Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 4 Issue 3 (2024)

Enhanced Predictive Analysis of Online Consumer Purchase Psychology using Deep Learning

Dr. Nawab Akram,

Associate Professor, Magadh Professional Institute, Danapur, Patna, Bihar, India. nawabakramlnmi@gmail.com

Dr. RVS Praveen,

Director, Digital Engineering & Assurance, LTIMindtree Limited, M/s. Divija Commercial Properties The Skyview-Building No. 20, Hyderabad, Telangana, India. praveen.rvs@gmail.com

Rohit Abhimalla,

Assistant Professor, Paari School of Business, SRM University AP, Guntur, Andhra Pradesh, India. dr.rohitka@gmail.com

Dr. Sindhu. V,

Assistant Professor, Department of Computer Science, Bangalore, Karnataka, India. sindhu.v@christuniversity.in

Dr. Manasi Vyankatesh Ghamande,

Assistant Professor, DESH, Vishwakarma Institute of Technology, Pune, India. manasi.ghamande@vit.edu

Abstract –Financial fraud, which involves fraudulent practices to acquire financial gains, has recently become a major issue in businesses and organizations. It is inefficient, expensive, and time-consuming to discover fraudulent activities through manual verifications and inspections. The intelligent detection of fraudulent transactions is made possible by artificial intelligence through the evaluation of enormous amounts of financial data. Key components for ensuring operational integrity and limiting financial losses in the financial services business include fraud detection and risk assessment. Due to the increasing complexity of fraud schemes, traditional techniques of detection that depend on static rules and historical data are no longer adequate. In order to better detect fraud and evaluate risk in the financial services sector, this study explores the application of predictive analytics and machine learning (ML). Real-time data and adaptive algorithms are used to evaluate the performance of ML techniques such as supervised learning, unsupervised learning, and ensemble methods in detecting fraudulent actions. The results show a considerable improvement in detection accuracy and risk assessment over older methods. This paper also explores the possible obstacles of deploying these technologies, such as data privacy concerns, interpretability, and the need for ongoing model training.

Keywords**—**Fraud Detection, Machine Learning, Threat Cycle, Risk Management.

I. INTRODUCTION

Company owners and managers require a solid grounding in marketing and consumer psychology, especially when it comes to differentiating between needs and wants. A firm grasp of these tenets is necessary for manufacturers to better adjust to the ever-changing buying habits of consumers. The last say in consumer behavior is with the buyers. Hence, consumer behavior is a significant component that might impact a service or commodity's final purchase or usage decision. In order to boost sales, corporate professionals are increasingly expected to understand what aspects impact clients' purchasing decisions. Businesspeople should pay special attention to the purchasing decision because it will undoubtedly influence the marketing strategy used by the next firm [1]. The core of consumer decision making is an integration process that uses knowledge to evaluate two or more different behaviors and choose one. Ultimately, a consumer's internal state, particularly their beliefs and emotions, can have a significant impact on their purchasing decision. The extent to which a consumer's perception and response to their environment are impacted by psychological factors directly correlates to the strength of their desire to make a purchase decision. At any one time, a person's purchasing intention affects their consumption of a good. In general, consumer purchase intention is the consumer's ability and desire to obtain a given good at a specific location and time. Marketing managers often employ strategies to influence customers' intent to buy in order to promote the acquisition of both new and used products from a certain establishment, business, or even organization. The capacity to utilize sports psychology to influence game-day decisions has become an essential skill for enhancing performance due to the field's rising profile events. The relationship between sports psychology and the inclination of customers to purchase is an interesting field of study necessitate a more thorough understanding. Even while online shopping has grown exponentially, consumers are increasingly worried about a host of new issues. These encompass a wide range of issues, including as the safety of financial transactions, the integrity of personal information, the reliability of online agreements, the clarity of product descriptions, and the capacity to enforce legal claims [2]. To customers who favor more traditional means of purchasing, online buyers are more prone to experience feelings of insecurity. Online shopping has developed into a more sophisticated system than traditional transactions, making consumers more vulnerable to skewed perceptions of the process. With the proliferation of online shopping and internet access, more and more consumers are becoming accustomed to and even preferring to do their regular shopping online. Online shopping has become the norm for most individuals these days. Online shopping habits are the major subject of consumer behavior studies conducted within this paradigm. Before making a purchase, nearly all online shoppers look for relevant reviews [3]. This input from consumers, in the form of reviews, has a substantial influence on the purchasing intention or behavior of potential customers. Therefore, internet reviews can be a great tool for predicting how customers will behave while buying online. One of the factors that impacts the intention or decision to buy is the perceived risk of purchasing online. There is an inherent degree of uncertainty for consumers when it comes to selecting a brand and a payment method. Psychologists often use behavioral tests in a controlled environment to investigate the several factors that impact consumers' decision-making process when making purchases.

