Volume (1) No (12), 2025
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Research Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Smart Fraud Detection in Online Retail: Leveraging Classification Algorithms for E-Commerce Transactions

Mr. Himanshu Barhaiya
Department of Master Computer Application
LNCT
Bhopal MP India
hemansu008@gmail.com

***Abstract*—The quick increase in e-commerce transactions through the internet has turned fraud detection into a major obstacle to e-commerce operations' security and reliability. This paper suggests a good DNN-based framework for intelligent fraud detection, which is built on the analysis of a very imbalanced credit card fraud dataset. This process consists of numerous data preprocessing steps, which involve, but are not limited to, how the missing values are treated and how duplicates and outliers are eliminated, as well as how z-score normalization can be used to obtain a homogeneous scale. To address the problem of class imbalance, SMOTE is applied, and a minority fraud is oversampled, and Principal Component Analysis (PCA) is applied, and the most beneficial attributes are obtained. The improved data is divided into training and testing (80:20) before training the DNN model. The results of the experiment indicate that the given DNN is more advantageous compared to traditional machine learning algorithms, including ANN, XGBoost, and Decision Tree in terms of accuracy, precision, recall, and F1-score of 97.10%. Overall, the proposed DNN-based system is effective in understanding the complex fraudulent activities in online shopping, it is also highly accurate in detection and offers a scalable solution that can be used to manage the contemporary e-commerce fraud prevention.**

***Keywords*—Machine Learning, Fraud Detection, Financial Institutions, Deep Learning, Data Security, Transaction.**

I. INTRODUCTION

E-commerce has greatly changed the way most individuals purchase and how the traditional buyer-seller relationships are conducted [1]. Simultaneous digitalization and automation that have occurred in the e-commerce sector have allowed the sector to become better in terms of efficiency, performance and resilience [2]. In the digital platform, many anonymous and non-anonymous individuals that have voluntarily registered on the e-commerce platforms can be easily and targeted reached [3][4]. The e-commerce has been accompanied by an increase in fraud within this industry. Even though fraud may sometimes be perpetrated by buyers, it is in fact more common in cases where the sellers are the ones who commit fraud. Consequently, the fraud in e-commerce is two-sided. There are numerous methods that are used by fraudulent customers. Some leave fake addresses of delivery, refuse to take goods delivered to them due to their change of mind, etc. These actions render the vendors to deal with extremely difficult [5][6][7]. Nonetheless, there is one more important concern of online buyers, namely fraudulent behavior of retailers. There is a variety of ways in which sellers can deceive buyers [8][9]. They might sell faulty or damaged products, miss on deliveries, or even make efforts to deliver goods without making the payments even after receiving some money in advance. Another common trick [10][11] is to fake things and present them as more appetizing than they actually are. Such fraudulent practices undermine customer trust and could lead to the buyers incurring losses. Fraud is the intentional failure to adhere to the truth in order to make a financial gain by a person or group of people [12].

The experts indicate that one of the most promising solutions is machine learning (ML) [13]. They support their decision by the fact that the process of detection could be automated by the ML algorithms, and the accuracy could be improved, and the number of detection times could be reduced to a significant level with the help of large amounts of data and learning the previous patterns of fraud [14][15]. In addition, the adaptability of the ML algorithms allows them to continue learning and evolving, which is highly significant when discussing the ever-evolving world of financial fraud, where fraudsters are continuously working on new ways of escaping notice [16][17]. However, there are challenges associated with using ML technology to identify financial fraud. The success of the application relies a lot on data quality and volume, the choice of proper algorithms and the real-time deployment of these systems [18]. In addition to that, integrating ML into the current financial infrastructure requires substantial investment in both technology and human resources. This research makes several key contributions to the field of Online Retail for E-Commerce [19]:

- Design and deployment of an all-inclusive DL-based fraud detection system using a customized Deep Neural Network that can understand and capture intricate nonlinear behaviors in e-commerce transactions.
- The combination of several pre-processing methods, namely duplicates removal, outlier removal, z-score normalization, PCA for feature extraction, and SMOTE for class balancing to improve the reliability of the model was done.
- Designed the system with scalability in mind, making it appropriate for use in online retail settings for real-time fraud detection.

- Assessed model performance comprehensively using several assessment indicators, such as precision, accuracy, recall, and F1-score.

A. *Motivation of this Study*

The e-commerce quick rise has given consumers the convenience factor but at the same time it has also resulted in the fraudulent activities that are complex and high in number. The consumer and the seller engage in cheating to such an extent that it includes making false delivery claims and refusing to accept shipments, selling damaged goods or keeping products that are not supposed to be sold, thus severely decreasing the trust among people using online marketplaces. The traditional rule-based fraud detection systems are unsuitable for the modern digital platforms since they are unable to keep up with the evolving fraud tendencies. Among the different techniques that come under machine learning, deep learning is the most powerful one as it has the ability to find and unravel the hidden patterns thus increasing the overall detection accuracy which is the driving factor for having a strong and scalable fraud detection framework to ensure e-commerce operations secures.

