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

# Real-Time Fraud Detection in E-Commerce Using Machine Learning Models

**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

**Ismatha Begum,**
Assistant Professor, Department of IT, Guru Nanak Institute of Technology, Hyderabad, Telangana, India. ismatha23@gmail.com

**Dr Satish Chandra Ojha,**
Assistant Professor, Department of Marketing, Indian Institute of Management Bodh Gaya, Bihar, India, satish.o@iimbg.ac.in

**Shambhu Sharan Srivastava,**
Associate Professor, Department of Computer Science, School of Management Sciences, Varanasi, India. shambhuss@yahoo.com

**Dr. P. Chellammal**
Professor, Department of Computer Science and Engineering, J J College of Engineering and Technology, Trichy, Tamilnadu, India. chellammalkarthikeyan@gmail.com

**Subharun Pal,**
PG Scholar, Department of CEP, Indian Institute of Technology Patna, Bihta, Patna, Bihar, India. subharunpal@gmail.com

*Abstract* –The number of transactions is increasing since more and more individuals are buying things online. Our research also shows that online transactional misrepresentation is on the increase. There will soon be a meteoric rise in the application of machine learning for the purpose of detecting and preventing online fraud. Inventions always come with their fair share of troubles, and the web-based system during COVID-19 was no exception; there were just too many people utilizing it and making too many transactions online. An e-commerce business can expand in several ways. Online stores rely on fraud detection and prevention systems to be operational. Even while ML plays a significant role in these anti-fraud activities, the organizational context in which these ML models operate must be considered. This study takes an organization-centric approach to the problem of fraud detection by constructing an operational model of anti-fraud departments in e-commerce enterprises.

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

## I. INTRODUCTION

Businesses and consumers alike are impacted by the serious issues of spam and fraud in today's digital world. It is a complex and difficult process to identify the fake reviews. Reviewers who engage in fraudulent practices often receive compensation for their work. Because of this, it is quite difficult for the average consumer to discern between real and fake reviews simply by reading them. Another tactic used by some businesses to boost sales is to have satisfied customers post favorable reviews of their products. The growing level of competition in the business has made it necessary for every corporation to maintain its esteem and height on the lookout. As a result, both businesses and customers should make segmenting automated reviews a top priority [1]. In this paper, we will discuss the importance of a machine learning-based system that can identify spam, fake reviews, and non-reviews. The revolutionary change in how people express and share their opinions is entirely the fault of the social web. What you're looking at here is client-generated content, a term for user-submitted content as opposed to owner-provided content. This proposed e-commerce web app takes user reviews as input, stores them in a database, and then analyzes them to detect spammers. The identification of outliers is a crucial problem with multiple applications. Find the data points that explain the system's strange behavior; that's what outlier identification

is all about. Even though they only make up a tiny fraction of the population, accurately identifying and understanding such data points is critical to the system's well-being. Credit card fraud detection is a common example of a problem that is often presented as a difficulty to identify outliers. Credit card fraud is common in online marketplaces, thus it's important to have reliable systems to detect it[2]. Credit card fraud is already a difficult problem to solve, and the fact that fraudulent tendencies are always changing just makes things worse. The reason behind this is that hackers are continuously inventing new ways to trick their victims and the systems they use. In a time of quickly changing trends and decreasing amounts of labelled data points, it is an immense task to maintain a secure marketplace. To better serve their customers, numerous online marketplaces have implemented data-driven ranking and recommendation algorithms. Despite these algorithms' enhanced capacity to predict users' preferences and boost platform revenues, the data-driven training approach makes them vulnerable to poisoning. If a seller has access to particular user accounts and often clicks on both popular and seller-owned things, the collaborative filtering-based recommendation algorithms could be tricked into thinking there's a strong connection between the two items [3]. This means it has the potential to start promoting the seller's target item to consumers who engage with the popular item, increasing its exposure compared to before. Fortunately, the system's ranking or recommendation algorithms can only be tricked by dishonest sellers that use specific user identities and commit a certain amount of fraud on real e-commerce platforms. Bad sellers' aggregate conduct is usually more noticeable than normal users' because they persistently advocate many target items at once and can only handle so many accounts at once. We are able to detect fraudulent activity as a result of this.