II. Literature Survey

Cognitive distancing theory (CLT) explains how psychological distance influences people's thoughts and behaviors. Perceptions of local items or events are seen as more concrete and contextualized, while perceptions of distant objects or events are seen as more high-level, abstract, and stable, according to CLT [4]. Psychological distance is the subjective separation between an actor and an event in the actor's mental space, presuming that several degrees of separation can be unified inside one mental space [5]. Geographical, social, probabilistic, and chronological distance are the four primary categories of psychological distance identified by CLT research [6]. A person's social distance from their social target, the amount of time that elapses between an event and a judgment, the perceived likelihood that an event will occur, and the physical distance between an event and an individual are all factors that contribute to psychological distance [7]. As a result, the interplay between various elements has been the focus of several recent research. The interplay between social distance and time distance in evaluating the risk to customers [8]. Analyzing the interplay between geographical and interpersonal distance, customer responses to recommendation systems. When negative reviews are prevalent, consumers may see them as less helpful. When people talk about their "purchase intention" on social media, they're referring to their desire to make an online purchase [9]. For social commerce to achieve its aim of leveraging social networks to increase commercial benefits, it is essential that users increase the volume of transactions by sharing information or suggestions [10]. Since purchase intentions are a strong proxy for real purchases, many researchers have investigated what motivates customers to make these kinds of purchases. Prior studies have investigated the factors that influence online purchasing intentions, including social commerce structures, viewpoints on information quality, and collective skills [11]. And other research have found that when consumers engage with platforms. Relationship success in social commerce sites depends on platform features including social interactions, educational material, and real-time communications that help with decision-making [12]. In social commerce, consumers still encounter challenges when attempting to communicate their intent to buy through platform engagement, even though there is a wealth of literature on the subject. Furthermore, by investigating what is commonly thought to be an essential component of trade relationships psychological contracts and their impact on customers' purchase intentions, the current study addresses a gap in our understanding of social commerce [13]. A psychological contract is usually in place whenever two or more people do anything together. One of the key motivators for individuals to use e-services is their confidence in the service. Perceived institutional procedure success has a detrimental effect on consumers' trust in online retailers and their likelihood to make repeat purchases [14]. Compared to hedonic values, confidence, and privacy concerns, utilitarian values have a more positive impact on consumers' opinions toward online purchase. There are a few factors that impact the uptake of social commerce [15]. These include ease of use, trust in the service provider, and worries about security and privacy. In order to decrease end user risk, create confidence, and ultimately encourage consumer commitment, the significance of an online vendor's reputation. Consumers' perceptions of trust and pleasure in regard to their purchasing habits and intends to purchase are affected by their level of empowerment [16]. It is hard to build trust based on previous actions or the probability of future contacts because people often trade information online without knowing each other. Loyalty to a brand increases when consumers are pleased, invested, and trusting in it [17]. Users' impressions of an online store's expertise, honesty, and friendliness impact their level of trust in that store. The consumer's intent to buy something online is affected by these impressions. Customer satisfaction is defined as "a necessary response of humans to activities with computers as intermediaries". Customers are more likely to enjoy a game if they are enticed to play it [18]. The motivation to achieve a goal stems from an innate interest or enjoyment in the task at hand. An individual's intrinsic motivation stems from their choice for an activity, according to CET [19]. Consequently, individuals will be driven to participate from an internal source—their enjoyment of the activity itself.

People are more likely to play games repeatedly if they have a positive attitude about human-computer interaction, which is a crucial component of consumer satisfaction [20]. Customers can experience positive attitudes through rewards, immersion in games, competition, and a sense of self-control, among other things. The emotional, cognitive, and physical aspects of consumer enjoyment are encompassed by these components [21]. According to this research, "autonomy" occurs when a player actively participates in a gamified activity and keeps trying to get rewards no matter what obstacles they encounter. The thrill and excitement of competing, as well as the sense of accomplishment, are all components of the competitive spirit. The term "enjoyment" refers to the level of pleasure and amusement a consumer experiences while navigating a gamified online environment. Some studies have shown that satisfying specific psychological needs is essential for keeping one's intrinsic drive, or enjoyment, high.

III. Methodology

Financial institutions have recognized that using isolated security methods on individual delivery channels does not provide adequate protection against illegal account activity. Financial IT platforms are vulnerable to fraud due to security weaknesses that allow for large-scale monetary theft.

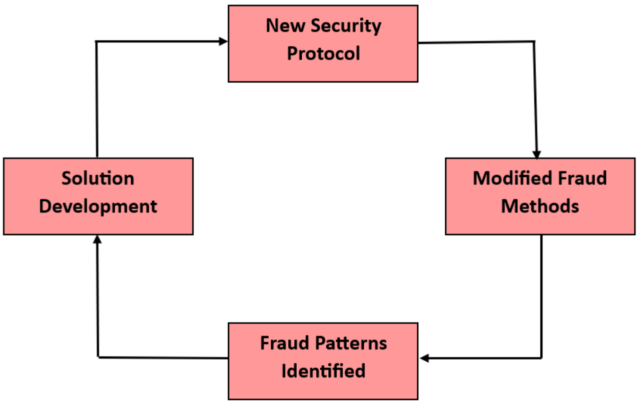

Fig. 1. Threat Cycle of Fraud

Malicious third parties can use weak authentication techniques including signatures, PINs, passwords, and Card Security Codes (CSCs) to conduct fraudulent financial transactions.

