Journal of Open Innovation: Technology, Market, and Complexity

Article

Systemic Acquired Critique of Credit Card Deception Exposure through Machine Learning

Rui Miguel Dantas [1], Raheela Firdaus [2,*], Farrokh Jaleel [3], Pedro Neves Mata [1], Mário Nuno Mata [1] and Gang Li [2]

[1] Lisbon Accounting and Business School, The Polytechnical Institute of Lisbon, 1069-035 Lisbon, Portugal
[2] Department of Management, School of Management and Economics, North China University of Water Resources and Electric Power, Zhengzhou 450002, China
[3] Department of Mechanical Engineering, International Islamic University, Islamabad 04403, Pakistan
* Correspondence: firdausraheela@gmail.com

Abstract: A wide range of recent studies are focusing on current issues of financial fraud, especially concerning cybercrimes. The reason behind this is even with improved security, a great amount of money loss occurs every year due to credit card fraud. In recent days, ATM fraud has decreased, while credit card fraud has increased. This study examines articles from five foremost databases. The literature review is designed using extraction by database, keywords, year, articles, authors, and performance measures based on data used in previous research, future research directions and purpose of the article. This study identifies the crucial gaps which ultimately allow research opportunities in this fraud detection process by utilizing knowledge from the machine learning domain. Our findings prove that this research area has become most dominant in the last ten years. We accessed both supervised and unsupervised machine learning techniques to detect cybercrime and management techniques which provide evidence for the effectiveness of machine learning techniques to control cybercrime in the credit card industry. Results indicated that there is room for further research to obtain better results than existing ones on the basis of both quantitative and qualitative research analysis.

Keywords: algorithms; credit card; database; fraud; financial organization; machine learning; dataset

check for updates

Citation: Dantas, R.M.; Firdaus, R.; Jaleel, F.; Neves Mata, P.; Mata, M.N.; Li, G. Systemic Acquired Critique of Credit Card Deception Exposure through Machine Learning. *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 192. https://doi.org/10.3390/joitmc8040192

Received: 23 August 2022
Accepted: 11 October 2022
Published: 21 October 2022

1. Introduction

Online services are more common in the new age. A major contributor of online services is in the banking industry, where services are shifting more towards digital banking. Digital banking provides an ease to the customer to transfer funds, online shopping and many other activities [1]. This digitization is not only fruit for the customer but banks are also facilitated with this change [2]. Cybercrime became more common and severe due to transfer of manual work into digital. Cyber security is now capturing more attention both in industry and research. Cyber-criminal activities are increasing at fast speeds and becoming more sophisticated. This paper is designed to present a systematic literature review of machine learning techniques and analysis techniques in management science to lower cybercrime. Cybercrime is a major challenge to many financial organizations. The impacts of cybercrime are dangerous to an organization's assets [3]. Cybercrime raised globally by 50% from 2018 to 2020 which is equal to an amount of under USD 1 trillion [4].

Fraud is an act of obtaining financial benefits or services illegally by wrong or criminal deception. Credit card fraud is the unauthorized way of performing transactions without the approval of a card's owner. There are many types of credit card fraud, such as application fraud and merchant fraud. Credit card fraud can be categorized in a number of ways, as bankruptcy, theft or stolen cards, counterfeit fraud and application fraud as said by [5]. Online financial fraud has gained prominence in the past few years to comply with the growing electronic commerce economy. A lot of fraud prevention strategies have been

designed to prevent this offence and yet the statistics provided by global fraud reports suggest that more research is needed in the area of credit card fraud.

Moreover, financial crimes are increasing every day and no organization is immune from being victims of fraudsters [6]. Financial organizations are under serious threat of online fraud due to an increase in the usage of electronic commerce. Credit card transactions are conducted by using cards physically, mostly at point of sale, and by internet, where digital transactions are processed with the details of a card and card holder [7]. Plastic cards are used for cashless payments to buy goods and services. Everyone can shop remotely by using credit cards from home. Banks and financial firms that deal with plastic cards should pay attention to this hot issue, which not only minimizes their revenue but is also dangerous for their reputations [8]. Electronic transactions have raised significantly due to the increase in trends of online shopping from e-businesses, such as Amazon, eBay, and Ali-baba [9].

In addition, a very innovative and interesting sub field of computers is artificial intelligence algorithms. These algorithms learn from data with little human intervention. The two major branches of machine learning are supervised and unsupervised learning. Machine learning techniques have successful application in many fields of life, such as weather prediction, computational intelligence, medical, and fraud detection. Machine learning algorithms are playing a vital role in overcoming cybercrime in the finance sector [10]. Credit card fraud detection is the best way to analyze the efficiency and effectiveness of machine learning algorithms. Many single and hybrid approaches of machine learning are applicable to the decline in credit card fraud [11].

The credit card fraud detection problem has a number of challenges, such as variation in a card holder's spending nature, high number of genuine and low number of fraudulent transactions, which is known as class imbalance, where a small number of transactions are timely checked by investigators [12].

This research aims to perform a systematic literature review on cybercrime from the perspective of machine learning, datasets used for analysis, performance metrics and data collection methods in finance and more focused towards the credit card industry. The study reviews current research articles from five famous databases. This is the first systematic literature available on cybercrime in finance with two dimensions, one is machine learning with details of datasets and second is analysis methods and data collection techniques with performance metrics. The previous studies only focused on one dimension. This paper is helpful for the researchers as it supports categorizations and aggregation of literature which can lead to a clear way forward for future research.

2. Literature Review

Fraud is an unauthorized access to a financial account either for personal benefits or financial loss. Financial fraud has captured attention with far reaching significance in the government, corporate administrations, and financial firms. Millions of pounds are lost each year in e-commerce due to credit card fraud. For instance, global brands suffered from USD 18.30 billion loss in credit card fraud in 2015. This value increased to USD 20.18 billion in the year 2016. For cards issued in the euro area only, the total value of fraudulent card transactions amounted to €1.03 billion. Only in Europe the card frauds in 2019 were 1.03billion in European currency [13]. A vital element is to achieve maximum success against opponents by using minimum resources. The rapid increase in the rate of credit card use is one of the main reasons that increases the number of financial fraud occurrences every year. Hence after, both the client and issuer of credit card suffer from heavy loss due to credit card fraud, and this situation will become worse in coming years [14]. Fraud detection is a method of monitoring the behavior of a user in order to avoid unwanted conditions in credit card transactions. Fraud detection is not only concerned with the uncovering of such unwanted activities, but also making decisions timely and quickly to avoid damage [15,16]. Recently, many fraud detection solutions

and software have emerged to prevent fraud in businesses, including credit card, retail, e-commerce, insurance, and industry.

Machine learning techniques are one of the prominent and popular approaches used to solve the problem of credit fraud detection. Authors in [17] have categorized ML techniques as supervised and unsupervised algorithms. SML algorithms require training data to build a model for the test data. These algorithms are very popular in credit card fraud detection but are difficult to implement and train. On the other hand, unsupervised learning techniques do not require training data, as these algorithms learn from input data without labeling; however, there is a problem of data storage while using unsupervised learning [10]. The third approach used by the researchers is hybrid ML algorithms. Hybrid learning, also known as blended learning, is a combination of algorithms to improve efficiency, but it is tricky to work it with advanced mining applications.

A vital part of cybercrime research in the banking sector is played by the datasets. A number of datasets are explored with deep investigation [18]. Digital services in the banking sector and cybercrime were the main catalysts of this study. Cybercrime committers are becoming smarter, so to control these crimes they should be detected efficiently and effectively in rum time. Research on the need for updated and more recent datasets was also performed [19]. The importance of datasets was also highlighted [20] by examining anomaly detection techniques. Another study focused datasets and their role with supervised machine learning algorithms to elaborate their importance in the training of algorithms [21]. Datasets were also tested on the efficiency of machine learning algorithms [22].

3. Research Gaps

The main research gaps filled by this article are summarized as follows. Firstly, this article focused not only on which machine algorithms are used by most researchers, but also the performance metrics. Secondly, the research focuses on the datasets used by previous researchers and proposes new issue for real time datasets which can be more beneficial for researchers for analysis, which will ultimately be helpful for organizations to control cybercrime.

4. Research Motivations

A vital element is to achieve maximum success against opponents by using minimal resources. Every economy is using its own tools to mitigate cybercrime in the financial sector, but the results are almost the same. It is not very important that an economy have a lot of resources, but the important factor is efficient and effective utilization of these resources against cybercrime [8]. Cybercrime in the credit card industry is growing with every sun rise. To overcome the threat on card holders and on merchants, financial organizations are trying to design automatic, efficient and user-friendly monitoring systems [5]. There is a need for satisfactory mitigation factors against cybercrime in finance [23].