B. *Justification* and Novelty

The basis of this research is the increase in the intricacy of cheating through the Internet in the retail sector, where standard machine learning methods usually fail to discern the fine, nonlinear patterns hidden in the high-dimensional transactional data. The proposed system closes these gaps by proposing a new combination of stringent pre-processing (like duplicate deletion, outlier elimination, and z-score normalization), SMOTE for class balance and PCA for dimensionality reduction, all of which are tuned to provide a deep learning model's performance boost. This study's primary novelty is the development of a potent Deep Neural Network (DNN) model especially for fraud detection that can learn complex representations on its own and much outperform the traditional methods. By using cutting-edge methods for data processing with a powerful deep learning architecture, this research offers a strong, precise, and real-time-ready solution for the detection of frauds in the e-commerce sector.

II. Literature Review

Extensive research studies on Smart Fraud Detection have been thoroughly reviewed and analyzed to guide and strengthen the development of this study.

Chopra *et al.* (2025) the initial model (represented by blue bars) is performing quite well (for instance, around 88% in accuracy, 90% in precision, 92% in recall, and 88% in F1-score), supporting the conclusion of good detection albeit with the problem of balancing FP and FN. On the contrary, the second model (represented by orange bars) gives even better metrics with around 94-95% accuracy, 93% precision, 95-96% recall, and 92-93% F1-score, thus signaling an excellent ability to tell valid transactions from fraudulent ones with very few misclassifications [20].

Fnu and Murri (2025) A comprehensive approach for spotting retail fraud is introduced by means of combined use of FP-Growth algorithm alongside ML models. The entire process of analysis entails the conduct of preprocessing of UCI's Online Retail II dataset in order to extract features that are going to be used for Customer Engagement Index (CEI) assessment of customer behavioral pattern. The implementation of the algorithms GMM, OPTICS, and DBSCAN is aimed at distinguishing between the normal and the abnormal customer behaviors while PCA is applied to make the visual representation easier by reducing the dimensions. The study explored LDA and GNB models, the latter of which exhibited remarkable performance with 96.27% accuracy and 96.49% precision, as well as 96.27% recall and 96.26% F1-score through GNB [21].

Mutemi and Bacao (2024) A text-based fraud detection framework is introduced which can handle these losses in a very effective manner. The text speaks of the four main parts of the framework: text preprocessing, representation, machine learning techniques for knowledge extraction, and model assessment. The approach uses data augmentation techniques to improve the classifiers' efficacy in identifying fraud. The recommended Fast Text and Random Forest classifier combination outperforms conventional keyword-based models, achieving an astounding AUC of 0.99 and an F1 score of 0.833 on the enlarged dataset. This scalable system, which is based on the best practices in fraud detection, address the issue of fraud, which is already being exacerbated by the explosive rise of online shopping [22].

Banoth and Madhavi (2024) developed a deep learning (DL)-based method using the Kaggle dataset. They provide a unique text2IMG conversion method that may produce tiny images. In order to correct class imbalance, the pictures are fed into a convolutional neural network (CNN), which uses the inverse frequency approach to compute class weights. By using DL and ML approaches, the suggested system's validity and resilience were confirmed. A 99.87% accuracy rate was achieved by the suggested CNN by taking advantage of its deep characteristics [23].

Gupta *et al.* (2024) The research paper examines the different machine learning algorithms in terms of efficiency, specifically with their application to the highly unbalanced analysis of September 2013 credit card transactions made by European cardholders. The first findings show that the best results are obtained using the Isolation Forest method, achieving a detection accuracy of 99.38%, which is better than 99.26% for LOF and 54.82% for SVM, pointing out its effectiveness in separating fraudulent transactions with less stringent conditions [24].

Mutemi and Bacao (2023) Fraud detection has been a topic of extensive research in different areas, mostly in financial services. However, organized retail crimes research is still quite limited and rare in the literature. In an effort to provide more information on this subject, the authors put forth a scalable machine learning technique that is able to detect and separate ORC listings on a major online marketplace where it is the merchants who are engaged in organized retail crimes or fraud. The optimal detection model achieves a recall score of 0.97 on the holdout set and 0.94 on the out-of-sample testing set. These findings are derived from a small subset of 45 characteristics selected from a total of 58 features [25].

Table I summarizes recent research on smart fraud detection, including novel models, datasets, important discoveries, and difficulties encountered.