A. *Fraud Management*

Because scams can affect a financial institution's reputation, service quality, and bottom line, catching them early is a top priority. In order to combat the vulnerabilities in channelized authentication methods that fraudsters frequently take advantage of, numerous organizations are integrating "Swindle Management" with channelized security requirements. An all-encompassing fraud control architecture with multi-level security across all service networks is the end result of swindle management solutions' active screening of account activity data. To better serve their customers, financial institutions make use of modern technology including plastic credit/debit cards, ATMs, online banking, and mobile banking. The financial services within the business and data logic system levels are managed by allied network level servers, who direct received requests. For the purpose of verifying the legitimacy of users, network security policies and processes depend on the "what the user knows" and "what the user has" criteria. Users are needed to submit necessary security information, including passwords, identification numbers, and personal details, in order to access banking services. As an additional layer of protection, they might verify their identity with a physical security token or smart card. Data mining allows Reactive Fraud Management to execute complex computations on recorded transaction data. Cases of fraud can be identified by comparing them to established patterns of fraud or by noting unusual behavior in relation to the documented history. By necessitating transactional data prior to utilizing data analysis tools, the "store now, query later" strategy can lengthen the time it takes to detect fraud. Keep waiting for the transaction to close and the related fiscal value to rise in order to get a preventative response. To predict the arrival of fresh data, reactive fraud management systems use labeled priming data sets. To keep up with emerging fraud risks posed by unlabeled transaction requests, models must be retrained on a regular basis. Large financial losses and undetected fraud instances happen because of delays caused by adding a large number of tagged swindle cases to the training data set.

B. *Types of Financial Fraud Detection*

Financial fraud can take numerous forms, which are briefly described here.

1) Credit Card Fraud:

The crime of credit card fraud occurs when an unauthorized third party uses a credit card to make a purchase. When a physical card is lost or stolen, transactions can still be done remotely. There are several techniques for obtaining the cardholder's information. Phishing is when a fraudster impersonates a finance official to gain access to a user's information. Swipers or skimmers can be used to read a user's card directly, or entire databases can be obtained by breaching the financial institution's network security or enlisting an accomplice. Getting a replacement or fresh card for the user might be as easy as snooping on their mail. There has been a rise in organized crime related to credit card fraud due to the anonymity and accessibility of online methods. Credit card fraud is typically detected by analyzing a customer's spending patterns and flagging transactions that deviate from the norm.

2) Fiancial Fraud:

Financial statements provide facts about a company's expenses, loans, income, and profits. Management remarks on business performance and potential future difficulties may also be included. For investors and potential borrowers, a company's financial statements are a window into its health and performance. Falsifying a company's financial statements to make it look wealthier is known as financial statement fraud or corporate fraud. There are a number of motivations that might lead to financial statement fraud, including the desire to artificially inflate success, reduce tax liabilities, or appease manipulative managers. Financial statement fraud is challenging to detect due to a lack of understanding, infrequent occurrence, and the ability of educated industry professionals to conceal their deception.

3) Insurance Fraud:

Anyone in the insurance chain is capable of committing fraud at any point. False insurance claims, whether caused by inflated injuries or losses or entirely made up, constitute insurance claims fraud. Falsifying or intentionally provoking accidents that result in excessive repair and injury expenses is one example of automobile insurance fraud. People commit crop insurance fraud when they exaggerate the amount they will lose because of things like falling agricultural prices or natural disasters.
Insurance fraud can involve excessive billing, multiple claims, bribes to brokers, and "upcoding" of items.

4) Mortagage Fraud::

Financial fraud can take several forms, one of which is mortgage fraud, which is forging documents pertaining to real estate or mortgages. The practice of exaggerating a property's value in order to secure a loan is widespread.

5) Money Laundering:

Criminals engage in money laundering when they want to launder the proceeds of their illicit activity into more reputable companies. This masks their illicit activity by making it appear as though the funds are coming from legitimate sources. Because it gives criminals access to financial resources, money laundering is bad.