Now, it is necessary to analyze new methods, trends and counter-measures in the credit card cybercrime detection era to empower the systems that are working on credit card fraud. For smoothness in remote financial transactions, to guarantee the authenticity of credit card transaction and to shield the privacy of card holders, there is a need for more investigation [24]. Machine learning techniques enable efficient cybercrime detection, primarily with real-world datasets. These techniques are helpful to capture unauthorized transactions in a timely and effective way, which will ultimately reduce the amount of loss occurring every year in financial organizations due to cybercrime [11].

Research Objectives

Based on the research motivations as we discussed above, to shed some light on applications and successes of machine learning algorithms in credit card fraud detection, we have systematically analyzed previous literature on machine learning algorithms in cybercrime detection in the area of finance. The aim of this study is to provide the latest knowledge for financial organizations. The research work in this article will be helpful

for developers and organizations to design and implement more efficient and effective monitoring tools for cybercrime detection in the credit card industry.

The call of articles by some high impact journals is evidence for the need of more research on this topic. Subsequent research queries are:

RQ1. What were the major algorithms that captured the attention of previous researchers?
RQ2. What were the research objectives in the area of cybercrime in credit cards?
RQ3. What were the performance metrics used in most research papers?
RQ4. What were the data collection techniques used by the researchers?
RQ5. What are the keywords of the articles, either related to our research or not?
RQ6. What research gaps can lead to new research in the future?

To explain these queries the authors performed a synthetic literature assessment dedicated to machine learning and cybercrime in credit cards. Research articles were collected over ten years of publications from five major databases and then an extraction process was conducted on up to date papers.

5. Methodology

This section will enlighten our methodology for this paper. The main objective of this investigation is to systematically scrutinize the newfangled research on cybercrime in credit cards and machine learning algorithms, their key elements and relationship studies that are up to date. The purpose and effect of the studies are examined in a systematic manner to accomplish the goal. A systematic literature review is a trustworthy, transparent and applicable way of collecting and examining large amounts of evidence from previous research. Structured content analysis, according to [25], is used in this research. A scientific technique to assess the themes of communication is known as structured content analysis. It is very helpful in providing a comprehensive literature review.

Figure 1 explains our proposed methodology. As it indicates, we started with a selection of articles from five English databases and then our search was filtered from 2010 to 2019, so that the authors gathered updated and new articles. Authors applied screening on search, with a focus on full length articles. After removing duplication, abstracts of articles were analyzed to obtain closer articles to this study. Then, by applying RQ1, RQ2, RQ3, RQ4 and RQ5, further articles were refined.

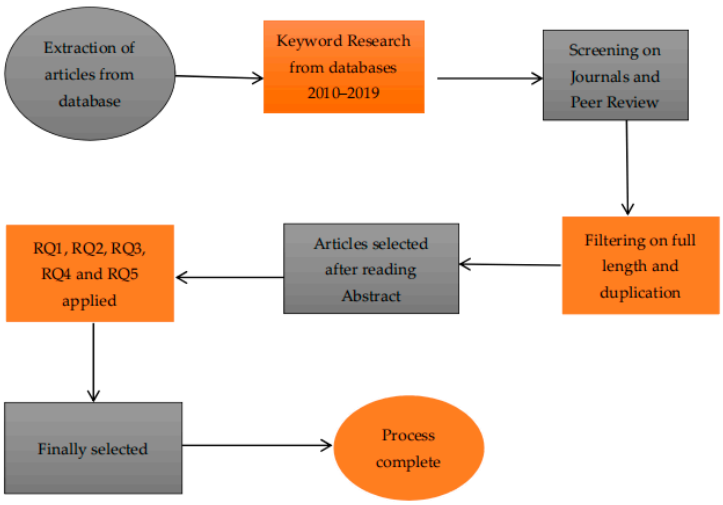

Figure 1. Detailed methodology.

Furthermore, four main steps recommended by [25] to lead a literature review were also applied in this research. These steps are: data gathering, descriptive investigation, identification of category and examining data. These steps are followed and then elaborated in the results of the findings.

5.1. Data Collection

This section describes the data to be examined and also components of examination. Articles included in this research work are important studies published in the English language. For this purpose, authors have selected five databases, named as Emerald Insight, IEEE, Springer, Science Direct and Wiley Library. Further research is refined by the keywords which are most appropriate for our paper. The time span included in this research is 10 years, from 2010 to 2019. The details of the keywords, followed by our research, are shown in Table 1.

Table 1. Structured keywords.

Database	Keywords
Emerald	Financial fraud (only this is performed with Springer), cybercrime, cybercrime in finance, online credit card fraud, artificial intelligence in financial fraud, uncovering fraud by machine learning, plastic card fraud detection, fraud investigation in credit cards by machine learning.
IEEE	Financial fraud, cybercrime, cybercrime in finance, online credit card fraud, artificial intelligence in financial fraud, uncovering fraud by machine learning, plastic card fraud detection, fraud investigation in credit cards by machine learning.
Springer	Financial fraud, cybercrime, cybercrime in finance, online credit card fraud, artificial intelligence in financial fraud, uncovering fraud by machine learning, plastic card fraud detection, fraud investigation in credit cards by machine learning.
Science Direct	Financial fraud, cybercrime, cybercrime in finance, online credit card fraud, artificial intelligence in financial fraud, uncovering fraud by machine learning, plastic card fraud detection, fraud investigation in credit cards by machine learning.
Wiley	Financial fraud, cybercrime, cybercrime in finance, online credit card fraud, artificial intelligence in financial fraud, uncovering fraud by machine learning, plastic card fraud detection, fraud investigation in credit cards by machine learning.

5.2. Descriptive Analysis

In descriptive analysis, assessment of formal characteristics is performed to provide history for the evaluation of each study. The attributes that are included in this research paper are title, publication date, author's details, techniques used in the paper, purpose of the article, journal, datasets used in the article, performance metrics used by the authors and future directions of articles.

5.3. Category Identification

As consistent with the previous literature methodology, articles are selected with supervised, unsupervised machine learning techniques, regression and correlation models, and some literature review articles. Researchers also examined articles with data collection both with the real time datasets and with secondary, as datasets are very important in experiments and evidence gathering. Mostly, studies are based on performance of machine learning algorithms in fraud detection and some studies are survey based.

5.4. Material Evaluation

In this section the articles selected from previous sections are identified. The main focus of each article is the use of machine learning algorithms, supervised or unsupervised, single or hybrid approach, in fraud detection during online credit card transactions. Research focused from the very general broad concept of financial fraud and then became narrower towards credit card fraud. All the results are combined from articles into this single study.

5.4.1. Selection of Articles

To achieve fine results, inclusion and exclusion strategies are applied in this article, as shown by Figure 2.

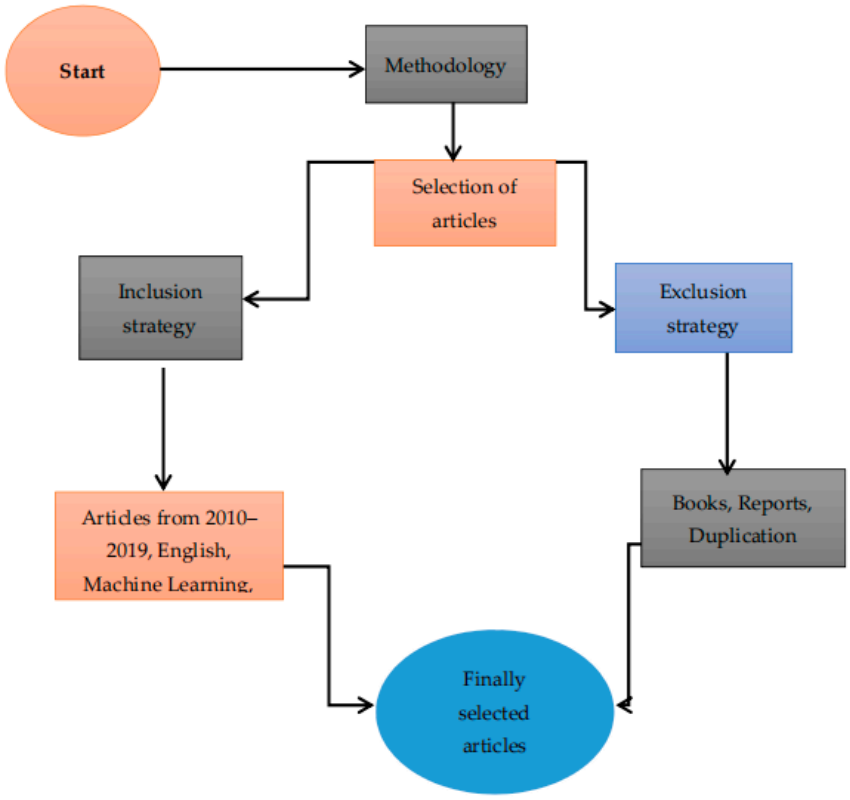

Figure 2. Selection of articles.