## II. Literature Survey

Here, we provide an overview of the relevant literature on group-based fraud detection. Classification algorithms, cohesive sub graph mining approaches, and fraud detection methods are the focus of our introduction. One way to represent the group-based fraud detection is as a classification problem involving edges or vertex [4]. Several studies utilizes matrix decomposition and balancing theory to forecast bipartite graph edge signs but suffers from handling ungrounded labelled nodes. Utilizes knowledge graphs prediction of edge signs is a problem that recommender systems can resolve [5]. The use of graph neural networks is another approach and edge sign prediction on unipartite using graph embedding not able to make use of community data, although bipartite graphs are an option [6]. In use vertex classification algorithms at the moment pay close attention to evaluating vertex attributes and exchanging neighborhood data to find a solution, yet they are unable to detect fraudulent activity and may not do a good job of identifying sophisticated scams [7]. So far as we are aware, RICD is the breakthrough strategy for the "Ride Item's Coattails" assault. RICD is a uses near-biclique to classify the con artists and eventually uncover the scams. On the other hand, RICD pays little attention to attribute details and necessitates hand-editing. The STARS attack RTV is the most advanced technique for detection [8]. Utilizing all available rating data, our approach successfully identifies con artists. But this method's supervised version, named RTV-SUP, on the other hand, makes poor use of label information while using employs attributes derived from unsupervised learning and an easy way to learn from data (like logistic regression and help identify dishonest individuals [9]. A method called FraudBuster to detect financial scams that steal little sums of money over time. Any change to the user's spending profile that deviates from the taught model, as determined by their model of the user's spending behavior over time, is considered fraudulent by their technique [10]. Instead of looking at the transaction sequences that have been the subject of earlier study, we examine the users past actions sequentially when determining the legitimacy of a transaction [11]. This allows us to confirm if a transaction is legitimate. We were able to reduce the amount of feature engineering required by recasting the fraud detection process as a sequence classification problem using this paradigm [12]. This is carried out so as to demonstrate validity. To make their classifier's answer more clear, they utilize attention-based recurrent neural networks (RNNs) to increase its performance and attention ratings [13]. In order to encapsulate the metadata of each transaction, several variables are used. Some examples of features that come under this category are the day of the week, the hour of the day, the quantity, and the device identifier. The rapid development of AI and the broad availability of affordable cloud computing have both contributed to the increasing use of data mining and machine learning techniques for fraud detection in many different industries [14]. However, researchers have not yet examined the efficacy of these strategies on online marketplaces and social media platforms such as eBay [15]. Our team has seen those prior reviews, have a tendency to generalize about every approach and domain. Such comprehensive coverage does not allow for a thorough understanding of machine learning techniques and their applications in the e-commerce domain [16]. Credit card fraud and other forms of financial fraud are the only foci of the majority of literature reviews on fraud. Furthermore, when it comes to systematic literature reviews, a lot of these studies fall short when it comes to providing evidence for replication [17]. In order to study

data mining and machine learning's role in e-commerce fraud detection on online marketplaces, we suggest doing a systematic literature review following the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) methodology [18]. These are the gaps in the current research that we are aware of. This is a crucial topic to concentrate on because the frequency and expense of fraud are on the increase [19]. Staying informed on the latest research and trends in the field is essential for identifying untapped research opportunities and educating professionals in the field about the most effective machine learning and data mining techniques for detecting fraud. The use of machine learning (ML) is critical for the detection, prevention, and mitigation of e-commerce fraud [20]. Popular examples include Google, LinkedIn, and eBay. The ML models employed to identify online store fraud are actually not operating in a vacuum, but rather as a component of a larger anti-fraud team staffed by analysts and detectives who investigate individual instances and identify trends. To achieve this goal, anti-fraud departments must include fraud detection models into their daily operations [21]. There is a lot of information out there about fraud detection in general, but we couldn't locate anything that details the procedures that anti-fraud teams use every day. Consequently, academic studies on fraud detection systems and their practical applications in businesses are not aligned. As a result, finding out if new types of fraud exist becomes much more complicated detection methods into the routine work of fraud departments, or if they address issues that are genuinely pertinent to the actual world.