C. Machine Learning Methods for Financial Fraud Detection

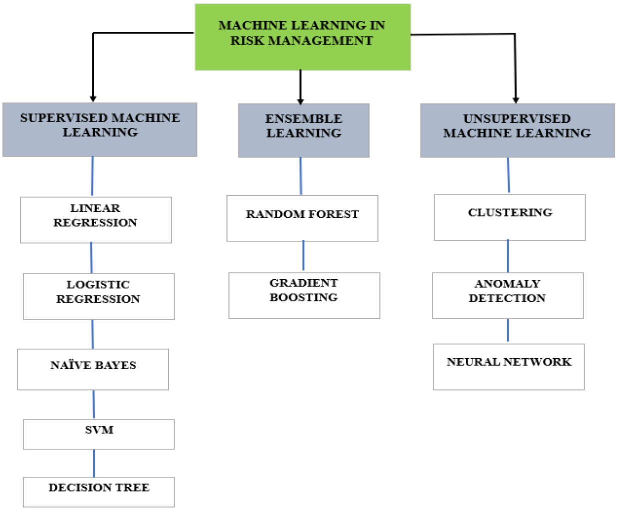

Fig. 2. Machine Learning in Risk Management

1) *Supervised Learning*

As part of its training procedure, the algorithm takes an input-output data set and uses it to build a mathematical model. Textual examples are used to teach algorithms, with input and output predetermined. The input set and the intended outputs are both obtained through the usage of an algorithm in this study. The Algorithm checks its real findings against the correct results to find out what happened. The model is then suitably updated after that. Classification, regression, prediction, and gradient boosting are supervised learning methods that estimate values using patterns. This interpretation is usually used in contexts where past data is used to forecast future occurrences. The two main functions of supervised learning are regression and classification. Decision trees, Naïve Bayes, and close neighbors are all instances of supervised machine learning. Supervised learning has a general foundation in classification. By studying real-world workflow examples, trainees can better understand and quantify how different classes behave when plotting vectors. The goal of inductive machine learning is to build a generalizable classifier by studying rules in previously encountered cases or by the execution of multiple tasks simultaneously.

a) *Linear Regression*

This algorithm for learning is the simplest. Predictive analysis is another possible application of this statistical tool. Continuous and quantitative factors like age, price, and sales are used to develop predictions. The method reveals a straight line between the dependent and independent variables in linear regression. It reveals the relationship between the values of the independent and dependent variables, showing how the latter changes.

b) *Logistic Regression*

Known as supervised learning, this method is utilized by most algorithms. To predict a collection of categorical dependent variables, this technique employs a set of independent variables. It analyzes the performance of a categorical dependent variable. Therefore, a value that can be categorized is necessary. Given a probability value between zero and one, it provides alternatives to yes/no and other binary choices. Except for that one detail, this method is very similar to Linear Regression. You can use Logistic Regression for classification problems and Linear Regression for regression concerns. Using this method, you can categorize observations based on many data sets and determine which factors are most important. Instead of fitting a regression line, we use an S-shaped logistic function that can take on two possible maximum values: 0 and 1. The logistic function's curve represents the probability of an event, such the color of an apple or a ball. Given its versatility, this approach ranks high among learning algorithms. It excels at both continuous and discrete dataset classification and prediction.

c) *Naïve Bayes*

This algorithm is designed for supervised learning. Its basis is Bayes' theorem. This is used for classification-related computing tasks. It can be used to classify texts in a multi-level dataset. Because of its simplicity and effectiveness, the Naive Bayes Classifier has become a go-to tool for training machine learning models. One such approach is a probability-based classifier, which relies on an object's likelihood to produce conclusions. Classifying articles, doing sentiment analysis, and detecting spam are just a few of its many applications.

d) *SVM*

This kind of supervised learning is useful for solving problems related to regression and classification. Its primary utility in machine learning is as a classification dataset. This method aims to find the optimal partition for partitioning an n-dimensional space into cases so that additional data points can be classified more easily in the future. A hyperplane is the term used to describe this dividing plane. Using SVM, the most extreme locations that can be used for hyperplane creation are taken into consideration. Due to the fact that these outliers are referred to as support vectors, the method is known as Support Vector Machine.

e) *Decision Tress*

Even while this supervised learning method works well for classification problems as well as regression ones, regression is where it truly excels. This classifier makes use of a hierarchical framework. Each leaf node represents the final result, the branches reflect the decision-making mechanisms, and the core nodes represent the qualities of the dataset. Two kinds of nodes can be seen in a decision tree: the decision node and the leaf node. A leaf node with no further branches is the end result of a decision node with numerous branches making a choice. The rating is based on features of the dataset that was provided. This graphical depiction shows all possible solutions to an issue given a particular set of parameters. This grows outward from the root node in the same manner that a tree's structure does. A classification and regression tree (CART) is a tool for retrieving variables from a tree. A decision tree divides a tree into subtrees based on yes/no answers.

2) *UnSupervised Learning*

Building a model using only inputs is known as unsupervised learning. Rather of relying on tagged output, this learning method uses unlabeled data. Association rules and K-means are algorithms that fall under this category. Unsupervised learning can be seen in Figure 3. A plethora of algorithms are utilized in unsupervised learning. The algorithms that are most commonly used include: Techniques for acquiring models using latent variables. Neural networks, clustering, and detecting anomalies

a) Clustering

Sorting items into categories according to how similar they are is an integral part of the process. Pattern recognition, machine learning, image analysis, computer graphics, and data retrieval are just a few examples of the many data mining applications that rely on the painstaking investigation of mathematical data. A lot of labor, not an algorithm, is required of the group's study. It is feasible to acquire knowledge on cluster formation comprehensively by use of algorithms that approach the subject from diverse angles. Data space shows perfect clusters with little overlap and separation.