5.4.2. Inclusion Strategy

Research papers published from January 2010 to January 2019 are collected for the synthesis of literature. Manuscripts not published until January 2019 are not included. Articles focusing on financial fraud and then articles addressing fraud detection through machine learning are included. Only articles published in the English language on both addressing management and machine learning are selected.

5.4.3. Elimination Strategy

Books, technical reports and working papers are excluded in the review to maintain the quality of literature. After reading abstracts and conclusions the articles that do not address cybercrime in the financial sector and fraud detection through machine learning are also eliminated from this research.

6. Results

This segment illuminates the outcomes after conducting descriptive investigation, data assessment and category identification. Citations, year of publication, and methodology for study are highlighted by the descriptive investigation.

Structured keywords are used in the search process as shown by Figure 3. The articles from 2010 to 2019 are included at the start of search; books, magazine, reports, and articles are all from Emerald (6689), IEEE (1171), Springer (46,917), Science Direct (24,845) and Wiley (33,995). Filters are applied to this large amount of data by only selecting articles and after that article details are as Emerald (4458), IEEE (1143), Springer (11,424), Science Direct (14,991) and Wiley (14,503). Applying the extraction process again by up to date, full length articles, by RQ1, RQ2, RQ3, RQ4, RQ5 and RQ6, after removing duplication, 25 articles are selected from each database. Again, screening is conducted by reading the abstracts of all 125 articles and then 10 articles are selected from every database that are more appropriate to this topic.

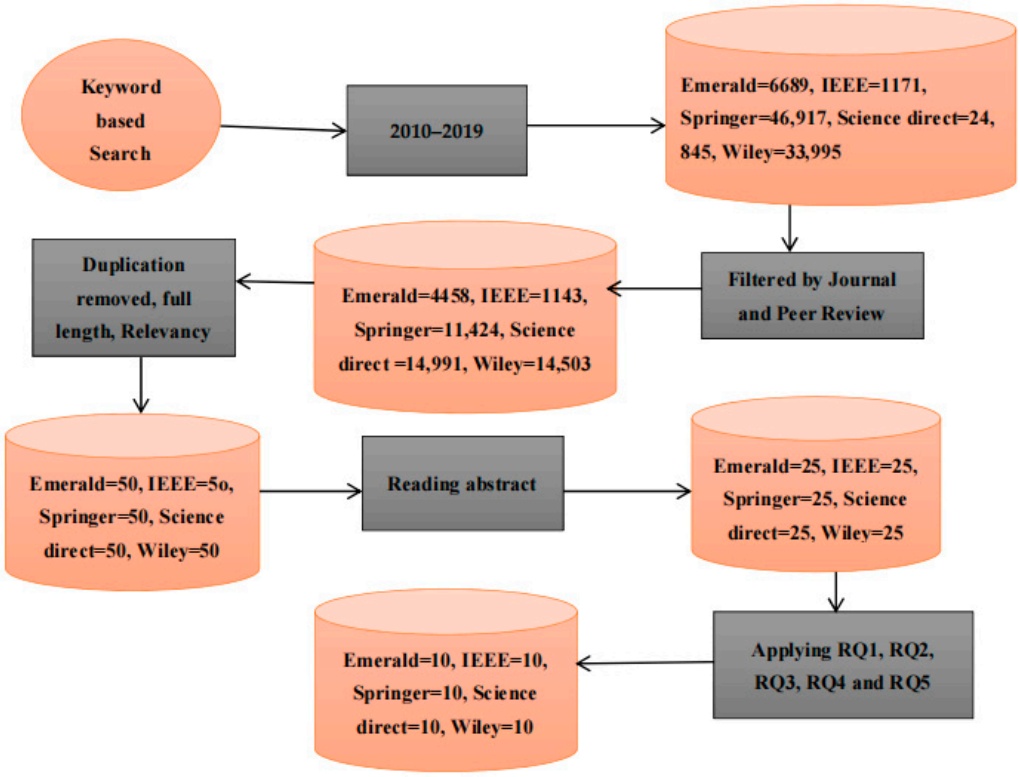

Figure 3. Detailed process.

6.1. Analysis Based on Frequency of Publication

The section elaborates the results from descriptive analysis, publications over time span and database. Figure 4 clarifies that the Wiley database has published higher numbers of research articles addressing the issue of credit card fraud and machine learning, Science Direct falls second in publication frequency. The third database on number of publications is Springer, and fourth is Emerald. IEEE has the lowest number of publications.

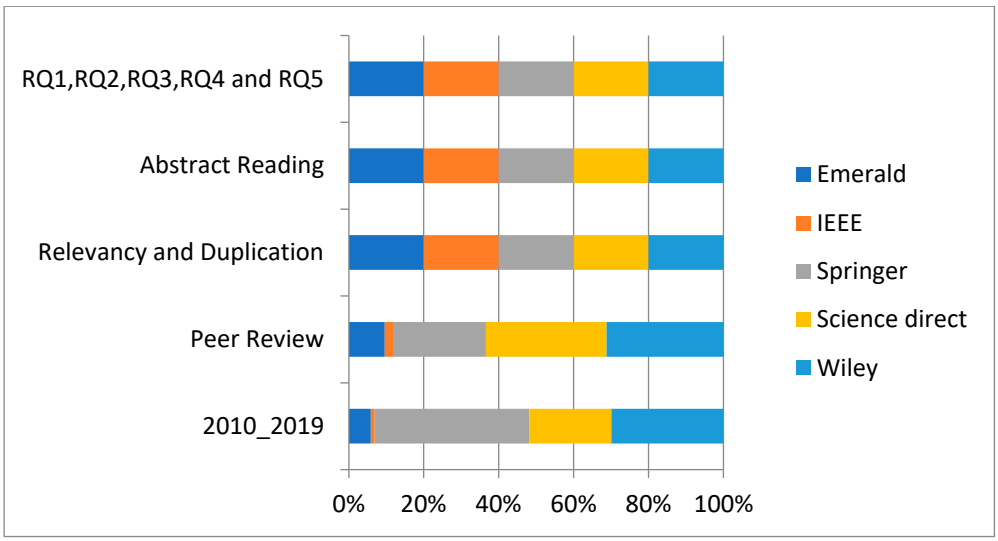

Figure 4. Detail of publication frequency.

Figure 4 signifies the graphical representation of articles in five databases from the year span of 2010 to 2019.

6.2. Analysis Based on Techniques

The techniques used in previous research, title and year of research and details of researcher are highlighted in these sections.

Table 2 provides information about the most commonly used machine learning techniques. Supervised machine learning techniques were used by majority of researchers. The frequently explored machine learning techniques in the area of fraud detection are neural network, PCA, support vector machine, Naïve Bayes, adaptive boosting, random forecast, artificial immune system, artificial neural network, decision tree, genetic algorithm, k-nearest neighbor, bagging, Multi-layer Perceptron and Hidden Markov model. Single and hybrid approaches were used to find fraud more effectively and efficiently.

Table 2. Techniques in previous research.