TABLE I. OVERVIEW OF RECENT STUDIES ON SMART FRAUD DETECTION IN ONLINE RETAIL FOR E-COMMERCE TRANSACTIONS USING MACHINE LEARNING

Author	Proposed Work	Dataset	Key Findings	Challenges / Limitations
Chopra et al. (2025)	Compared two fraud detection models using classification metrics.	Kaggle	Model 1 achieved ~88% accuracy, 90% precision, 92% recall, 88% F1; Model 2 achieved 94–95% accuracy, 93% precision, 95–96% recall, 92–93% F1, showing superior discrimination ability.	False positives and false negatives were still difficult to balance, and the dataset's specifics were not completely revealed
Fnu & Murri (2025)	Fraud detection using FP-Growth with ML models (GMM, OPTICS, DBSCAN, PCA, LDA, GNB).	UCI Online Retail II Dataset	GNB achieved outstanding performance: 96.27% accuracy, 96.49% precision, 96.27% recall, 96.26% F1	High dimensionality required PCA; clustering methods sensitive to noise; CEI computation complexity
Mutemi & Bacao (2024)	Text-based fraud detection framework using text preprocessing, augmentation, and ML classifiers.	Textual retail/online fraud dataset	Fast Text + Random Forest achieved F1 = 0.833 and AUC = 0.99; outperformed keyword-based models	Text augmentation complexity; domain-specific language differences may reduce generalizability
Banoth & Madhavi (2024)	Deep learning-based fraud detection using novel Text-to-Image conversion and CNN with inverse frequency class weights	Kaggle dataset	Proposed CNN achieved 99.87% accuracy; demonstrated robust performance under class imbalance	Text-to-Image method requires high computation; applicability to non-text data not tested
Gupta et al. (2024)	Comparative evaluation of ML models on imbalanced credit card fraud dataset	European Credit Card Fraud Dataset (Sept 2013)	Isolation Forest achieved highest accuracy at 99.38%, outperforming LOF (99.26%) and SVM (54.82%)	Severe class imbalance; accuracy alone may not reflect minority fraud detection effectiveness
Mutemi & Bacao (2023)	ML-based system for identifying online marketplace listings related to Organized Retail Crime (ORC)	Proprietary ORC dataset (45 selected features from 58)	Best model achieved recall of 0.97 (holdout) and 0.94 (out-of-sample), showing high fraud detection accuracy	Limited dataset size; ORC listings are rare → model performance may vary on larger platforms

III. RESEARCH METHODOLOGY

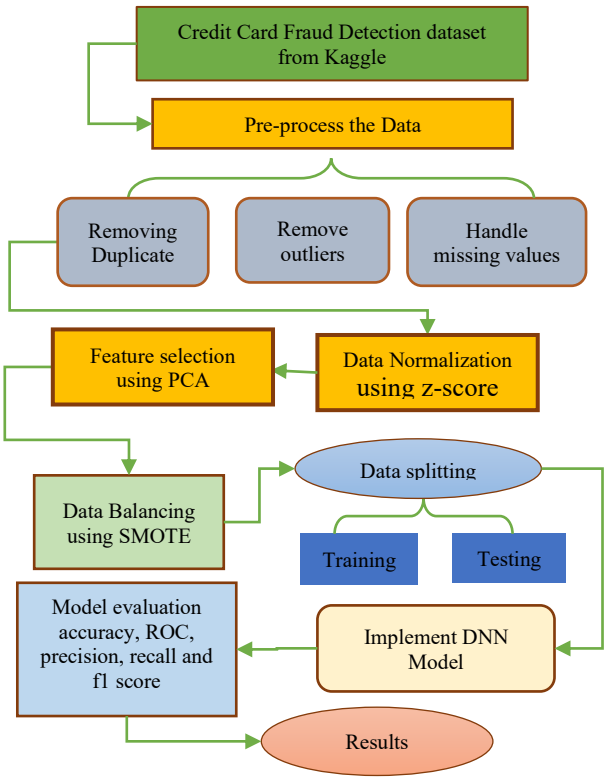

Fig. 1. Machine learning-based Smart Fraud Detection in Online Retail flowchart

The suggested methodology of Deep Neural Network (DNN) for CCFD. The initial data was heavily pre-processed which included imputing the missing values, removing the duplicates, and eliminating the outliers, Z-score normalization came next. For feature extraction, Principal Component Analysis (PCA) was employed to minimize dimensionality while preserving the most significant components. The problem of class imbalance was addressed by increasing the number of samples in the minority class using the Synthetic Minority Over-Sampling Technique (SMOTE). Furthermore, 80% of the dataset was used for training and 20% for testing. The produced dataset, which included the most crucial characteristics to reliably differentiate between fraudulent and genuine transactions, was utilized to train the DNN model. Figure 1 depicts the flowchart of ML-based Smart Fraud Detection in Online Retail.

The suggested flowchart for Smart Fraud Detection in Online Retail, which focuses on E-Commerce transactions, is explained in detail below.