b) Anomaly Detection

Organizing objects causes them to be more tightly packed into one set or cluster than into other groupings. Pattern recognition, machine learning, image analysis, computer graphics, and data retrieval are just a few examples of the many data mining applications that rely on the painstaking investigation of mathematical data. A lot of labor, not an algorithm, is required of the group's study. A cluster and its components can be obtained by various methods.

c) Neural Networks

Computer systems that mimic the structure and function of the brain are known as neural networks. Without task-specific rules, these networks can do tasks based on occurrences. Take photos of cats as an example. They learn to recognize them by comparing them to ones that are labeled as "cat" or "no cat" and then applying that knowledge to other images. Cats, including their fur, tails, cheekbones, and other feline-like facial characteristics, are completely foreign to them. From the models they analyze, they generate visual characteristics. A network of linked nodes called artificial neurons, which stand in for the actual neurons in an animal's brain, is the basis of the ANN. Synapses in the brain's blood supply are just one example of a link that can communicate with other neurons. The artificial neural network (ANN) gathers and evaluates signals before sending them on to connected neurons. In artificial neural network (ANN) applications, the input sum to each neuron is used to determine its output, and the signal to edges is a real value. In general, the symmetry of neurons and edges is congruent with how they learn. A link's signal intensity can be adjusted by changing its weight. Neurons can only transmit signals when their interaction signals exceed a certain threshold. It is common practice to organize neurons in layers. The input is changed in multiple ways via multiple levels. A series of steps is required for the symptoms to progress from the input layer (the first of many layers) to the extraction layer (the last). Originally, the ANN method aimed to mimic human problem-solving abilities; but, as it progressed, its concentration shifted to individual tasks, leading it to stray from its biological roots. Computer vision, machine translation, speech recognition, social networking, medical diagnostics, and even human services like painting are just a few of the many areas that have found usage for ANNs.

3) Ensemble Learning

Ensemble learning is the process of combining multiple machine learning algorithms to achieve better results than each algorithm could achieve on its own. The predictions provided by each learner are integrated using a combination rule to produce a single, more accurate forecast, rather than relying on a single model. There are two main types of ensemble methods: sequential and parallel ensembles. The parallel approaches use a combiner to combine the predictions of many base classifiers that have been trained independently. Bagging and the random forest algorithm, which is an extension of it, are common parallel ensemble techniques. In order to promote variety among the ensemble members, parallel ensemble algorithms make use of the parallel generation of base learners.

a) Random Forest

One of the ensemble models, random forest (RF), creates several prediction models and combines them to produce the final prediction model. From the original data, the RF creates multiple bootstrap samples (training data), and only part of the independent variables are used to train decision trees in each bootstrap sample. After averaging or voting the bootstrap tree predictions, final predictions are then produced when additional data points are presented for independent variables. The RF regression problem uses averaging. Using the remaining data that are not included in the bootstrap samples, the OOB error is a metric that assesses the prediction ability of the bootstrap trees. By employing the ideal number of bootstrap samples where the OOB error is reduced, the predictive performance of the model can be enhanced.

b) Gradient Boosting

A machine learning method called "boosting" can turn a weak classifier into a powerful one. This kind of ensemble meta-algorithm is employed to lessen variance and bias. On the other hand, classifiers that get somewhat better results than random guessing are considered weak learners, and classifiers that achieve much better results called strong learners; it is on the latter that boosting ensemble techniques are founded. Regarding the query of whether a class full of poor students may produce a single exceptional student. A group of subpar learners could generate a single excellent learner. A major influence on statistics and machine learning, which resulted in the creation of various boosting algorithms, such as XGBoost and AdaBoost. Boosting involves iteratively applying a base learning algorithm to changed input data. By adjusting the training set to incorporate misclassified cases from the previous round, boosting techniques teach a weak learner with the input data, compute predictions, and then train a new weak learner. The iterative learning process is carried out until a predetermined quantity of baseline learners is collected and combined. The goal of boosting is to reduce bias, not variation. Base learners with low variance and high bias, such decision stumps (a decision tree with one internal node), benefit from this. Because misclassified samples are given more weight, the base learner tends to focus on them. Extra weight is given to samples that the algorithm determines to be prejudiced against in order to rectify the bias in the underlying classifier. Overfitting, which can occur when boosting algorithms place too much emphasis on noisy samples, makes them unsuitable for learning from noisy data. Despite this, boosting-based ensemble approaches are highly effective in applied machine learning.

D. *Structure of Financial Fraud Detection*

A large number of training and test datasets improves the model's accuracy. Figure 4 shows the channels and patterns used by fraudsters. From data sets, extract these patterns. Instructions for the detection system to tell legitimate purchases apart from fraudulent ones. In real-time, trained models can reject or hold transactions for further investigation in order to detect and prevent fraud.All businesses can benefit from a reduction in global fraud if they work together to share their fraud experiences.