Title and Year	Author	Techniques
Neural networks: the panacea in fraud detection [6]	Maria Krambia-Kapardis	Neural networks
Building our defence against credit card Fraud: a Strategic view [8]	Hendi Yogi Prabowo	Crime triangle, historical and bench marking analyses
Parameters of automated fraud Detection techniques during online transactions [5]	Vipin Khattri, Deepak Kumar Singh	Literature review
"Predicting susceptibility to cyber-fraud victimhood" [26]	M.T. Whitty	Exploratory Factor Analysis, bi-variate associations
"Predicting fraudulent financial reporting using artificial neural network" [27]	N.Omar, Zulaikha 'A.Johari, M.Smith	AAN, fraud triangle theory
Factor analysis of financial crime-related issues [23]	G. Babatunde et al.	PCA
A hybrid firefly and support vector machine classifier for phishing email detection [28]	O.A. Adewumi et al.	FFA with SVM
Top 10 data mining techniques in business applications: a brief survey [29]	W.C. Lin et al.	Regression, DT, NN, K-NN, MLPNN, NB, SVM, K-MEANS, C4.5
Influential factors of online fraud occurrence in retailing banking sectors from a global perspective An empirical study of individual customers in the UK and China February [30]	Y.Sun, I. Davidson	Correlation, demographic data analysis
Analysis on the new types and countermeasures of credit card fraud in mainland China [24]	F. Bai, X.Chen	Literature survey
Resilient Identity Crime Detection [31]	C.Phua et al.	CD and SD
Anomaly Detection via Online Oversampling Principal Component Analysis [32]	Y.J. Lee, et al.	OsPCA
Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy [12]	A.D. Pozzolo et al.	Random forest
Credit Card Fraud Detection Using Ada Boost and Majority Voting [11]	K.Randhawa et al.	AdaBoost and majority voting
Comparison with Parametric Optimization in Credit Card Fraud Detection [33]	M.F. A.Gadi, et al.	NN, BN, NB, AIS and DT
CoDetect: Financial Fraud Detection With Anomaly Feature Detection [34]	D. HUANG et al.	Graph mining techniques.
Transaction Fraud Detection Based on Total Order Relation and Behavior Diversity [9]	L. Zheng, et al.	Logical graph
Machine Learning- and Evidence Theory-Based Fraud Risk Assessment of China's Box Office [35]	S.QIU 1,2 AND H.Q. HE1	Regression models.

Table 2. *Cont.*

Title and Year	Author	Techniques
Credit Card Fraud Detection Using RUS and MRN Algorithms [36]	A.Charleonnan	AdaBoost.M1, RUS, MLP, NB, MRN
Fraud Detection in Big Data using Supervised and Semi-supervised Learning Techniques [37]	G.E. M.Acosta et al.	Balanced random forest
A survey of machine-learning and nature-inspired based credit card fraud detection techniques [38]	Akinyelu A.O.	Hidden Markov model, NN, meta-learning, SVM, frequent item set mining, ANN, Bayesian network and neural network, decision tree and logistic regression, R-frequency and time-dependent score, bagging and ensemble, transaction aggregation logistic regression, genetic algorithm, artificial immune system
Payment Card Fraud Detection Using Neural Network Committee and Clustering [39]	A. S. Bekireva et al	Single neural network, neural network committee
A machine learning based approach for phishing detection using hyperlinks information [40]	A.K Jain B. B.Gupta1	LR, NB, RF, SVM, NN, C4.5, SMO
Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance [41]	A.Somasundaram	SMOTE, bagging
Effective detection of sophisticated online banking fraud on extremely imbalanced data [42]	W.Wei et al.	Neural network, decision forest
Fraud detection within bankcard enrollment on mobile device based payment using machine learning [43]	H.ZHO et al.	LR, RF, GBDT
Hybrid Approach for Improvising Credit Card Fraud Detection Based on Collective Animal behavior and SVM [44]	V. Dheepa and R.Dhanapal	SVM, collective animal behavior approach
Logistic Regression Learning Model for Handling Concept Drift with Unbalanced Data in Credit Card Fraud Detection System [45]	P. Kulkarni and R. Ade	MLPNN, back propagation algorithm, LR
Scalable Machine Learning Techniques for Highly Imbalanced Credit Card Fraud Detection: A Comparative Study [46]	R.A. Mohammed et al.	RF, BE, GNB
Neural Network Rule Extraction to Detect Credit Card Fraud [47]	N.F. Ryman-Tubb and P.Krause	Neural network, S Oracle-based adaptive algorithm
End-to-end neural network architecture for fraud scoring in card payments [48]	J.A. Gómez a et al.	Artificial neural networks
Financial fraud detection using vocal, linguistic and financial cues [49]	C.S. Throckmorton et al.	LR, NB, KNN, GLRT
A data mining based system for credit-card fraud detection in e-tail [50]	N.Carneiroa et al.	Random forests, support vector machines, logistic regression
Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments [51]	C.C. Lin et al.	LR, DT, ANN
A customized classification algorithm for credit card fraud detection [52]	A.C. de Sá et al.	BNC

Table 2. *Cont.*

Title and Year	Author	Techniques
Feature engineering strategies for credit card fraud detection [53]	A.C. Bahnsen et al.	LR, DT, RF, Bayes
Sequence classification for credit-card fraud detection [54]	J.Jurgovsky et al.	RF and the LSTM7
Application of credit card fraud detection: Based on Bagging ensemble classifier [55]	M.Zareapoor, P.Shamsolmoali	Decision tree algorithms
Some Experimental Issues in Financial Fraud Mining [56]	J.West1 and M.Bhattacharya2	GP1-2, GA1-2, FL, ACO, NN1-2, SVM, DT1, DT2, Fn, LAZY, RULE
ConvNets for Fraud Detection analysis [57]	A.houiekha *, H.Ibn EL Hajb	SVM, Conventional Neural Network, random forest, gradient boosting
Isolation-based anomaly detection using nearest-neighbor ensembles [58]	T. R. Bandaragoda1 et al.	INNE (Isolated Neig. Neigh) Science Direct
Hybrid approaches for detecting credit card fraud [59]	Y. Kültür. M. U.Çağlayan	DT, RF, BN, NB, SVM, K-models
A systematic review on intrusion detection based on the Hidden Markov Model [60]	A.A.R.Abbaset al.	Hidden Markov models (HMM)
Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China [61]	X.P.SONG et al.	BPNN, LR, SVM, C5.0,DT
Predicting credit card delinquencies: An application of deep neural networks [62]	T.Sun1 M. A. Vasarhelyi2	LR, NB, ANN, DT, NN
Machine learning methods for detecting patterns of management fraud Number [63]	D.g.whiting et al.	RF, GB, RE
Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results [64]	N. Wong et al.	AIS
An overview of the use of neural networks for data mining tasks [65]	F. Stahl∗ and I. Jordanov	Neural networks
Imbalanced SVM-Based Anomaly Detection Algorithm for Imbalanced Training Datasets [66]	G. Wang et al.	SVM
Advancing the assessment of automated deception detection systems: Incorporating base rate and cost into system evaluation [67]	D.P. Twitchell, C.M. Fuller	Bayes' theorem

6.3. Analysis Based on Purpose and Datasets

The purpose of research is the main focus while reading any study. In this section, purpose of previous research is provided in a very comprehensive way. Datasets play a vital role in the findings of any research. Here the datasets are the main focus, and they were taken from primary or secondary sources for the analysis of credit card fraud detection.

Table 3 elaborates the aim of the previous researchers and data used by them for analysis. In the field of fraud detection, real world data are more important, but it is difficult to obtain real fraud data from any organization. The main objectives were to deal with challenges, such as dealing with class imbalance, to find the most efficient techniques, handling of complex problems, customer behaviors and feature engineering. Primary as well as secondary data were used for experiments to provide evidence on the role of machine learning methodologies to solve the problem of financial fraud. Primary data were the concern of the majority of researchers.

Table 3. Purpose and datasets in previous work.

Purpose	Datasets
The aim of [6] is to find out the role of ANN to distinguish credit card deceptions.	Primary data
The main purpose in [8] is to examine developments in the prevention of credit card crimes in the USA, UK, Australia and Indonesia over the time domain of 2003–2007 and practical implementation of systems to prevent fraud.	Primary and secondary data
The objective of [5] is to facilitate researchers with a clear vision towards the designing of online fraud detection systems.	Secondary data
The main focus of [26] is to provide a theoretical framework to forecast the vulnerabilities of online fraud victimhood.	Primary data
The efficiency of artificial neural networks to forecast deceitful activities in financial writing was inspected [27] in small size firms in Malaysia.	Secondary data
Factor analysis was performed [23] to list the issues related to financial crimes in Nigeria. Data were collected through questionnaire from six geopolitical zones of Nigeria.	Primary data
Nature inspired base machine learning techniques were explored and reported [28] to detect phishing emails.	Primary data
The main objective of [29] was to find out the most widely used data mining methodologies in business applications.	Secondary data
Online fraud was focused [30] in the retailing banking sectors of two countries, the United Kingdom and the People's Republic of China.	Primary data
An investigation was conducted in [24] to explore new countermeasures and to adapt comprehensive methodology against credit card fraud in China.	Secondary data
Two new layers were introduced in [31], where fraud detection system one is communal detection and system two is spike detection. The first one detects the social relationship while the other detects duplicates to increase the suspicion score.	Real time primary data
An oversampling technique was proposed in [32] to detect outliers by using principal component analysis. In new techniques there is no need to store entire datasets and the new proposed technique is better with large scale problems.	Primary data
Most suitable performance measures were illustrated [12] for fraud detection. The two most common problems of inequality between two classes and concept drift were discussed in [12]. The effect of these issues was measured on data by real transactions.	Real world data
Two famous methods used in combination are AdaBoost and majority voting for credit card fraud detection processes [11] A publicly available dataset was used for evaluation of single and hybrid models.	Primary data
Apart from Naïve Bayes, other classification techniques perform better with cost sensitivity. Decision table and artificial immune system are the best methods according to [33].	Real world data
Network and feature information were used to propose a novel fraud detection framework [34]. The proposed framework can simultaneously recognise suspicious activities and use pattern recognition for the features associated with fraud activities.	Artificial and real data
To represent a logical relationship between features of transactions, a logical graph was proposed. Entropy diversity coefficient was calculated to measure the variation in transactions and path-based transition probability was implemented.	Primary data
A framework was developed for fraud risk assessment based on evidence theory. Logistic regression was utilized to measure the probability for evidence theory. Fraud risk factor was put forward by [35].	Real world data
Information related to customer behavior during credit card transactions was collected in Taiwan to analyze the prediction for the efficient detection of risk in credit card payment [36] techniques	Primary data
To deal with three main problems in credit card fraud detection, namely class imbalance, labeled and unlabeled samples and capability to process large datasets. A framework was proposed [37] which can handle all these issues.	Primary data
Some restrictions and achievements of credit card fraud detection methods were in analyzed in [38]. It provides essential knowledge for researchers in the domain of credit card deception.	Secondary data
Fraud detection can be performed by examining current and previous attributes of the same card [39].	Secondary data