## III. METHODOLOGY

The prevalence of online fraud has increased in recent years, leading numerous companies to focus on detecting and preventing it. Billions of dollars in damages still occur every year as a result of fraudulent transactions, despite the fact that fraud detection and prevention technology have come a long way. There has to be better fraud detection and prevention systems in place since next year there will be a dramatic increase in online transactions. Figure 1 depicts the multi-step procedure for creating a data-driven categorization model.

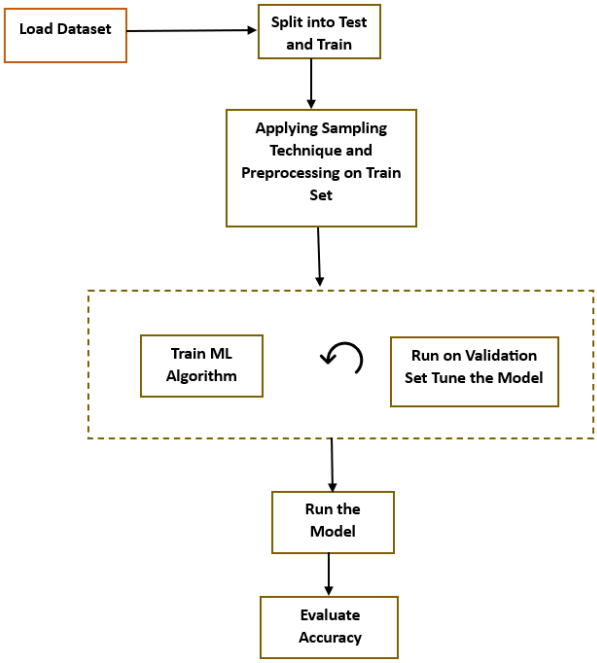


Fig. 1. The Flow Process of Developing the Model

### A. *Types of Fraud*

Telephone fraud, computer intrusion, credit card fraud, theft/counterfeit fraud, bankruptcy fraud, application fraud, and behavioral fraud are some of the many forms of e-commerce fraud.

*1) Credit Fraud:*

There are two main categories of credit card fraud: offline and online.

*a) Offline Fraud:*

It is considered offline fraud to use a stolen or counterfeit physical card in any situation.

*b)Online Fraud:*

Internet, shopping, phone, or web-based fraud occurs as a result of this.

*2) Computer Intrusion:*

Any individual, whether a hacker or an insider with knowledge of the system's architecture, who gains unauthorized access to a system or environment is considered an intruder. There may have been an effort to access or intentionally alter data without authorization[22].

*3) Telephone Fraud:*

The term "telephone fraud" refers to the intentional misuse of mobile phone and fixed line services with the goal to cause harm to another party.

*4) Theft Fraud/Counterfeit Fraud:*

Using someone else's cards without their permission is theft. After the owner contacts the bank with feedback, the bank will then take steps to verify the thief. The same holds true for the use of credit cards remotely, which might result in fraudulent charges. Using your codes on other websites and a copied card number, you can avoid using real cards and signatures.

*5) Bankruptcy Fraud:*

Personal bankruptcy fraud is notoriously hard to predict, and it's possible that a consumer will fall prey to it when a bank issues them an order. Furthermore, it is becoming increasingly difficult to receive loans that are not desired. A potential strategy for avoiding this sort of fraud is to do a pre-check with the credit bureau. In this way, the banking records of its customers were exposed.

*6) Application Fraud:*

An example of application fraud would be submitting a credit card application with inaccurate information. It is possible to tell applications that belong to the same person by looking for similarities in their details (called duplicates) or by looking for differences between users' details (called identity fraudsters)[23].

*7) Behavioral Fraud:*

Placing orders using cardholder present credentials when valid card details have been fraudulently obtained is behavioral fraud.

## B. *Methods Used For Credit Card Fraud Detection:*

*1) Bayesian Networks:*

The interdependencies among the variables in a probabilistic model are illustrated in Figure 2 by a directed acyclic graph (BN). For instance, in the case where event A directly impacts event D, then event B immediately follows suit, and the chain reaction continues thereafter. The arcs on the graph show the relationships between the random variables, whereas each node represents a variable. As an added bonus, e is independent.

