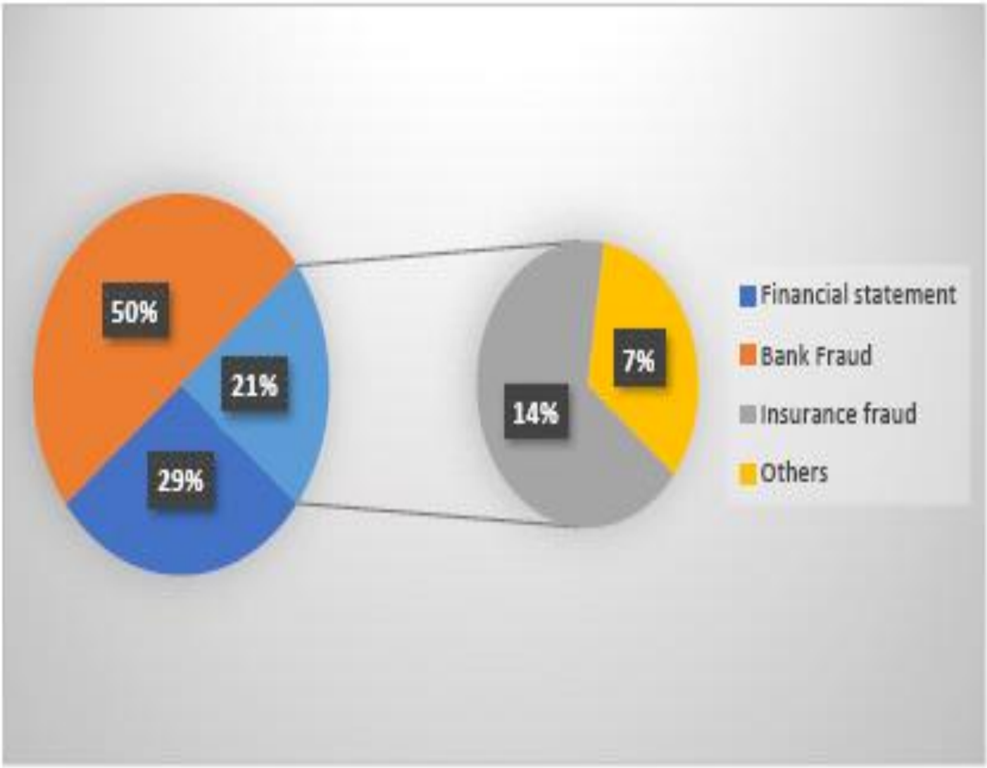

Figure 11. Frequency Fraud Types

Stranded on the answer of our study, we disseminated fiscal fraud into four poles apart orders, similar as fiscal statements bank fraud, insurance fraud, and others. The number of reviewed papers and the type of scheme are handed in Figure 4. It can be shown that out of the review papers, 50 papers contemplate on bank deception, 29 addressed fiscal statements, and 21 concentrated on insurance deception and other fraud at 14 and 7, independently. This show that the maturity of studies have a significant meeting point on depository fraud and fiscal statement fraud, while insurance frauds that take account of healthiness insurance and bus indemnity aren't as constantly linked in the reviewed papers. It also shows that the two most widespread frauds include credit card and fiscal fraud, which become visible to be the most current types of fraud. On the other hand, our review didn't cover magnate laundering, stock, goods, or mortgage fraud for some reason. One of the factor is the difficulty in acquiring these data and the powerlessness to reveal results if they're related to perceptive subjects.

6. Limitation Validity

In this SLR, various ML ways and flimflam types were associated. We develop our protocols to promote external and internal validity as astronomically as possible while replying the RQs. still, there are still some bounds and validity risks that can be caught and mentioned also.

1. This SLR is only limited to conference and journal papers that bat engine knowledge(ML) in the terrain of detecting financial fiddle . By using our quest way in the early stands of the review, several on-relevant disquisition papers were linked and barred from this review. This ensures that the named disquisition papers satisfied the criteria for the study. Still, it's believed that using farther sources, analogous as span-new root books, would have yon enhanced this retrospection.

7. Conclusions

Financial Fraud can be committed in different financial aspects analogous as insurance, banking, taxation, and marketable precincts. recently, financial fraud has advance increasingly worrisome among companies and industriousness. Despite several labors to annihilate monetary fiddle , its durability negatively affects the economy and council as truly large amounts of capitalist are lost to fraud every day). With the appearance of artificial intelligence, machine- knowledge- predicated approaches can be used intelligently to descry crooked deals by assaying a large number of financial data. In this paper, we presented a study that completely reviewed and synthesized the being literature on ML- predicated fraud discovery. In particular, this paper espoused the Kitchenham methodology, which uses well- defined protocols to prize, synthesize, and report results. Several studies have been gathered predicated on the specified quest strategies for popular electronic libraries. After the addition/ rejection criteria, 87 were named. In this review, popularly used ML ways for fraud discovery, the most common fraud type, and the evaluation criteria are epitomized. predicated on the reviewed papers, results showed that SVM and NN are the popular ML algorithms used for fraud, and credit card fraud is the most popular fraud type in the literature. The paper ultimately presented the pivotal issues, gaps, and limitations in the area of financial fraud discovery and suggests areas for future disquisition. We linkedAppl.Sci.