Table 3. *Cont.*

Purpose	Datasets
A new methodology was proposed [40] that can detect attacks by analyzing hyperlinks of websites. This approach was compared with classification algorithms.	Primary data
The problem of financial crimes was fixed [41] by suggesting incremental learning and transaction window bagging.	Real world data
Various advance data mining techniques were incorporated and relevant information synthesized by the framework proposed [42] for effective fraud detection in the banking sector.	Primary data
Improved gradient boosting decision tree was used with real data to improve the detection of fraud [43].	Real world data
A hybrid approach was proposed in [44] for uncovering credit card deception with the combination of clustering and classification techniques.	Real data
A universal framework using logistic regression was projected in [45] that can tackle issues related to learning for the evaluation of credit cards.	Secondary data
The usability of various machine learning techniques was evaluated [46] as scaleable algorithms to handle the problem of class imbalance data.	Primary data
In mission critical areas of business, it was demonstrated with experiments [47] that neural networks can perform better and be a good tool for transparent fraud detection.	Real world data
The main focus of [48] was on artificial neural network to solve the problem of online crime detection.	Primary data
An important information for financial fraud detection can be provided by numerical data of finance, linguistic data and non-verbal data [49].	Real time data
Integrating the manual and programmed classification provides more knowledge about design and compares various techniques that are based on machine learning. A complete system was designed and implanted for risk scoring by [50].	
All features of the fraud triangle theory, including incentives, opportunity and rationalization, were analyzed by using data mining techniques. The second goal was to reveal whether experts agree (or not) with results by novel techniques. Authors in [51] used questionnaires and data mining techniques.	Real time data
Results of the study [52] were compared with another seven algorithms and the techniques were examined for classification problems. The firm's economy was improved up to 72.64%.	
With the use of Von Mises Distribution a new group of characteristics were suggested [53]*, 2016) by improved transaction strategy and with the analysis of periodic behaviour.	Primary data
Long short-term network was implemented [54] to collect transaction behaviours for the improvement in accuracy to catch online financial crimes.	Real world data
Three main methods were used to assess the methodology suggested in [55] for credit card fraud discovery.	Primary data
Some issues related to experiments, such as a focus on detection technique, attribute assortment and performance metrics, and skillful simulations in credit card fraud detection, were examined in [56].	Primary data
Various experimental techniques were used to assess the performance of the model proposed in [57] for the detection of genuine and fraudulent activities in mobile communication.	Primary data
Four vulnerabilities were identified in [58], as incapability to find anomalies, anomalies with irrelevant features, masked anomalies and multi model anomalies. An alternative mechanism was proposed by using the nearest-neighbor ensemble.	Primary data
By using ensemble techniques and three new methodologies, OPT, PES and WGT, a model was suggested [59] to overcome the issue of credit card fraud.	Real world data
Evaluation of merits and boundaries of architectural models and application were discussed in [60]. Six models from literature were used to choose the exact type for a specific application.	Secondary data
Outcomes in [61] elaborated risk factors and a rule-based system to help overcome error rates.	Real world data
Two main contributions of networks [62], first to develop an efficient system for credit card fraud prediction and second to evaluate neural network in credit card fraud detection.	Primary data

Table 3. *Cont.*

Purpose	Datasets
New developed statistical techniques were explored in [63] to handle complex problem domains and for efficient detection of fraud in the credit card domain	Secondary data
For effective security management, the main focus of [64] was artificial immune system. The solution was evaluated by a case study of fraud in credit card transactions.	Real world data
The main considerations of [65] were both supervised and unsupervised techniques for the implementation of neural networks in credit card deceptions.	Secondary data
To assign different weight to positive support vector, new imbalanced support vector machine based suspicious activity detection was proposed in [66].	Primary data
To illustrate the importance of contextual information, both theoretical and experimental evidence was provided by [67].	Real world data

6.4. Analysis Based on Performance Measures and Future Directions

This section throws light on the performance measures adopted by scholars in the past and will suggest directions for the next research.

Table 4 explains the performance measures and future directions in the area of credit card fraud detection. Performance evaluation was conducted by examining accuracy, true positive rate, false positive rate, true negative rate, F-Measure, recall, precision, mean, median, mode, standard deviation, error rate, area under curve, alert rate, beta and by value of P. The future directions suggested by the researchers are base rate, classifier performance, cost, class imbalance, knowledge extractions, integration of different techniques, new strategies of feature engineering, and techniques that can be examined with multi-dimensional data, focusing on execution time to identify fraud.

Table 4. Performance measures and future directions in previous studies.

Performance Measure	Future Work
Precision (%) [6]	ANN had achieved high accuracy in fraud predictions. If the same parameters were used as [6] then accuracy can be 95%. ANN can save audit costs.
New developments in credit card fraud prevention [8]	Four pillared house of fraud prevention was not able to provide complete standards for real time fraud prevention systems. Further research can be conducted to find more extensions of the house to overcome the likelihood of fraud [8].
Regularity of constraints used in fraud detection [5]	More research can be performed for trade-off time consumption to detect fraud. By using parameters [5], classier implementation of a system to capture fraud can be a future piece of work.
B, SE, b, t, p [26]	More studies can be performed to evaluate routine tasks in details and how distinguishing can be achieved between fraudulent and non-fraudulent content [26].
Standard error of approximation, R, R2, Adjusted R2 [27]	The prediction model used in [27] can be compared with other fraud detection techniques as future research. More research is needed to discover better indicators for risk and the need for a new, trustworthy model.
N mean SD, component Total % of cumulative variance [23]	Further research can be conducted on control measures suggested by [23] which are necessary to control financial crimes.
Accuracy, false positive rate, false negative rate [28]	New features can be introduced and more techniques can be investigated in future research which can provide good results [28].
Accuracy, techniques used in the eight application areas [29]	There is room for research based on new parameters and techniques for example pattern identification in textual data and context analysis [29].
Correlation significance value hypo [30]	According to [30] more research can be carried out with merchants and businesses which use online transactions and are suffering from fraud on a daily basis.