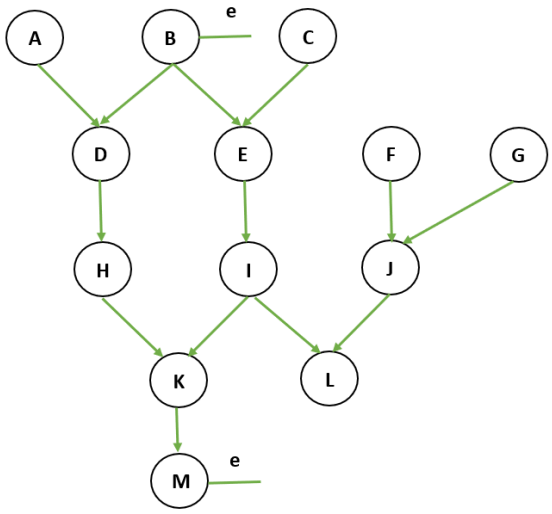


Fig. 2. Bayesian Networks

Equation 1 calculates the conditional probability of an event $Q_p$ given an event $T$, where $I(Q_p|T)$is the probability of $Q$ when $T$ occurs; this form of the Bayesian theorem is the mathematical basis for BN.

$$I(Q_p|T) = \frac{I(T|Q_p)I(Q_p)}{I(T)} \tag{1}$$

In the fraud detection challenge, the BN is unknown, hence learning it from the data is important to construct the BN graph. Equation 1 and the BN graph allow us to identify the dependent variables (conditional probability) that are required for fraud to take place. Prior to calculating the conditional probability, we can use Equation 2 to ascertain the probability of fraud.

$$I(k_p, \dots, k_d) = \prod_{p=0}^{d} I\left(k_p \middle| parents(K_p)\right) \tag{2}$$

where Figure 1 shows a graph that determines $parents(K_p)$.

*2) Neural Networks:*

A real-life neuron is modeled after in a Neural Network (NN) by a network of interconnected basic processing units (called "nodes"). The processing capabilities of the network are represented by the weights, which are the inter-unit connection strengths learnt from a series of training patterns[24]. The standard formula for a neuron's output, Equation 3, is a linear combination of the following variables $k_z$:

$$\begin{aligned} net &= g_1 * k_1 + g_2 * k_2 + \cdots + g_N * k_N \\ &= \sum_{z=1}^{N} (g_z k_z)\, x = g^T * k \end{aligned} \tag{3}$$

in which $g_z$ is a weight linked to the input $k_z$. This weight serves as a measure of the extent to which an input impacts the output value. An activation function, which might be linear, ramp, sigmoid, step, hyperbolic tangent, or gaussian, is used with the computed result (net). With its capacity to categorize non-linearly separable regions, the MultiLayer Perceptron (MLP) NN model was chosen for our fraud detection strategy. The Levenberg-Marquardt method was used for the training since it is both fast and effective. To run a series of tests to find the optimal configuration for NNs, which is a two-layer network with a hidden layer of 10 neurons and an output layer of 1 neuron.

*3) Random Forest:*

Based on the idea of using trees for product classification, Breiman presented the Random Forest (RF) algorithm. According to Breiman, an algorithm is defined as follows: A RF is a classifier that uses a collection of tree-structured classifiers $y(k, \theta_x)x = 1$ where the $\theta_x$ are independent, identically distributed random vectors and each tree votes for the most popular class at input $k$. With a high probability of success $I(y(k) = H)$, the classifier is probably effective. The variables of the problem are shown by the vector $K$, while the response is represented by the vector $H$. Everyone can see the dataset.

$$\left(\left(k_{1,1}, \dots, k_{1,d}\right), \left(k_{2,1}, \dots, k_{2,d}\right), \dots, \left(k_{x,1}, \dots, k_{x,d}\right)\right) \quad (4)$$

where m is the number of features and $H$ is the number of trees.

*4) Isolation Forest Algorithm:*

When it comes to e-commerce fraud detection, the data points are minimal and distinct, which is why the Isolation Forest supervised technique is applicable. Abnormalities can be "isolated" due to these characteristics. For each provided dataset, the Isolation Forest (or iForest) constructs an ensemble of iTrees; instances with shorter average path lengths on the iTrees are included in the fraud detection in E-Commerce output. Here, the only two parameters are the sub-sampling size and the quantity of trees to construct. The iForest's detection performance converges rapidly with a small number of trees and requires a small amount of precision[3].