A. Data Collection

The dataset utilized was the Kaggle Credit Card Fraud Detection. The dataset contained two days' worth of transactions conducted by European credit cards. Out of 284,807 transactions throughout the designated period, 492 fraudulent transactions were found in the dataset. The following data visualizations, which include bar graphs and heatmaps were utilized to examine feature correlations, fraud distribution, etc.:

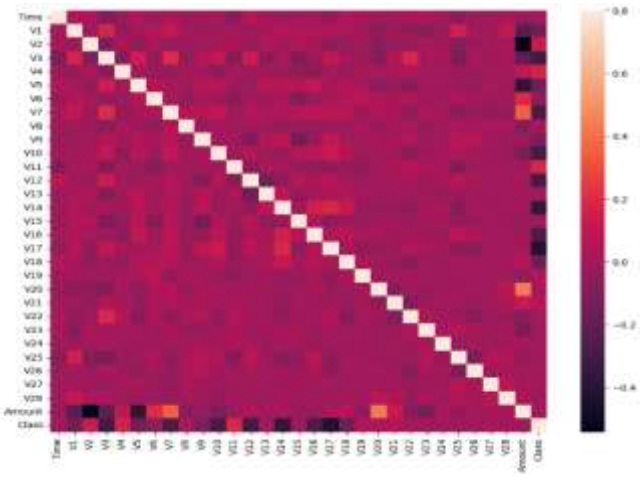

Fig. 2. Features correlation heatmap

In Figure 2, a correlation heatmap illustrates the characteristics of a dataset used to detect credit card fraud. A diagonal indicates that each feature is perfectly correlated with itself (correlation coefficient value of 1). Most of the features, particularly the ones whose identities have been hidden (V1 to V28), are only weakly correlated with one another, indicating a low level of multicollinearity. However, correlations of the target variable "Class" with the features "Amount" and "Time" are minor but observable, and thus they are likely to affect the predictive models of fraud.

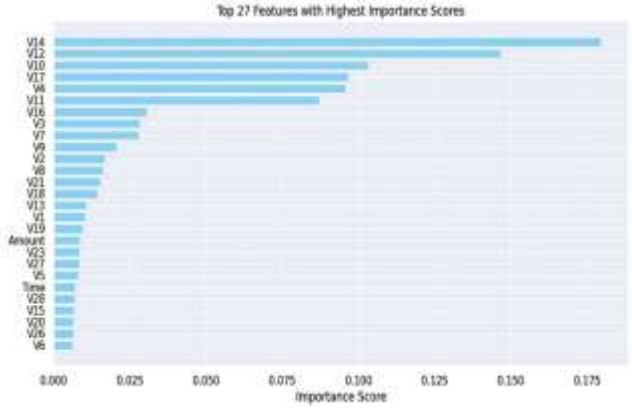

Fig. 3. Top 27 features with the highest importance score

The bar chart in Figure 3 shows the top 27 features that have the most impact in a fraud detection model as indicated by their importance scores. The feature V14 tops the list as the most powerful, next come V12, V10, V4, and V11, whose contributions, in fact, are disproportionately more than the rest of the feature set. The majority of the anonymized predictors (V1 to V28) are the leading ones, while the "Amount" and "Time" show a much less degree of significance.

B. Data Pre-processing

The initial step for data preparation was the gathering of the CCFD dataset, after which concatenation and data cleaning were performed. Data pre-processing also included transformation and normalization of the data to achieve uniformity and better model performance. The comprehensive pre-processing operations are described as follows:

- **Handle missing values:** Missing values can be dealt with in more than one way [26]. A complex imputation procedure, filling in missing values with zeros or the mean, and just eliminating the cases with missing data are some of the several strategies.
- **Removing Duplicate:** The first step in data pre-processing is removing duplicate entries, which is essential to provide data quality, consistency, and accuracy for later analysis or model training [27]. Duplicate records may come from various sources, like mistakes made in data entry, joining of data from different sources, or bad data collection methods.
- **Remove Outliers:** Data input inconsistency, wrong observations, or very extreme cases are the usual reasons for outliers, which are the points in the dataset that show the largest difference compared to the main data.

C. Data Normalization using z-score

The process of transforming or standardizing data to have a comparable distribution is called data normalization [28]. The most popular method for normalizing data is called min-max normalization and z-score normalization. z-score normalization, a standardization method using 0 as the mean and 1 as the standard deviation, has been used in this investigation. By using a unit standard deviation, this scaling strategy modifies the values that are centered around the average value. Equation (1) gives the definition of the z-score normalization.

$$E' = \frac{E - \bar{M}}{\sigma_M} \quad (1)$$

Where,

$\bar{M}$ is the mean, σ_M is the standard deviation, and E' and E are new and old for every data item.

D. Feature selection using PCA

Feature selection is an important step wherein the relevant variables are selected from a dataset through understanding [29]. The research includes many features, but have picked only those that are needed for improving performance measurement. A helpful method for choosing features is Principal Component Analysis (PCA), by revealing and uses the major components that account for the greatest variance in the dataset.

E. Data Balancing using SMOTE

Data balancing can be classified as one of the major techniques in ML, especially in cases where an imbalance among classes is present, i.e., There are noticeably fewer samples in one or more classes than in others [30]. SMOTE, though a statistical approach, still effectively manages to double the minority class samples in a dataset, thus making it balanced. The process that was involved included the algorithm creating new examples out of the existing, less represented classes that were given as input.