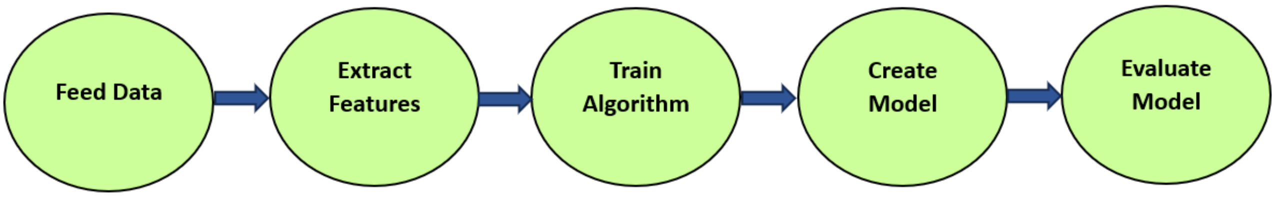

Fig. 3. Structure of Fraud Detection

Because they can be trained using test data to produce efficient and accurate results, Machine Learning Algorithms are good at detecting fraud. Monitoring, learning, detection, prevention, and continuous improvement are all essential components of a complete life cycle strategy for making decisions in real-time. Machine learning algorithms, when used properly, can drastically cut down on fraudulent transactions. Companies should not keep quiet about their fraudulent past if they want to remain ahead of scammers. Although there are systems in place to identify and notify companies of fraudulent transactions, they are siloed and do not facilitate learning across organizations to lessen the impact of fraud. A real-time platform that can automatically learn from previous incidents and alert global organizations is what the industry is aiming for. This is We provide a global model that makes use of AI and ML techniques. Collaborative fraud prevention initiatives cannot be successful without a central fraud management platform. The plan is to build an OS that will let businesses all across the globe communicate fraud trends, find fraudulent transactions before they happen, and secure their apps even more. Businesses should implement digital handshakes to improve communication. All parties concerned can agree on a uniform format for these entities to exchange the database. Since fraudulent transactions are now publicly available, the remaining organizations can take proactive security measures to prevent huge losses. A number of sectors can benefit from this centralized structure, including banking, telecommunications, the stock market, the internet, and social engineering. A dynamic, intelligence-driven approach to risk management is necessary for industries to prevent, detect, react to, and recover from cyber assaults.

IV. Results and Discussion

The performance of the classifiers was evaluated in this work using a variety of assessment metrics. The results of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) informed most of the measurements. Whereas "FP" denotes the anticipated number of valid transactions, "TP" denotes the anticipated number of fraudulent transactions. Amount of fraudulent transactions predicted as legitimate (TN) and number of legitimate transactions predicted as fraudulent (FN) are two different things. Metrics like accuracy, precision, recall, TPR, and FPR were used to evaluate classifier performance in the study. All of the assessment metrics are defined and presented in Table 1.

TABLE I. PERFORMANCE PREDICTION(%)

Metric	Formula and Description
True Positive Rate	$TRP = TP/(TP + FN)$
False Positive Rate	$FRP = FP/(FP + TN)$
Precision	$Precision = TP/(TP + FP)$
Recall	$Recall = TP/(TP + FN)$
F-Measure	$F - Measure = 2TP/(2TP + FP + FN)$
Accuracy	$Acc = (TP + TN)/(TP + TN + FP + FN)$

Using the IEEECIS dataset, Figure 5 displays the accuracy values of different methodologies applied to financial fraud detection. Greater values for accuracy indicate greater performance, since it measures how well each approach detects cases of fraud.