The maximum iTree height increases on the order of $d$, but the average height increases on the order of log $d$, therefore it is not easy to get this kind of score from the path length $y(k)$. The typical iTree path length is given by Equation (5).

$$v(d) = 2Y(d-1) - \left(\frac{2(d-1)}{d}\right) \quad (5)$$

To normalize the path length $y(k)$, we utilize the average path length $v(d)$and the harmonic number $Y(p)$, which may be approximated by adding 0.577 to ln $(p)$which is Euler's constant. As seen in Eq. (6), the instance $k$ 's anomaly score is $l$.

$$l(k, d) = 2^{-\frac{U(y(k))}{v(d)}} \quad (6)$$

the average of $y(k)$ from the set of isolation trees is denoted as $U\big(y(k)\big)$.

$$U\big(y(k)\big) \to v(d), l \to 0.5 \quad (7)$$
$$U\big(y(k)\big) \to 0, l \to 1 \quad (8)$$
$$U\big(y(k)\big) \to d - 1 \to l \to 0.5 \quad (9)$$

A connection between $l$ and $U\big(y(k)\big)$. It is possible to do the following evaluations using the score $l$: If an instance's return s is very close to 1, then it is definitely an anomaly. If the $l$ -value is substantially smaller than 0.5, it is reasonable to presume that the instances are typical. Once all instances produce s ~ 0.5, it becomes evident that the total sample is normal. Here are the steps that the Isolation Forest algorithm takes: In order to separate observations, the isolation forest uses a random feature selection process followed by a recursive selection of random split values. The number of branches in a tree representing recursive splitting is directly proportional to the length of the path from the root node to the branch's terminal node. The normalcy metric is this path length averaged throughout a forest of trees that are grown in the same way. If the average length is short, then isolating the sample is easier and the sample is more likely to be abnormal.

## IV. Results and Discussion

This proposed used a number of evaluation indicators to determine how well the classifiers performed. Most of the measures were based on the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) results. Whereas "FP" stands for the expected quantity of legitimate transactions, "TP" stands for the expected quantity of fraudulent ones.

One common metric for evaluating binary classifier performance is the ROC curve. Classifier sensitivity to false positive rate is depicted by the receiver operating characteristic (ROC) curve. The false positive rate is the result of subtracting the specificity of the classification model from one. You can see the proposed model's receiver operating characteristic (ROC) curve in Figure 3. The logistic regression classification model yields a score of 0.95 and an area under the curve of 0.95.

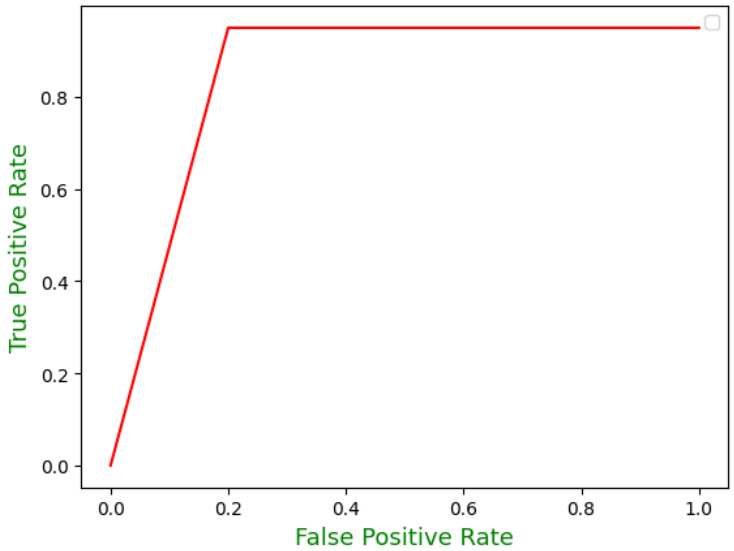


Fig. 3. ROC Curve of the Proposed Model

Value of transactions as seen by the cumulative distribution function (CDF) in Figure 4. Of all valid transactions, 64% have values below $25 USD, whereas 40% are associated with chargebacks. Valid transactions typically have smaller values than chargeback ones, as can be seen from the data.