F. Data Splitting

The dataset was split into samples for testing (20%) and training (80%). The model was then trained using the training sample.

G. Proposed Deep Neural Network (DNN) Model

A well-known DL method among academics is the deep neural network (DNN). The DNN's network structure is composed of input, hidden, and output layers, all of which are fully connected [31]. In the layer that follows, every neuron is coupled to every other neuron, however these neurons are not connected to each other across layers. The output is subject to an activation function after every network layer, which amplifies the impact of network learning [32]. Consequently, DNN may alternatively be viewed as a massive perceptron composed of several smaller perceptron's. For example, the formula for computing ith layer forward propagation is Equation (2):

$$x_{i+1} = \sigma(\sum w_i x_i + b) \quad (2)$$

where the input value is denoted by x, the weight coefficient matrices by w, and the bias vector by b. ReLU is typically employed as an activation function in a multi-class network; the formula is as follows, Equation (3):

$$\sigma(x) = max(0, x) \quad (3)$$

The formula for choosing cross-entropy as a classification loss function is as follows (4):

$$C = -\frac{1}{N}\sum_x \sum_{i=1}^{M}(y_i log p_i) \quad (4)$$

where N is the number of categories, M is the number of input data sets, yi is the chance of classifying into category i, and pi is the probability that the actual category and the categorization i match.

H. Evaluation Metrics

A number of performance indicators were used to assess the suggested architecture's efficacy [33]. The actual values were compared with the trained models' anticipated results. True-Negatives (TN), False-Negatives (FN), True-Positives (TP), and False-Positives (FP) were established through this comparison. An overview of the matrix's accuracy, precision, recall, and F1-score can be seen below:

Accuracy: The percentage of data points from the dataset (given input samples) that the trained model accurately predicted. It is expressed in the form of Equation (5):

$$Accuracy = \frac{\text{TP+TN}}{\text{TP+Fp+TN+FN}} \quad (5)$$

Precision: The accuracy of a model is calculated by dividing the total number of positive events by the number of properly anticipated positive cases. The capacity of the classifier to forecast the positive classifications is shown by Equation (6):

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \quad (6)$$

Recall: This measure is the proportion of correctly anticipated positive outcomes to all cases that ought to have produced positive results. It is expressed mathematically as Equation (7):

$$Recall = \frac{\text{TP}}{TP+FN} \quad (7)$$

F1 score: It assists in regulating memory and accuracy by combining the two metrics' harmonic mean. Its range is [0,1]. It is expressed mathematically as Equation (8):

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

Where TP is the quantity of transactions that were accurately identified as fraudulent. Transactions appropriately identified as non-fraud are denoted with TN. FN is the number of fraudulent transactions that were mistakenly recorded as non-fraudulent. FP is the quantity of genuine transactions that are incorrectly labelled as fraudulent.

IV. Results and Discussion

Hardware support for the experiment was given through a cluster of different resources and cloud services. An Intel(R) Core (TM) i5-2520MCPU operating at 2.50 GHz and 12 GB of RAM were installed on the local computer to provide effective processing and memory for the activities at hand. The model was evaluated using the primary performance measures of accuracy, precision, recall, and F1-score following training on the Credit Card Fraud Detection dataset from Kaggle, as shown in Table II. At 97.10% accuracy, 98% precision, 97.10% recall, and 97.40% F1-score, the DNN model demonstrated outstanding performance and received high scores on all important assessment measures. Those results suggest that the model was very effective in spotting the fraudulent transactions and at the same time maintaining an equilibrium between recall and precision.

TABLE II. Experiment Results of Proposed Models for of Smart Fraud Detection in Online Retail on Credit Card Fraud Detection Dataset

Performance matrix	Deep neural network (DNN)
Accuracy	97.10
Precision	98
Recall	97.10
F1-score	97.40

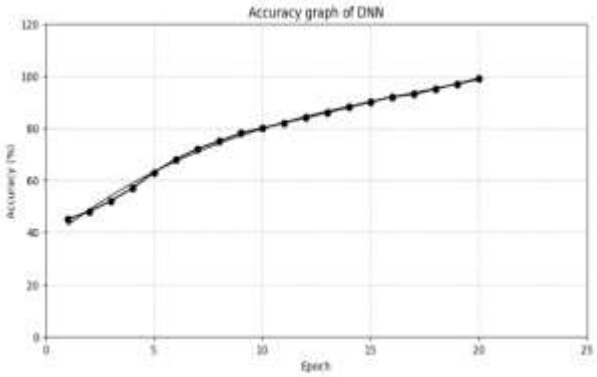

Fig. 4. Accuracy curves for the DNN Model

Figure 4 shows the recommended DNN model's training accuracy for each training period. The model begins with an initial accuracy of approximately 45 and keeps increasing steadily, which is evidence of good learning and better performance as the number of epochs grows. Cloning and predicting the transactions is accurate, and the model has a 99 percent accuracy by the 20th epoch.