Table 4. *Cont.*

Performance Measure	Future Work
Analysis of new trends (theoretical) [24]	Analysis of new methodologies and techniques in credit card delinquencies is needed, so that an improvement can be seen to handle this issue timely and smoothly, which can protect clients from being victimized by fraudsters [24].
F-Score, ROC and value of threshold [31]	The main boundaries were pointed out in [31] were imbalanced class and time limitations.
ROC curve (AUC), time, TP, FP in [32]	Principal component analysis might not very good for the estimation of principal directions in [32] with the high dimensional data. Thus, further other techniques can be analyzed with high dimensions.
Alerts raised by the FDS (Pk, CPk, AUC), sum of classifiers' ranks [12]	The employment of knowledge on ranked approach can be considered as future work [12] that can be exactly intended to replace linear accumulation of forecasting.
Accuracy, sensitivity, MCC [11]	Techniques that were discussed in [11] can be extended to online learning models and more models can be explored on the base of learning. Fast fraud detection can be made possible by using online learning models which will help the financial sector to reduce fraudulent transactions.
Cost function [33]	By focusing on cost sensitivity and skewness of the data further in depth, optimization of attributes can be performed as a future work of [33]. SVM and other techniques can be considered in the pool of comparison.
Detection accuracy, time with different rank size [34]	Tasks related to finance can be represented as similarity and feature tensors [34]. Future studies can focus on how to combine tensors into a co detect framework for better fraud detection.
TPR and FPR, precision, recall, F-Measure, AUC, time of transaction, accuracy [9]	To describe users' spending behavior more accurately in models [9], machine learning algorithms can be focused. Models can also be extended on users' feedback.
p-value [35]	The approach designed in [35] can be applied in fraud detection.
Accuracy, sensitivity, specificity [36]	Other performance measures can be explored [36].
TP, FN, TN, FP, Acc, G-mean, wtdAcc, area under the ROC curve AUC [37]	Proposed strategy [37] based on a meta-classification approach can be applied to datasets of different sizes to observe the results.
Accuracy, AUC, FP rates, false negative rates, classification speed [38]	Class imbalance problems with many fraud detection algorithms can be explored [38].
ROC, TPR, TNR, TPR + TNR [39]	Further legitimate models could be evaluated to analyze usage of credit cards [39].
True positive, true negative, precision, F1 measure, accuracy [40]	Non-HTML websites can also be considered for detection with high accuracy in future work [40].
Area under curve (AUC), FPR, TPR, TNR, FNR [41]	To improve efficiency and effectiveness of feature engineering, strategies can be proposed as a future work [41].
Alert volume, detection rate [42]	The framework proposed in [42] can be explored by integrating with existing fraud detection systems.
Accuracy, recall, precision, F1 score [43]	Historical credit card transactions to evaluate transactions in training data and other clustering techniques may be considered in the future [43].
Sensitivity, specificity, TP, FP, TN and FN, ROC Plot, F-Measure [44]	The process [44] can help to predict suspicious activities in future by detecting patterns in crimes.
Accuracy, kappa and mean of relative error and absolute error [45]	To overcome the drawback of current research, future work by [45] can be carried out on smart techniques, those that can deal with a large volume of real-world applications on non-stationary situations, such as Gaussian distribution.
Sensitivity, specificity, precision, F-score, AUC, ROC [46]	Future research can focus on imbalance class handling in big data environments [46].
Accuracy, precision [47]	The SOAR extraction methodology [47] can be used in other areas where transparency is more vital in fraud detection.
VDR and TFPR, ROC [48]	Further research can be conducted on long short-term memory and NN [48].
ROC, precision and accuracy [49]	There is a need to evaluate other validation methods in the future [49].

Table 4. *Cont.*

Performance Measure	Future Work
ROC curves, precision, recall [50]	To carry out future research, cost based performances in [50] can be used to train an algorithm for a learning process which can provide better results for business applications.
Fraud triangle, mean, S.E, p [51]	The models offered by [51] can be compared with other supervised and unsupervised techniques as a future study.
F1, precision and recall [52]	To achieve accuracy and efficiency at the same time, a multi objective optimization framework can be implemented [52].
Cost saving [53]	Calculation time for various attributes and response time can be explored by using the framework proposed in [53].
AUCPR, AUCPR [54]	The two directions [54] for future research are, for one, fraud detection, and other is more general and can be evaluated in other fields.
MCC, fraud catching rate, false alarm rate, BCR [55]	The strategy proposed by [55] can be used with real time systems as future research.
FPR, accuracy, TP, FP [56]	Using real-world data, comparison can be performed with different metrics and techniques on the basis of feature selections [56].
Accuracy [57]	The work in [57] can be useful for fraud problems, especially online credit card fraud.
AUC, S.D, execution time [58]	Further research [58] can be performed on why neural networks are always good with small datasets.
TPR, TNR, NPV, alarm rate [59]	More studies can be performed on focusing the associative memory to OPWEM and can analyze the outcome to overcome fraudulent transactions. OPWEM can be modified to overcome fraudulent transactions online [59].
TP, FP, TN, FN, accuracy, space complexity, scalability, time complexity, training time, detection robustness, F-Measure [60]	Challenges of HMM identified by [60] can be used to conduct further research.
Mean SD, F p, accuracy (%) error rate (%), ROC, AUC [61]	Others studies can be conducted with different countries and machine learning techniques can be explored for detection of violators [61].
Overall accuracy, recall, precision, specificity, F 1, FNR, FPR, AUC, model building time [62]	Integration of various machine learning techniques and data can be used from long time frames as future research [62].
AUC, average, positive, negative [63]	An interesting new work [63] can be conducted to explore high-risk firms, even if they are not currently suffering from fraud.
FPR, FP ratio detection rate [64]	By using the knowledge of extraction rules using historical transaction data, the transaction processing can be made more efficient [64].
Accuracy [65]	Other data mining fields can be explored with NN [65].
Sensitivity, specificity, precision, F1 measure [66]	In future [66], comparison of LMSVM in a multi class situation can be compared with other detection techniques.
TP, FP, TN, FN, F-Score, AUC, ROC, mean median SD [67]	Base rate estimations, classifier performance, cost, and usefulness of classifier can be major issues to be addressed in further research [67].

6.5. Analysis Based on Journals and Conferences

Journal and conferences of papers and frequency of publications with respect to the journals are presented in this segment.

The studies that are included for this systematic review of credit card fraud and machine learning are represented by Table 5. These studies are from different journals and conferences. Minimum studies included from one journal or conference is one and maximum is five.

Table 5. Articles published in journal and conferences.

Journal/Conf.	Number of Papers
Managerial Auditing J.	01
Journal of Money Laundering Control	01
Journal of Financial Crime	05
Kybernetes	02
Information & Computer Sec.	01
Inst. Of elec. and electronic Trans. on Know. & dataEng	02
Inst. Of elec. and electronic Trans. on NN &LearningSys.	01
Inst. Of elec. and electronic ACCESS	03
2008 7th Int. Conf. on Machine Learning and Applications	01
Inst. Of elec. and electronic Trans on computational social sys.	01
2016 Mang. and Innovation Tech. Int. Conf. (MITiCON-2016)	01
2017 Inst. Of elec. and electronic Colombian Conference on Communications and Computing (COLCOM)	01
Intr. Jor. Sys. Assur Eng. Mang.	01
Opt. Mem. & NN (Information Optics)	01
Journal of Ambient Intelli. & Humanized Comp.	01
Springer Science+Business Media	01
Front Inform Technol Electron Eng	01
Int. Symp. on Sec in Computing and Comm.	01
Proceedings of 2nd Int. Conf. on Computer and Com Tech.	01
Pattern Recognition Letters	01
Decision Support Systems	02
Knowledge-Based Systems	01
Eng. App. of A.I	01
Exp. Sys. With Applications	02
Procedia Computer Science(International conference on intelligent computing)	01
The Int. Conf on ComputationalSci.	01
First Int. Conf. on Intel. Computing in Data Sci.	01
Computational Intelligence	02
Expert Systems	01
Stat Anl. Data Min.	01
Journal of Forecasting, J. Forecast	01
Intell Sys Acc Fin Mgmt	01
Information Systems Journal	01
WIREs Data Mining and Knowledge Discovery	01
ETRI Journal,	01
Info Systems J.	01

7. Conclusions

Research on fraud detection through machine learning is a very hot topic from the last decade. Definitely, it is observed that a significant amount of research is conducted, but

still there is a need for more future research in fraud detection. This article is an effort to enhance the knowledge of the research by conducting a wide-ranging systematic literature review. Five databases are selected and then articles from these databases are included in this research. An extraction process is applied to the articles and finally selected articles are included in the study. The research gaps in previous research show that several issues need to be focused on more in the future, such as class imbalance, high dimensions, consumption time, accurate fraud detection, and decrease in false alarms.

The systematic literature review is conducted in this article as a review of past research from the five famous databases. The main focus is on the machine learning techniques, performance metrics and datasets used by the researchers. The analyses examined in this article clearly shows that primary datasets are real time and primary datasets are more important to generate better outcomes. The secondary datasets presented by the different online repositories are less effective for the presentation of actual problems.

The databases are summarized and then categorized according to the perspective of cybercrime and machine learning techniques, performance metrics, and datasets that would be helpful for financial firms to understand cybercrime in detail. This article can contribute to cybercrime awareness and security frameworks of banks. Moreover, this paper emphasized the availability of primary datasets in the area of machine learning. Machine learning algorithms can provide effective results with real datasets and researchers can also pay more attention to the actual problem after the analysis of real data. This study is beneficial for future researchers as it highlights the need for real datasets. Mostly researchers focused on neural networks and regression analysis in management analysis methods.