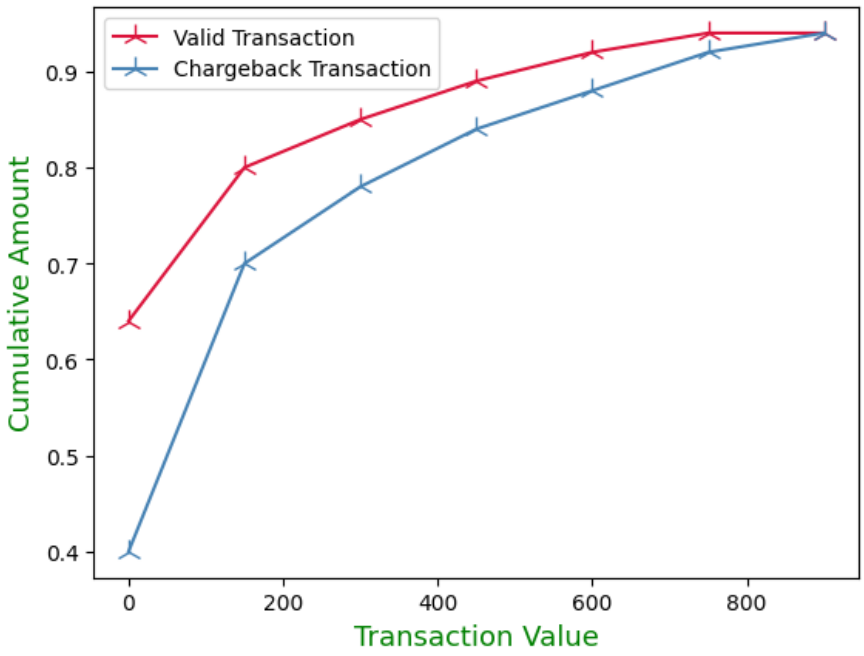


Fig. 4. Cumulative Distribution Function (CDF) of Transaction Value

In Fig. 5, we can observe how the three ML approaches fared when applied to the Credit Data. The lower values of the three parameters plainly indicate that the Neural Network is the best option. The lower values show that the Neural Network model does a better job at predicting fraud, and there is less of a discrepancy between the actual and predicted values.

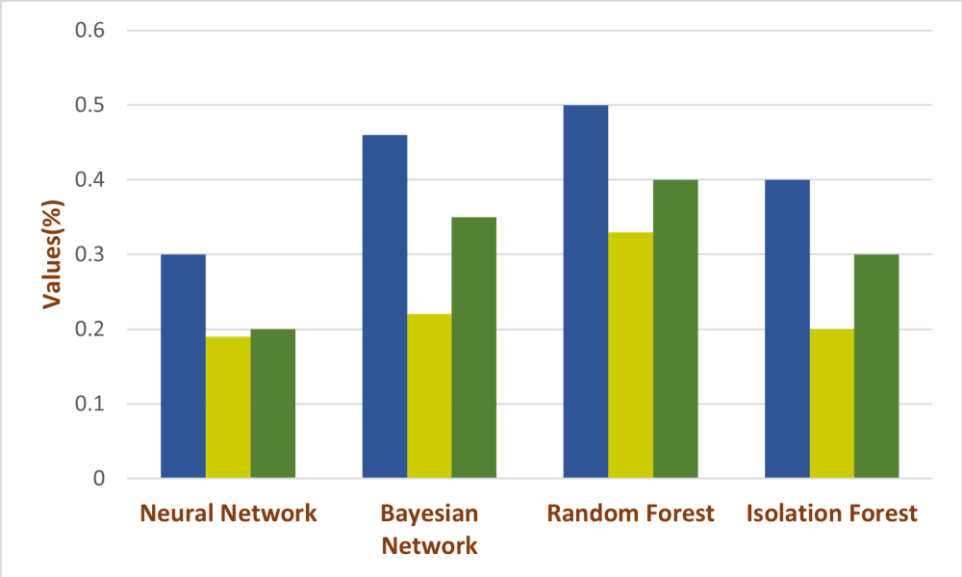


Fig. 5. RMSE, MSE and MAE Comparison of Different Models

Normal and fraudulent transactions are depicted in Figure 6. Data regarding the frequency of each class is presented in the chart.