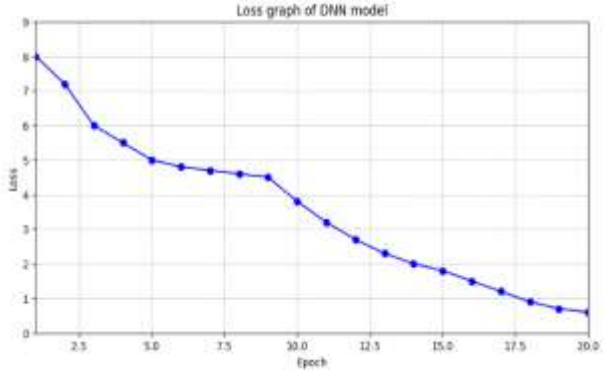

Fig. 5. Loss curves for the DNN Model

The suggested Deep Neural Network (DNN) model's loss decrease across 20 training epochs is displayed in Figure 5. The loss of the model is very high at the beginning, about 8, which means there is a big prediction error. The loss is then consistently decreased at each training period, which is a convincing demonstration of the model's capacity to learn and reduce error. By the 20th epoch, the loss is less than 1, which means that the model has become very close to the ideal prediction and that its performance has improved significantly.

A. Comparative Analysis

Online retail performance comparison of predictive models of smart fraud detection indicates that deep learning methods are usually better than the traditional algorithms in table II. Models such as ANN, and XGBoost have moderate accuracy which implies that they are limited to depicting sophisticated patterns of fraud. The SVM and MLP performance are more robust and balanced with high precision and recall and would be applicable in real-time instances of

fraud detection. The weakness of the Decision Tree model is however that it has low recall even though it is very accurate implying that it is weak in detecting fraudulent transactions. All in all, the Deep Neural Network (DNN) is the most effective model to use in determining fraud in e-commerce transactions with an accuracy of 97.10% and good metric balance.

TABLE III. ACCURACY COMPARISON OF DIFFERENT PREDICTIVE MODELS OF SMART FRAUD DETECTION IN ONLINE RETAIL FOR E-COMMERCE TRANSACTIONS

Models	Accuracy	Precision	Recall	F1-score
ANN[34]	88.93	82.40	78.76	80.54
SVM[35]	94.9	95.9	95.1	95.1
MLP[36]	95.8	97.6	93.9	95.8
DT[37]	96.5	83	64	72
XGBoost [38]	88	88	86	87
DNN	97.10	98	97.10	97.40

The suggested framework makes use of a Deep Neural Network (DNN) structure, which is very smart for fraud detection and is capable of accurately and reliably capturing very complex and nonlinear patterns in online retail transactions. By getting to know the very subtle relationships among the features, the model very much cuts down the number of FP as well as FN, assuring more accurate fraud detection. Its main benefits are: generalization that is better than others, robustness to noisy or unbalanced data, irrespective of data quality, and the capability to automate. With low human feature engineering, high-level representations may be obtained. Thus, DNN model is highly effective and highly scalable to real-time fraud detection in the present online shopping environment.

V. CONCLUSION AND FUTURE STUDY

The transaction volumes are the largest ever witnessed through the online retail platforms and the result has been that fraud detection has become the most essential condition that needs to be fulfilled so that e-commerce operations can be safe and reliable. The current study provided the detailed and well-organized DNN-based intelligent fraud detection system using the Credit Card Fraud Detection dataset. The data was highly pre-processed, including: missing value treatment, elimination of duplicates and outliers, z-score normalization, use of PCA to extract features and class balancing by use of SMOTE. Once these steps have been performed, it turned out to be an effective input of deep learning. The recommended DNN managed to predict quite successfully, with 97.10 percent accuracy, 98 percent precision, 97.10 percent recall, and 97.40 percent F1-score, significantly higher than the performance of such traditional models as ANN, XGBoost, and Decision Tree. It was also demonstrated that the model had a high learning capacity, rapid convergence and stability throughout the training of the model as indicated by the performance curves. Therefore, it can be stated that DNNs are highly effective in identifying contemporary internet retail deal frauds that are highly complex and nonlinear, and they do not only identify it, but also offer greater generalization, adaptability, and reliability.

The development of future research on this work could be guided to several promising directions. Firstly, more advanced deep learning models such as CNN-LSTM composites or Transformer-based models can also be tried to ensure that the sequential analysis of the flows of transactions becomes even more compelling. Also, the use of explainable AI (XAI) techniques can be mentioned as one of the ways of providing transparent interpretation of the fraud decisions that will consequently result in the rise of trust and compliance with legal regulations.

REFERENCES

[1] A. Srivastava, P. K. Bala, and B. Kumar, "New perspectives on gray sheep behavior in E-commerce recommendations," *J. Retail. Consum. Serv.*, 2020, doi: 10.1016/j.jretconser.2019.02.018.

[2] K. M. R. Seetharaman, "Digital Transformation in Retail Sales: Analyzing the Impact of Omni-Channel Strategies on Customer Engagement," *J. Glob. Res. Math. Arch.*, vol. 10, no. 12, 2023, doi: 10.5281/zenodo.15280578.