Furthermore, this article is very helpful for future research and mentions the main issues in the area of fraud detection by using machine learning. These issues are handling complex data, imbalance data, real time response, customer behavior, spending patterns, customer awareness, and impact of social, technological and administration. The performance metrics that can be explored in future research include false positive, true positive, response time of algorithms and area under the curve. Our paper is very helpful as it is knowledge based on up to date research which can be very beneficial for researchers, especially those dealing with fraud detection.

Future work in this study can be conducted as it has a lot of limitations. This study focused on a list of keyword-based searches for this literature review. Further in depth studies will be performed with more keywords. Secondly, we just focus on five repositories; other repositories will be explored in future. Thirdly, our time span is from 2010 to 2019; time will be expanded with further research. Furthermore, this research is focused on articles, but in the next work financial reports, non-English articles and PhD theses could be included. There should be proper reporting on cybercrime to generate more real time datasets. The banks should be more aware that if there will be a greater availability of real time data, then there will be more understanding of the issue. The researchers can also conduct in depth analysis on primary data availability in the field of fraud detection.

Author Contributions: Conceptualization, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Data curation, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Formal analysis, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Funding acquisition, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Investigation, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Methodology, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Project administration, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Resources, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Software, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Supervision, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Validation, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Visualization, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Writing—original draft, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L.; Writing—review & editing, R.M.D., R.F., F.J., P.N.M., M.N.M. and G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Instituto Politécnico de Lisboa.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Instituto Politécnico de Lisboa for providing funding for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khan, S.; Zeeshan, A.; Gul, A.; Haider, S.A. Islamic banking system of Pakistan: Comparison between perception and experience. *Acad. Strateg. Manag. J.* **2021**, *20*, 1–6.
2. Hyde, A.M. E-BANKING: Review of literature. *Prestige e-J. Manag. Res.* **2015**, *2*, 2.
3. Kandasamy, K.; Srinivas, S.; Achuthan, K.; Rangan, V.P. IoT cyber risk: A holistic analysis of cyber risk assessment frameworks, risk vectors, and risk ranking process. *Eurasip. J. Inf. Secur.* **2020**, *8*, 8. [CrossRef]
4. Cremer, F.; Sheehan, B.; Fortmann, M.; Kia, A.N.; Mullins, M.; Murphy, F.; Materne, S. Cyber risk and cybersecurity: A systematic review of data availability. *Geneva Pap Risk Insur. Issues Pract.* **2022**, *47*, 698–736. [CrossRef] [PubMed]
5. Khattri, V.; Singh, D.K. Parameters of automated fraud detection techniques during online transactions. *J. Financ. Crime* **2018**, *25*, 702–720. [CrossRef]
6. Krambia-Kapardis, M.; Christodoulou, C.; Agathocleous, M. Neural networks: The panacea in fraud detection. *Manag. Auditing J.* **2010**, *25*, 659–678. [CrossRef]
7. Jie, W.; Poulova, P.; Haider, S.A.; Sham, R.B. Impact of internet usage on consumer impulsive buying behavior of agriculture products: Moderating role of personality traits and emotional intelligence. *Front. Psychol.* **2022**, *13*, 951103. [CrossRef]
8. Prabowo, H.Y. Building our defence against credit card fraud: A strategic view. *J. Money Laund. Control.* **2011**, *14*, 371–386. [CrossRef]
9. Zheng, L.; Liu, G.; Yan, C.; Jiang, C. Transaction Fraud Detection Based on Total Order Relation and Behavior Diversity. *IEEE Trans. Comput. Soc.* **2018**, *5*, 796–806. [CrossRef]
10. Alloghani, M.; Al-Jumeily, D.; Mustafina, J.; Hussain, A.; Aljaaf, A.J. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In *Supervised and Unsupervised Learning for Data Science. Unsupervised and Semi-Supervised Learning*; Berry, M., Mohamed, A., Yap, B., Eds.; Springer: Cham, Switzerland, 2020. [CrossRef]
11. Randhawa, K.; Loo, C.K.; Seera, M.; Lim, C.P.; Nandi, A.K. Credit Card Fraud Detection Using AdaBoost and Majority Voting. *IEEE Access* **2018**, *6*, 14277–14284. [CrossRef]
12. Dal Pozzolo, A.; Boracchi, G.; Caelen, O.; Alippi, C.; Bontempi, G. Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy. *IEEE Trans. Neural Networks Learn. Syst.* **2018**, *29*, 8.
13. European Central Bank, Euro System, Executive Summary. Available online: https://www.ecb.europa.eu/pub/cardfraud/html/ecb.cardfraudreport202110~{}cac4c418e8.en.html (accessed on 11 December 2020).
14. Awoyemi, J.O.; Oluwadare, S.A. Credit card fraud detection using Machine Learning Techniques: A Comparative Analysis. In Proceedings of the 2017 International Conference on Computing Networking and Informatics (ICCNI), Lagos, Nigeria, 29–31 October 2017.
15. Ali, A.; Iqbal, S.; Haider, S.A.; Tehseen, S.; Anwar, B.; Sohail, M.; Rehman, K. Does governance in information technology matter when it comes to organizational performance in Pakistani public sector organizations? Mediating effect of innovation. *SAGE Open.* **2021**, *11*, 21582440211016557. [CrossRef]
16. Vipul, P.; Umesh, K.L. A Survey on Different Data Mining & Machine Learning Methods for Credit Card Fraud Detection. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2018**, *3*, 320–325.
17. Chauhan, T.; Rawat, S.; Malik, S.; Singh, P. Supervised and Unsupervised Machine Learning based Review on Diabetes Care. In Proceedings of the 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 19–20 March 2021; pp. 581–585. [CrossRef]
18. Kilincer, I.F.; Ertam, F.; Abdulkadir, S. Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Comput. Netw.* **2021**, *188*, 107840. [CrossRef]
19. Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J. Survey of intrusion detection systems: Techniques, datasets and challenges. *Cybersecurity* **2019**, *2*, 20. [CrossRef]
20. Alsaedi, A.; Moustafa, N.; Tari, Z.; Mahmood, A.; Anwar, A. TON_IoT telemetry dataset: A new generation dataset of IoT and IIoT for data-driven intrusion detection systems. *IEEE Access* **2020**, *8*, 165130–165150. [CrossRef]
21. Vaccari, I.; Chiola, G.; Aiello, M.; Mongelli, M.; Cambiaso, E. MQTT set, a new dataset for machine learning techniques on MQTT. *Sensors* **2020**, *20*, 17. [CrossRef]
22. Mahfouz, A.; Abuhussein, A.; Venugopal, D.; Shiva, S. Ensemble classifers for network intrusion detection using a novel network attack dataset. *Future Internet* **2020**, *12*, 180. [CrossRef]
23. Iwasokun, G.B.; Akinyede, R.O.; Fadamiro, C.F.; Bello, O.A. Factor analysis of financial crime-related issues. *J. Financ. Crime* **2019**, *26*, 113–130. [CrossRef]
24. Bai, F.; Chen, X. Analysis on the new types and countermeasures of credit card fraud in mainland China. *J. Financ. Crime* **2013**, *20*, 267–271. [CrossRef]
25. Seuring, S.; Gold, S. Conducting content-analysis based literature reviews in supply chain management. *Supply Chain Manag. Int. J.* **2012**, *17*, 544–555. [CrossRef]