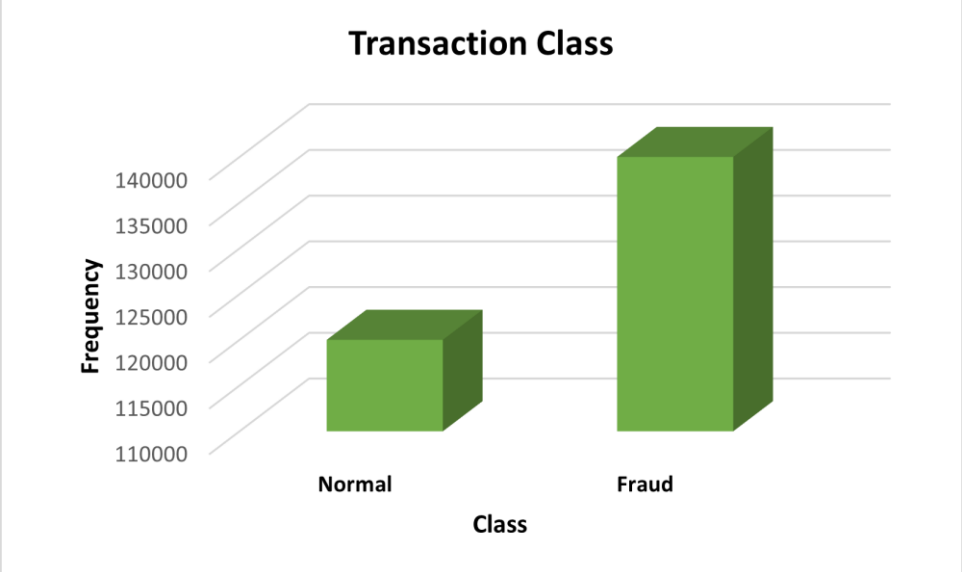


Fig. 6. ROC Curve of the Proposed Model

## V. Conclusion and Future Directions

Businesses and consumers engage in electronic commerce when they transact business through the purchase and sale of goods and services as well as the transmission of information. The growth of online shopping has been phenomenal, and it shows no signs of slowing down. There is a yearly loss of hundreds of millions of dollars due to fraud since the

prevalence of possible fraudulent actions has increased with the advent of e-commerce. Lots of data is kept and transported from one place to another nowadays, thanks to the Internet and E-commerce. Criminals may gain unauthorized access to send data. Fraud is on the rise, costing businesses and governments throughout the world billions of dollars annually. A number of contemporary methods for identifying fraudulent activity are often suggested and used in various industries.

REFERENCES

[1] J. Song, X. Qu, Z. Hu, Z. Li, J. Gao, and J. Zhang, "A subgraph-based knowledge reasoning method for collective fraud detection in E-commerce," *Neurocomputing*, no. xxxx, 2021, doi: 10.1016/j.neucom.2021.03.134.

[2] S. Ray, "Fraud Detection in E-Commerce Using Machine Learning," *BOHR Int. J. Adv. Manag. Res.*, vol. 1, no. 1, pp. 7–14, 2022, doi: 10.54646/bijamr.2022.02.

[3] and S. M. Porwal, Utkarsh, "Credit card fraud detection in e-commerce: An outlier detection approach," 2018.

[4] L. Akoglu, R. Chandy, and C. Faloutsos, "Opinion fraud detection in online reviews by network effects," *Proc. 7th Int. Conf. Weblogs Soc. Media, ICWSM 2013*, pp. 2–11, 2013, doi: 10.1609/icwsm.v7i1.14380.

[5] N. Kumar, D. Venugopal, L. Qiu, and S. Kumar, "Detecting Anomalous Online Reviewers: An Unsupervised Approach Using Mixture Models," *J. Manag. Inf. Syst.*, vol. 36, no. 4, pp. 1313–1346, 2019, doi: 10.1080/07421222.2019.1661089.

[6] P. Kaghazgaran, J. Caverlee, and M. Alfifi, "Behavioral analysis of review fraud: Linking malicious crowdsourcing to Amazon and beyond," *Proc. 11th Int. Conf. Web Soc. Media, ICWSM 2017*, no. Icwsm, pp. 560–563, 2017, doi: 10.1609/icwsm.v11i1.14953.

[7] W. Kudo, M. Nishiguchi, and F. Toriumi, "GCNEXT: graph convolutional network with expanded balance theory for fraudulent user detection," *Soc. Netw. Anal. Min.*, vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1007/s13278-020-00697-w.