[3] M. Zhou *et al.*, "Understanding consumers' behavior to adopt self-service parcel services for last-mile delivery," *J. Retail. Consum. Serv.*, vol. 52, p. 101911, Jan. 2020, doi: 10.1016/j.jretconser.2019.101911.

[4] S. K. Tiwari, "Quality Assurance Strategies in Developing High-Performance Financial Technology Solutions," *Int. J. data Sci. Mach. Learn.*, vol. 05, no. 01, pp. 323–335, Jun. 2025, doi: 10.55640/ijdsml-05-01-26.

[5] S. Wang, "A comprehensive survey of data mining-based accounting-fraud detection research," in *2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010*, 2010. doi: 10.1109/ICICTA.2010.831.

[6] A. Parupalli, "The Evolution of Financial Decision Support Systems: The Evolution of Financial Decision Support Systems: From BI Dashboards to Predictive Analytics," *KOS J. Bus. Manag.*, vol. 1, no. 1, pp. 1–8, 2025.

[7] N. Malali, "Exploring Artificial Intelligence Models for Early Warning Systems with Systemic Risk Analysis in Finance," in *2025 International Conference on Advanced Computing Technologies (ICoACT)*, IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/ICoACT63339.2025.11005357.

[8] S. K. Chintagunta and S. Amrale, "Enhancing Cloud Database Security Through Intelligent Threat Detection and Risk Mitigation," *Tech. Int. J. Eng. Res.*, vol. 9, no. 10, pp. 49–55, 2022, doi: 10.56975/tijer.v9i10.159996.

[9] B. R. Ande, "Federated Learning and Explainable AI for Decentralized Fraud Detection in Financial Systems," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 35s, pp. 48–56, 2025, doi: 10.52783/jisem.v10i35s.5921.

[10] S. B. V. Naga, S. Thangavel, S. K. Kuchoor, N. Narukulla, and L. K. Yenduri, "Optimizing Online Marketing Strategies with Machine Learning and Deep Learning Innovations," in *Impact of Digital Transformation on Business Growth and Performance*, 2025, pp. 483–512. doi: 10.4018/979-8-3693-9783-1.ch018.

[11] S. R. Kurakula, "Designing Enterprise Systems for the Future of Financial Services: The Intersection of AI, Cloud-Native Microservices, and Intelligent Data Processing," *Eur. J. Comput. Sci. Inf. Technol.*, vol. 13, no. 20, pp. 91–103, 2025.

[12] V. Verma, "Deep Learning-Based Fraud Detection in Financial Transactions: A Case Study Using Real-Time Data Streams," vol. 3, no. 4, pp. 149–157, 2023, doi: 10.56472/25832646/JETA-V3I8P117.

[13] T. Shah, "Leadership in digital transformation: Enhancing customer value through AI-driven innovation in financial services marketing," *Int. J. Sci. Res. Arch.*, vol. 15, no. 3, pp. 618–627, Jun. 2025, doi: 10.30574/ijsra.2025.15.3.1767.

[14] R. R. Popat and J. Chaudhary, "A Survey on Credit Card Fraud Detection Using Machine Learning," in *Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018*, 2018. doi: 10.1109/ICOEI.2018.8553963.

[15] P. Chandrashekar, "Data-Driven Loan Default Prediction: Enhancing Business Process Workflows with Machine Learning," *Int. J. Emerg. Res. Eng. Technol.*, vol. 6, no. 4, pp. 18–26, 2025.

[16] K. B. Thakkar and H. P. Kapadia, "The Roadmap to Digital Transformation in Banking: Advancing Credit Card Fraud Detection with Hybrid Deep Learning Model," in *2025 2nd International Conference on Trends in Engineering Systems and Technologies (ICTEST)*, 2025, pp. 1–6. doi: 10.1109/ICTEST64710.2025.11042822.

[17] H. Kali, "Optimizing Credit Card Fraud Transactions

identification and classification in banking industry Using Machine Learning Algorithms," *Int. J. Recent Technol. Sci. Manag.*, vol. 9, no. 11, pp. 85–96, 2024.

[18] R. Jain, S. K. Das, and Y. Makin, "Behavioral Risk Tolerance in U.S. Retirement Planning Vs. Property Insurance: A Comparative Analysis," *Int. J. Appl. Math.*, vol. 38, pp. 41–70, 2025.

[19] K. M. R. Seetharaman, "Analysing the Role of Inventory and Warehouse Management in Supply Chain Agility: Insights from Retail and Manufacturing Industries," *Int. J. Curr. Eng. Technol.*, vol. 12, no. 06, pp. 583–590, Jun. 2022, doi: 10.14741/ijcet/v.12.6.13.