26. Whitty, M.T. Predicting susceptibility to cyber-fraud victimhood. *J. Financ. Crime* **2019**, *26*, 277–292. [CrossRef]
27. Omar, N.; Johari, Z.A.; Smith, M. Predicting fraudulent financial reporting using artificial neural network. *J. Financ. Crime* **2017**, *24*, 362–387. [CrossRef]
28. Adewumi, O.A.; Akinyelu, A.A. A hybrid firefly and support vector machine classifier for phishing email detection. *Kybernetes* **2016**, *45*, 977–994. [CrossRef]
29. Lin, W.-C.; Ke, S.-W.; Tsai, C.-F. Top 10 data mining techniques in business applications: A brief survey. *Kybernetes* **2017**, *46*, 1158–1170. [CrossRef]
30. Sun, Y.; Davidson, I. Influential factors of online fraud occurrence in retailing banking sectors from a global prospective: An empirical study of individual customers in the UK and China. *Inf. Comput. Secur.* **2015**, *23*, 3–19. [CrossRef]
31. Phua, C.; Smith-Miles, K.; Lee, V.; Gayler, R. Resilient Identity Crime Detection. *IEEE Trans. Knowl. Data Eng.* **2012**, *24*, 533–546. [CrossRef]
32. Lee, Y.J.; Yeh, Y.R.; Wang, Y.C.F. Anomaly Detection via Online Oversampling Principal Component Analysis. *IEEE Trans. Knowl. Data Eng.* **2013**, *25*, 7. [CrossRef]
33. Gadi, M.F.A.; Wang, X.; do Lago, A.P. Comparison with Parametric Optimizationin Credit Card Fraud Detection. In Proceedings of the Seventh International Conference on Machine Learning and Applications, San Diego, CA, USA, 11–13 December 2008.
34. Huang, D.; Mu, D.; Yang, L.; Cai, X. CoDetect: Financial Fraud Detection With Anomaly Feature Detection. *IEEE Access* **2018**, *6*, 19161–19174. [CrossRef]
35. Qiu, S.; He, H.-Q. Machine Learning- and Evidence Theory-Based Fraud Risk Assessment of China's Box Office. *IEEE Access* **2018**, *6*, 75619–75628. [CrossRef]
36. Charleonnan, A. Credit Card Fraud Detection Using RUS and MRN Algorithms. In Proceedings of the 2016 Management and Innovation Technology International Conference (MITiCON-2016), Bang-San, Thailand, 12–14 October 2016.
37. Melo-Acosta, G.E.; Duitama-Muñoz, F.; Arias-Londoño, J.D. Fraud Detection in Big Data using Supervised and Semi-supervised Learning Techniques. In Proceedings of the IEEE Colombian Conference on Communications and Computing (COLCOM), Cartagena, Colombia, 16–18 August 2017.
38. Adewumi, A.O.; Akinyelu, A.A. A Survey of Machine-Learning and Nature-Inspired Based Credit Card Fraud Detection Techniques. *Int. J. Syst. Assur. Eng. Manag.* **2017**, *8*, 937–953. [CrossRef]
39. Bekirev, A.S.; Klimov, V.V.; Kuzin, M.V.; Shchukin, B.A. Payment Card Fraud Detection Using Neural Network Committee and Clustering. *Opt. Mem. Neural Netw.* **2015**, *24*, 193–200. [CrossRef]
40. Jain, A.K.; Gupta, B.B. A machine learning based approach for phishing detection using hyperlinks information. *J. Ambient Intell. Humaniz. Comput.* **2018**, *10*, 2015–2028. [CrossRef]
41. Somasundaram, A.; Reddy, S. Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance. *Neural Comput. Appl.* **2018**, *31*, 3–14. [CrossRef]
42. Wei, W.; Li, J.; Cao, L.; Ou, Y.; Chen, J. Effective detection of sophisticated online banking fraud on extremely imbalanced data. *World Wide Web* **2013**, *16*, 449–475. [CrossRef]
43. Zhou, H.; Chai, H.F.; Qiu, M.L. Fraud detection within bankcard enrollment on mobile device based payment using machine learning. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 1537–1545. [CrossRef]
44. Dheepa, V.; Dhanapal, R. Hybrid Approach for Improvising Credit Card Fraud Detection Based on Collective Animal Behaviour and SVM. In *International Symposium on Security in Computing and Communication*; Springer: New York, NY, USA, 2013.
45. Ade, P.K. Logistic Regression Learning Model for Handling Concept Drift with Unbalanced Data in Credit Card Fraud Detection System. In Proceedings of the Second International Conference on Computer and Communication Technologies, Hyderabad, India, 24–26 July 2015.
46. Mohammed, R.; Wong, K.-W.; Shiratuddin, M.F.; Wang, X. Scalable Machine Learning Techniques for Highly Imbalanced Credit Card Fraud Detection: A Comparative Study. In *PRICAI 2018: Trends in Artificial Intelligence*; Springer: New York, NY, USA, 2018.
47. Krause, N.F.-T. *Neural Network Rule Extraction to Detect Credit Card Fraud. International Conference on Engineering Applications of Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2011.
48. Gómez, J.A.; Arévalo, J.; Paredes, R.; Nin, J. End-to-end neural network architecture for fraud scoring in card payments. *Pattern Recognit. Lett.* **2018**, *105*, 175–181. [CrossRef]
49. Chandra, S.; Throckmorton, W.J. Financial fraud detection using vocal, linguistic and financial cues. *Decis. Support Syst.* **2015**, *74*, 78–87.
50. Carneiro, N.; Figueira, G.; Costa, M. A data mining-based system for credit-card fraud detection in e-tail. *Decis. Support Syst.* **2017**, *95*, 91–101. [CrossRef]
51. Lin, C.C.; Chiu, A.A.; Huang, S.Y.; Yen, D.C. Detecting the Financial statement fraud: The anaylsis of the differences between the data mining and expert's judgments. *Knowl.-Based Syst.* **2015**, *89*, 459–470. [CrossRef]
52. de Sá, A.G.; Pereira, A.C.; Pappa, G.L. A customized classification algorithm for credit card fraud detection. *Eng. Appl. Artif. Intell.* **2018**, *72*, 21–29. [CrossRef]
53. Bahnsen, A.C.; Aouada, D.; Stojanovic, A.; Ottersten, B. Feature Engineering Strategies for Credit Card Fraud Detection. *Expert Syst. Appl.* **2016**, *51*, 134–142. [CrossRef]
54. Jurgovsky, J.; Granitzer, M.; Ziegler, K.; Calabretto, S.; Portier, P.E.; He-Guelton, L.; Caelen, O. Sequence classification for credit card fraud detection. *Expert Syst. Appl.* **2018**, *100*, 234–245. [CrossRef]

55. Zareapoor, M.; Shamsolmoali, P. Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. *Procedia Comput. Sci.* **2015**, *48*, 679–685. [CrossRef]
56. West, J.; Bhattacharya, M. Some Experimental Issues in Financial Fraud Mining. The International Conference on Computational Science. *Procedia Comput. Sci.* **2016**, *80*, 1734–1744. [CrossRef]
57. Chouiekh, A.; Haj, E.H.I.E. ConvNets for Fraud Detection analysis. The First International Conference On Intelligent Computing in Data Sciences. *Procedia Comput. Sci.* **2018**, *127*, 133–138. [CrossRef]
58. Bandaragoda, T.R.; Ting, K.M.; Albrecht, D.; Liu, F.T.; Zhu, Y.; Wells, J.R. Isolation-basedanomaly detectionusing nearest-neighbor ensembles. *Comput. Intell.* **2017**, *34*, 968–998. [CrossRef]
59. Kültür, Y.; Çağlayan, M.U. Hybrid approaches for detecting credit card fraud. *Expert Syst.* **2016**, *34*, e12191. [CrossRef]
60. Ahmadian Ramaki, A.; Rasoolzadegan, A.; Javan Jafari, A. A systematic review on intrusion detection based on the Hidden Markov Model. *Stat. Anal. Data Min. ASA Data Sci. J.* **2018**, *11*, 111–134. [CrossRef]
61. Song, X.P.; Hu, Z.H.; Du, J.G.; Sheng, Z.H. Application of Machine Learning Methods to Risk Assessment of Financial Statement Fraud: Evidence from China. *J. For.* **2014**, *33*, 611–626. [CrossRef]
62. Vasarhelyi, T.S. Predicting credit card delinquencies: An application of deep neural networks Intelligent system in Accounting. *Financ. Manag.* **2018**, *25*, 174–189.
63. Whiting, D.G.; Hansen, J.V.; McDonald, J.B.; Albrecht, C.; Albrecht, W.S. Machine learning methods for detecting patterns of management fraud. *Comput. Intell.* **2012**, *28*, 505–527. [CrossRef]
64. Wong, N.; Ray, P.; Stephens, G.; Lewis, L. Artificial immune systems for the detection of credit card fraud: An architecture, prototype and preliminary results. *Inf. Syst. J.* **2012**, *22*, 53–76. [CrossRef]
65. Stahl, F.; Jordanov, I. An overview of the use of neural networks for data mining tasks. *Wiley Interdiscip. Rev. Data Mining Knowl. Discov.* **2012**, *2*, 193–208. [CrossRef]
66. Wang, G.P.; Yang, J.X.; Li, R. Imbalanced SVM-Based Anomaly Detection Algorithm for Imbalanced Training Datasets. *ETRI J.* **2017**, *39*, 621–631. [CrossRef]
67. Twitchell, D.P.; Fuller, C.M. Advancing the assessment of automated deception detection system: Incorporating the base rate and cost into system evaluation. *Inf. Syst. J.* **2018**, *29*, 738–761. [CrossRef]