[8] Y. Liu, Z. Sun, and W. Zhang, "Improving Fraud Detection via Hierarchical Attention-based Graph Neural Network," pp. 1–11, 2022.

[9] M. N. Ashtiani and B. Raahemi, "Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review," *IEEE Access*, vol. 10, no. July 2021, pp. 72504–72525, 2022, doi: 10.1109/ACCESS.2021.3096799.

[10] Idan Achituve Sarit Kraus Jacob Goldberger, "INTERPRETABLE ONLINE BANKING FRAUD DETECTION BASED ON HIERARCHICAL ATTENTION MECHANISM," 2019.

[11] C. Wang, "The behavioral sign of account theft: Realizing online payment fraud alert," *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 2021-Janua, pp. 4611–4618, 2020, doi: 10.24963/ijcai.2020/636.

[12] Y. Zeng, "applied sciences RLC-GNN : An Improved Deep Architecture for Spatial-Based Graph Neural Network with Application to Fraud Detection," 2021.

[13] Sakinah, "How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark," *Eng. Appl. Artif. Intell.*, vol. 76, pp. 130–157, 2019.

[14] J. S. Chang and W. H. Chang, "Analysis of fraudulent behavior strategies in online auctions for detecting latent fraudsters," *Electron. Commer. Res. Appl.*, vol. 13, no. 2, pp. 79–97, 2014, doi: 10.1016/j.elerap.2013.10.004.

[15] W. H. Chang and J. S. Chang, "A novel two-stage phased modeling framework for early fraud detection in online auctions," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11244–11260, 2011, doi: 10.1016/j.eswa.2011.02.172.

[16] J. Wang, R. Wen, C. Wu, Y. Huang, and J. Xiong, "FDGars: Fraudster detection via graph convolutional networks in online app review system," *Web Conf. 2019 - Companion World Wide Web Conf. WWW 2019*, pp. 310–316, 2019, doi: 10.1145/3308560.3316586.

[17] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," *IEEE Access*, vol. 11, no. January, pp. 3034–3043, 2023, doi: 10.1109/ACCESS.2022.3232287.

[18] R. Saia, "Unbalanced data classification in fraud detection by introducing a multidimensional space analysis," *IoTBDS 2018 - Proc. 3rd Int. Conf. Internet Things, Big Data Secur.*, vol. 2018-March, no. IoTBDS 2018, pp. 29–40, 2018, doi: 10.5220/0006663000290040.

[19] S. Kakkar, "Analysis of Discovering Fraud in Master Card based on Bidirectional GRU and CNN based Model," *2023 Int. Conf. Self Sustain. Artif. Intell. Syst.*, no. Icssas, pp. 50–55, 2023, doi: 10.1109/ICSSAS57918.2023.10331770.

[20] R. Rajkumar, N. Kogila, S. Rajesh, and A. R. Begum, "Intelligent System for Fraud Detection in Online Banking using Improved Particle Swarm Optimization and Support Vector Machine," in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Jun. 2023, pp. 644–649. doi: 10.1109/ICCES57224.2023.10192690.

[21] K. Yamini, V. Anitha, S. Polepaka, R. Chauhan, Y. Varshney, and M. Singh, "An Intelligent Method for Credit Card Fraud Detection using Improved CNN and Extreme Learning Machine," in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Jun. 2023, pp. 810–815. doi: 10.1109/ICCES57224.2023.10192774.

[22] P. J. Rana, "A Survey on Fraud Detection Techniques in Ecommerce," vol. 113, no. 14, pp. 5–7, 2015.

[23] R. Jhangiani, D. Bein, and A. Verma, "Machine Learning Pipeline for Fraud Detection and Prevention in E-Commerce Transactions," *2019 IEEE 10th Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2019*, pp. 0135–0140, 2019, doi: 10.1109/UEMCON47517.2019.8992993.

[24] E. Caldeira, G. Brandao, and A. C. M. Pereira, "Fraud analysis and prevention in e-commerce transactions," *Proc. - 9th Lat. Am. Web Congr. LA-WEB 2014*, no. December, pp. 42–49, 2014, doi: 10.1109/LAWeb.2014.23.