[20] P. Chopra, K. Jain, S. Gandhi, S. Mahadik, V. Ravalji, and P. Goel, "Developing Machine Learning Model for Real-Time Fraud Detection in Online Transactions," in *2025 International Conference on Networks and Cryptology (NETCRYPT)*, IEEE, May 2025, pp. 1568–1573. doi: 10.1109/NETCRYPT65877.2025.11102633.

[21] H. Fnu and S. Murri, "Algorithmic Approach for Fraudulent Transaction Detection using Market Basket Analysis with Big Data," in *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, IEEE, Apr. 2025, pp. 1–7. doi: 10.1109/ICDCECE65353.2025.11035609.

[22] A. Mutemi and F. Bacao, "Balancing act: Tackling organized retail fraud on e-commerce platforms with imbalanced learning text models," *Int. J. Inf. Manag. Data Insights*, vol. 4, no. 2, p. 100256, Nov. 2024, doi: 10.1016/j.jjimei.2024.100256.

[23] S. Banoth and K. Madhavi, "A Novel Deep Learning Framework for Credit Card Fraud Detection," in *2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART)*, 2024, pp. 191–196. doi: 10.1109/SMART63812.2024.10882509.

[24] P. Gupta, S. Shukla, V. Kikan, and A. Kumar, "Comparative Analysis of Machine Learning Algorithms for Detecting Fraudulent Transactions in Highly Imbalanced Credit Card Data," in *2024 IEEE International Conference on Smart Power Control and Renewable Energy (ICSPCRE)*, 2024, pp. 1–5. doi: 10.1109/ICSPCRE62303.2024.10675130.

[25] A. Mutemi and F. Bacao, "A numeric-based machine learning design for detecting organized retail fraud in digital marketplaces," *Sci. Rep.*, 2023, doi: 10.1038/s41598-023-38304-5.

[26] G. Sarraf and V. Pal, "Adaptive Deep Learning for Identification of Real-Time Anomaly in Zero-Trust Cloud Networks," vol. 4, no. 3, pp. 209–218, 2024, doi: 10.56472/25832646/JETA-V4I3P122.

[27] A. R. Bilipelli, "Forecasting the Evolution of Cyber Attacks in FinTech Using Transformer-Based Time Series Models," *Int. J. Res. Anal. Rev.*, vol. 10, no. 3, pp. 383–389, 2023.

[28] D. Patel, "Enhancing Banking Security: A Blockchain and Machine Learning- Based Fraud Prevention Model," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, pp. 576–583, 2023, doi: 10.14741/ijcet/v.13.6.10.

[29] R. Quddus Majumder, "Evaluating the Correlation Between Leverage and Profitability in the Retail Sector: A Comparative Study of Listed Companies Across Five Years," *Int. J. Res. Anal. Rev.*, vol. 12, no. 2, pp. 82–88, 2025, doi: 10.56975/ijrar.v12i2.314575.

[30] Y. Macha and S. K. Pulichikkunnu, "An Explainable AI System for Fraud Identification in Insurance Claims via Machine-Learning Methods," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 3, no. 3, pp. 1391–1400, Jul. 2023, doi: 10.48175/IJARSCT-11978X.

[31] V. Pal and S. K. Chintagunta, "Transformer-Based Graph Neural Networks for RealTime Fraud Detection in Blockchain Networks," pp. 1401–1411, 2023, doi: 10.48175/IJARSCT-11978Y.

[32] S. J. Wawge, "A Survey on the Identification of Credit Card Fraud Using Machine Learning with Precision, Performance, and Challenges," *Int. J. Innov. Sci. Res. Technol.*, vol. 10, no. 4, April, pp. 3345–3352, May 2025, doi: 10.38124/ijisrt/25apr1813.

[33] G. Mantha, "Transforming the Insurance Industry with Salesforce: Enhancing Customer Engagement and Operational Efficiency," *North Am. J. Eng. Res.*, vol. 5, no. 3, 2024.

[34] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *J. Big Data*, 2022, doi: 10.1186/s40537-022-00573-8.

[35] H. Sinha, "An examination of machine learning-based credit card fraud detection systems," *Int. J. Sci. Res. Arch.*, vol. 12, no. 2, pp. 2282–2284, Aug. 2024, doi: 10.30574/ijsra.2024.12.2.1456.

[36] K. Hayat and B. Magnier, "Data Leakage and Deceptive Performance: A Critical Examination of Credit Card Fraud Detection Methodologies," *Mathematics*, vol. 13, no. 16, p. 2563, Aug. 2025, doi: 10.3390/math13162563.

[37] P. Sundaravadivel, R. A. Isaac, D. Elangovan, D. KrishnaRaj, V. V. L. Rahul, and R. Raja, "Optimizing credit card fraud detection with random forests and SMOTE," *Sci. Rep.*, vol. 15, no. 1, p. 17851, May 2025, doi: 10.1038/s41598-025-00873-y.

[38] M. A. Alrasheedi, "Enhancing Fraud Detection in Credit Card Transactions: A Comparative Study of Machine Learning Models," *Comput. Econ.*, Aug. 2025, doi: 10.1007/s10614-025-11071-3.