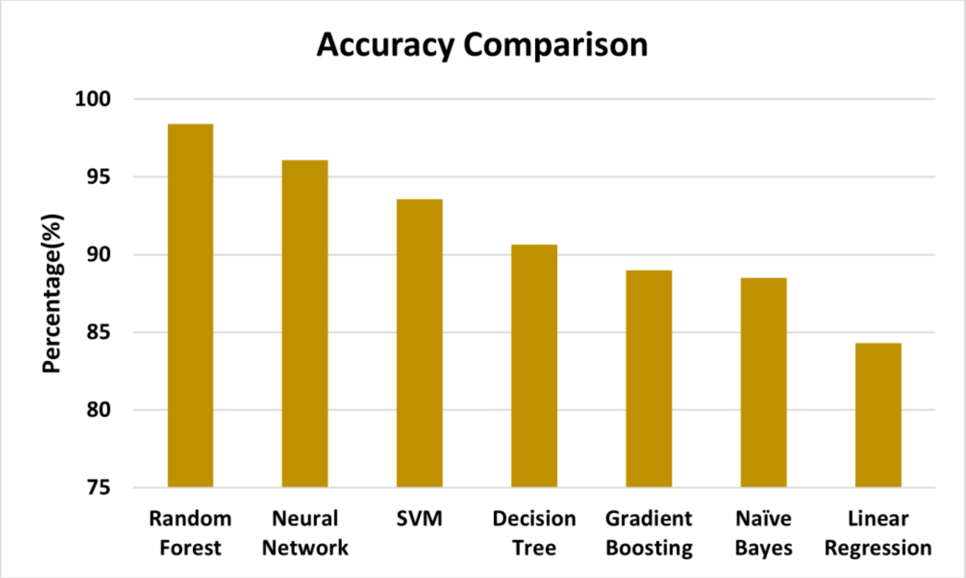

Fig. 4. Performance of Existing Methods

In conclusion, the Random Forest algorithm is the most effective strategy for detecting credit fraud in the specified dataset, with an accuracy of 0.98400 and an F1 score of 0.99194. The Random Forest technique was followed by the Neural Network, which achieved an accuracy of 0.97090 and F1 score of 0.98522. Additionally, we compared this finding to other references shown in Figure 5. The study found that three supervised machine learning techniques outperformed other algorithms: Random Forests, Neural Networks and SVM. Our findings and references support the performance evaluation of Random Forests and AdaBoost algorithms.

TABLE II. ACCURACY PREDICTION(%)

Metric	Accuracy
Random Forest	98.40
Neural Network	96.09
SVM	93.58
Decision Tree	90.63
Gradient boosting	89.00
Naïve Bayes	88.49
Linear Regression	84.30

Table II provides a comparison of the accuracy performance of various machine learning models. The most dependable models for this specific task are Random Forest and Neural Networks, which significantly surpass the others; in contrast, less effective models include Linear Regression and Naïve Bayes.

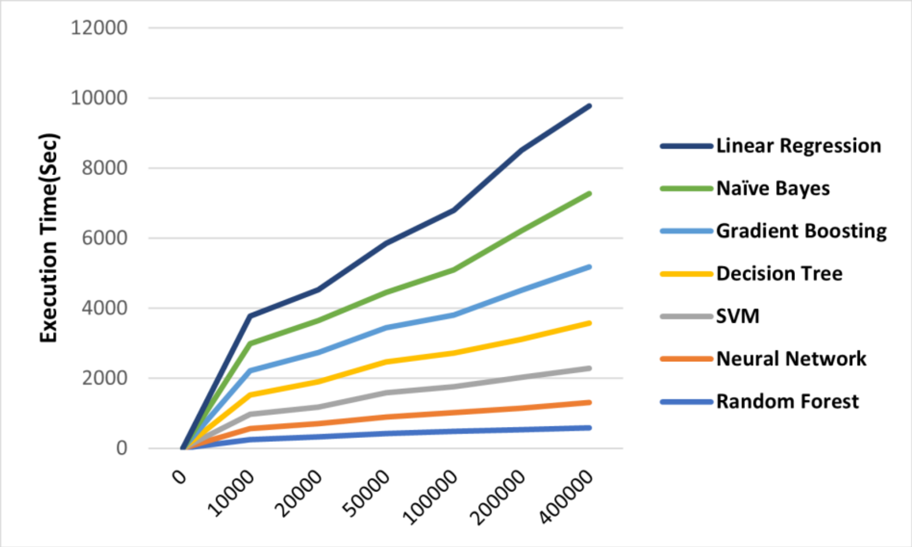

Fig. 5. Comparison of Existing and Proposed models' Computational Complexity

Furthermore, we simulated all datasets to assess the computational complexity of our model; Figure 6 displays the average execution times that resulted from this. The recommended model, Random Forest, has the quickest execution time, proving that it can efficiently handle high- and low-dimensional data and transactions. The model's performance has been fine-tuned, and it can now identify fraudulent users and transactions considerably faster, increasing its use in real-time fraud detection scenarios.

V. Conclusion and Future Directions

One of the most important parts of running a contemporary bank is detecting fraud. Every method showed some promise in detecting different forms of financial fraud, even though they varied in how well they did it. Computational systems, such as support vector machines and neural networks, are effective against fraudsters' evolving strategies because they can adapt to new techniques. There are still a lot of unanswered questions about intelligent fraud detection. To completely comprehend certain forms of fraud and data mining techniques, additional research is required. Parameter tuning allows for computational fraud detection cost-benefit analysis and can enhance the performance of current systems. A more precise foundation for intelligent detection systems could be the result of additional research into the many forms of financial fraud.

REFERENCES

[1] J. H. Lee, "Effect of Sports Psychology on Enhancing Consumer Purchase Intention for Retailers of Sports Shops: Literature Content Analysis," *J. Distrib. Sci.*, vol. 19, no. 4, pp. 5–13, 2021, doi: 10.15722/jds.19.4.202104.5.

[2] R. Wikansari, A. Ausat, R. Hidayat, S. Mustoip, and A. Sari, "Business Psychology Analysis of Consumer Purchasing Factors: A Literature Review," *Eur. Alliance Innov.*, 2023, doi: 10.4108/eai.17-12-2022.2333186.

[3] S. Kamalul Ariffin, T. Mohan, and Y. N. Goh, "Influence of consumers' perceived risk on consumers' online purchase intention," *J. Res. Interact. Mark.*, vol. 12, no. 3, pp. 309–327, 2018, doi: 10.1108/JRIM-11-2017-0100.