2022, 12, 9637 19 of 24 gaps in the disquisition by examining unexplored or less boned algorithms. former studies in financial shell game discovery concentrated on supervised type and regression styles, analogous as SVM, neural networks, and logistic regression. The usage of ensemble modes that take high ground of multiplex algorithms to classify samples is a rising trend in the field. Interestingly, we discovered that unsupervised knowledge approaches, analogous as clustering, were less employed in the present literature. Clustering is salutary for probing idle relations and correspondences. In addition, since there are a small quantum of flimflam particular things that command to be correlated, clustering could be successful. We recommend that future disquisitions compensate complementary attention to unsupervised practices, analogous as exception discovery, which can uncover new perceptivity. also, another avenue for future disquisition would be to use arising text- mining practices and word- bedding ways analogous asWord2Vec, Doc2Vec, or BERT to transform financial handbooks into vectors of features, which will also be used to make machine knowledge models.

References

[1]. Al Barsh, Y. I., et al. (2020). "MPG Prediction Using Artificial Neural Network." International Journal of Academic Information Systems Research (IJAISR) 4(11): 7-16.

[2]. Alajrami, E., et al. (2019). "Blood Donation Prediction using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 3(10): 1-7.

[3]. Alajrami, E., et al. (2020). "Handwritten Signature Verification using Deep Learning." International Journal of Academic Multidisciplinary Research (IJAMR) 3(12): 39-44.

[4]. Al-Araj, R. S. A., et al. (2020). "Classification of Animal Species Using Neural Network." International Journal of Academic Engineering Research (IJAER) 4(10): 23-31.

[5]. Al-Atrash, Y. E., et al. (2020). "Modeling Cognitive Development of the Balance Scale Task Using ANN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 74-81.

[6]. Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.

[7]. Abu Nada, A. M., et al. (2020). "Age and Gender Prediction and Validation through Single User Images Using CNN." International Journal of Academic Engineering Research (IJAER) 4(8): 21-24.

[8]. Abu Nada, A. M., et al. (2020). "Arabic Text Summarization Using AraBERT Model Using Extractive Text Summarization Approach." International Journal of Academic Information Systems Research (IJAISR) 4(8): 6-9.

[9]. Abu-Saqer, M. M., et al. (2020). "Type of Grapefruit Classification Using Deep Learning." International Journal of Academic Information Systems Research (IJAISR) 4(1): 1-5.

[10]. Afana, M., et al. (2018). "Artificial Neural Network for Forecasting Car Mileage per Gallon in the City." International Journal of Advanced Science and Technology 124: 51-59.

[11]. Al-Araj, R. S. A., et al. (2020). "Classification of Animal Species Using Neural Network." International Journal of Academic Engineering Research (IJAER) 4(10): 23-31.

[12]. Al-Atrash, Y. E., et al. (2020). "Modeling Cognitive Development of the Balance Scale Task Using ANN." International Journal of Academic Information Systems Research (IJAISR) 4(9): 74-81.

[13]. Alghoul, A., et al. (2018). "Email Classification Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 8-14.

[14]. Al-Kahlout, M. M., et al. (2020). "Neural Network Approach to Predict Forest Fires using Meteorological Data." International Journal of Academic Engineering Research (IJAER) 4(9): 68-72.

[15]. Alkronz, E. S., et al. (2019). "Prediction of Whether Mushroom is Edible or Poisonous Using Back-propagation Neural Network." International Journal of Academic and Applied Research (IJAAR) 3(2): 1-8.

[16]. Al-Madhoun, O. S. E.-D., et al. (2020). "Low Birth Weight Prediction Using JNN." International Journal of Academic Health and Medical Research (IJAHMR) 4(11): 8-14.

[17]. Al-Massri, R., et al. (2018). "Classification Prediction of SBRCTs Cancers Using Artificial Neural Network." International Journal of Academic Engineering Research (IJAER) 2(11): 1-7.

[18]. Al-Mobayed, A. A., et al. (2020). "Artificial Neural Network for Predicting Car Performance Using JNN." International Journal of Engineering and Information Systems (IJEAIS) 4(9): 139-145.

[19]. Al-Mubayyed, O. M., et al. (2019). "Predicting Overall Car Performance Using Artificial Neural Network." International Journal o f Academic and Applied Research (IJAAR) 3(1): 1-5.

[20]. Alshawwa, I. A., et al. (2020). "Analyzing Types of Cherry Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 4(1): 1-5.

[21]. Al-Shawwa, M., et al. (2018). "Predicting Temperature and Humidity in the Surrounding Environment Using Artificial Neural Network." International Journal of Academic Pedagogical Research (IJAPR) 2(9):1-6.