[4] D. C. Bojanic and R. B. Warnick, "The role of purchase decision involvement in a special event," *J. Travel Res.*, vol. 51, no. 3, pp. 357–366, 2012, doi: 10.1177/0047287511418364.

[5] T. Bornemann and C. Homburg, "Psychological distance and the dual role of price," *J. Consum. Res.*, vol. 38, no. 3, pp. 490–504, 2011, doi: 10.1086/659874.

[6] L. A. Cai, R. Feng, and D. Breiter, "Tourist purchase decision involvement and information preferences," *J. Vacat. Mark.*, vol. 10, no. 2, pp. 138–148, 2004, doi: 10.1177/135676670401000204.

[7] S. Chung and J. Park, "Exploring consumer evaluations in social media: The role of psychological distance between company and consumer," *Comput. Human Behav.*, vol. 76, pp. 312–320, 2017, doi: 10.1016/j.chb.2017.07.042.

[8] D. A. Drossos, F. Kokkinaki, G. M. Giaglis, and K. G. Fouskas, "The effects of product involvement and impulse buying on purchase intentions in mobile text advertising," *Electron. Commer. Res. Appl.*, vol. 13, no. 6, pp. 423–430, 2014, doi: 10.1016/j.elerap.2014.08.003.

[9] A. H. Busalim, A. R. Che Hussin, and N. A. Iahad, "Factors Influencing Customer Engagement in Social Commerce Websites: A Systematic Literature Review," *J. Theor. Appl. Electron. Commer. Res.*, vol. 14, no. 2, pp. 0–0, 2019, doi: 10.4067/s0718-18762019000200102.

[10] T. S. Chang and W. H. Hsiao, "Factors influencing intentions to use social recommender systems: A social

exchange perspective," *Cyberpsychology, Behav. Soc. Netw.*, vol. 16, no. 5, pp. 357–363, 2013, doi: 10.1089/cyber.2012.0278.

[11] J. Fogel and S. Zachariah, "Intentions to use the yelp review website and purchase behavior after reading reviews," *J. Theor. Appl. Electron. Commer. Res.*, vol. 12, no. 1, 2017, doi: 10.4067/S0718-18762017000100005.

[12] H. N, "Social commerce constructs and consumer's intention to buy," *Int. J. Inf. Manage.*, vol. 35, no. 35 (2), pp. 183–191, 2015.

[13] J. A. Hill, S. Eckerd, D. Wilson, and B. Greer, "The effect of unethical behavior on trust in a buyer-supplier relationship: The mediating role of psychological contract violation," *J. Oper. Manag.*, vol. 27, no. 4, pp. 281–293, 2009, doi: 10.1016/j.jom.2008.10.002.

[14] S. Bebber, G. S. Milan, D. De Toni, L. Eberle, and L. A. Slongo, "Antecedents of Purchase Intention in the Online Context," *J. Relatsh. Mark.*, vol. 16, no. 1, pp. 82–98, 2017, doi: 10.1080/15332667.2016.1242396.

[15] C. Bianchi, L. Andrews, M. Wiese, and S. Fazal-E-Hasan, "Consumer intentions to engage in s-commerce: a cross-national study," *J. Mark. Manag.*, vol. 33, no. 5–6, pp. 464–494, 2017, doi: 10.1080/0267257X.2017.1319406.

[16] X. Chen, Q. Huang, and R. M. Davison, "Economic and social satisfaction of buyers on consumer-to-consumer platforms: The role of relational capital," *Int. J. Electron. Commer.*, vol. 21, no. 2, pp. 219–248, 2017, doi: 10.1080/10864415.2016.1234285.

[17] F. Cui, D. Lin, and H. Qu, "The impact of perceived security and consumer innovativeness on e-loyalty in online travel shopping," *J. Travel Tour. Mark.*, vol. 35, no. 6, pp. 819–834, 2018, doi: 10.1080/10548408.2017.1422452.

[18] Y. Chen, X. Yan, and W. Fan, "Examining the effects of decomposed perceived risk on consumer online shopping behavior: A field study in China," *Eng. Econ.*, vol. 26, no. 3, pp. 315–326, 2015, doi: 10.5755/j01.ee.26.3.8420.

[19] T. Huang, Z. Bao, and Y. Li, "Why do players purchase in mobile social network games? An examination of customer engagement and of uses and gratifications theory," *Program*, vol. 51, no. 3, pp. 259–277, 2017, doi: 10.1108/PROG-12-2016-0078.

[20] A. J. Kim and K. K. P. Johnson, "Power of consumers using social media: Examining the influences of brand-related user-generated content on Facebook," *Comput. Human Behav.*, vol. 58, pp. 98–108, 2016, doi: 10.1016/j.chb.2015.12.047.

[21] H. Kim, K. S. Suh, and U. K. Lee, "Effects of collaborative online shopping on shopping experience through social and relational perspectives," *Inf. Manag.*, vol. 50, no. 4, pp. 169–180, 2013, doi: 10.1016/j.im.2013.02.003.