Applied Sciences Research Periodicals
September 2023, Vol. 1, No. 6, pp. 19-28

# Machine Learning Credit Card Fraud Detection System

Juliet Chinazo Onyema, Chidi Ukamaka Betrand, and Mercy Benson-Emenike

**Abstract:**
The Credit Card Fraud Detection system is a web-based fraud detector tool that can be used to flag potential fraud cases in daily transactions. It is vital that credit card companies are able to spot fraudulent transactions using the credit card so that customers are not charged for items that they did not purchase. This article illustrates the modeling of a data set using past credit card transactions with the data of the ones that turned out to be fraudulent which is achieved using machine learning. The model recognizes whether a new transaction is fraudulent or not by validating a user before any transaction is being made through sending a One-Time-Password (OTP) to the user, hence detects 90% to 100% of the fraudulent transactions thus minimizing financial crime. The Structured System and Design Methodology (SSADM) technique was employed in developing this system. In this process, we have focused on analyzing and pre-processing data sets as well as the deployment of multiple anomaly detection algorithm such as Isolation Forest algorithm and Local Outlier Factor algorithm. The model was built with Python and implemented in an e-commerce site which was built with HTML, CSS, JavaScript and its database as SQLite. After building and testing the system, it was discovered that in order for the system to be more accurate and precise, more transactional data need to be fed to the system.

*Keywords: Machine Learning, Credit Card, Detection, Fraud, Data Science, Algorithms.*

## INTRODUCTION

Recently, online transactions are on the increase due to high purchase of goods and services. In developed and developing countries, credit card is the most acceptable and common means of payment for online and offline transactions. It is discovered that some sensitive information about a credit card such as the credit card number, secure code, validity, Card Verification Value (CVV) number and name of card holder are often compromised so as carry out an illegal and fraudulent transactions which can lead to huge financial loss to both the issuing banks and to the individual owner [1].

Fraudulent Credit Card transaction is an illegal and unauthorized usage of one's account other than the rightful owner of the account. This article aims at reducing online credit card fraudulent activities drastically by implementing OTP code verification to validate every transaction as well as reporting fraudulent activities to the real card owners.

Another promising way of reducing successful credit card fraud is based on the analysis of existing purchase data of the cardholder [2]. Considering that humans tend to exhibit some specific behavioral profiles or patterns, card holders can as well be represented with a set of patterns containing information about the purchase details; the time, amount spent and so on. This helps to identify unusual patterns in fraudulent activity by combining real-time transactional data with the historical analysis of customer behavior which leads to complete auditing, transparency and traceability. However, deviation from such habitual patterns poses as a potential threat alert to the system.

Meanwhile, some loop-holes were discovered in the existing system, these include:

- No request of OTPs for subsequent online transactions once card details are saved online.
- Non confirmation of owner's details before making transactions.
- Non availability of 3D-security enabled in most cards.

In this article, the model used validates an individual before every transaction, analyzes the individual's spending pattern; how often the individual does online transactions (web, POS, Online) on a monthly basis, keeps record of every fraudulent credit card transaction, prevents card fraud as well as notifies an individual whenever there is a fraudulent attempt. This technique promises high predictive accuracy in fraud detection and prevention.

## BACKGROUND STUDY AND RELATED WORKS

### Overview of Credit Card

Credit Card is a thin rectangular piece of either metal or plastic issued by banks or financial institutions that allows cardholders to borrow funds with which to pay for goods and services with merchants that accept cards for payment [2]. On its front are the bank name, card number, card holder's name the chip and the expiry date then at its reverse are the magnetic strip, signature, hologram, and the Card Verification Code (CVC) as shown in Figure 1 below:


**Figure 1: Diagram of a Credit Card**

With credit cards, shopping online, reserving airline tickets and ordering from a catalog becomes a breeze. Mailing a cheque is almost a thing of the past as a credit card is faster, easier, and generally a more secure way of doing business. Sadly, it seems that fraudsters are keeping track and even thriving in this growing environment [3].

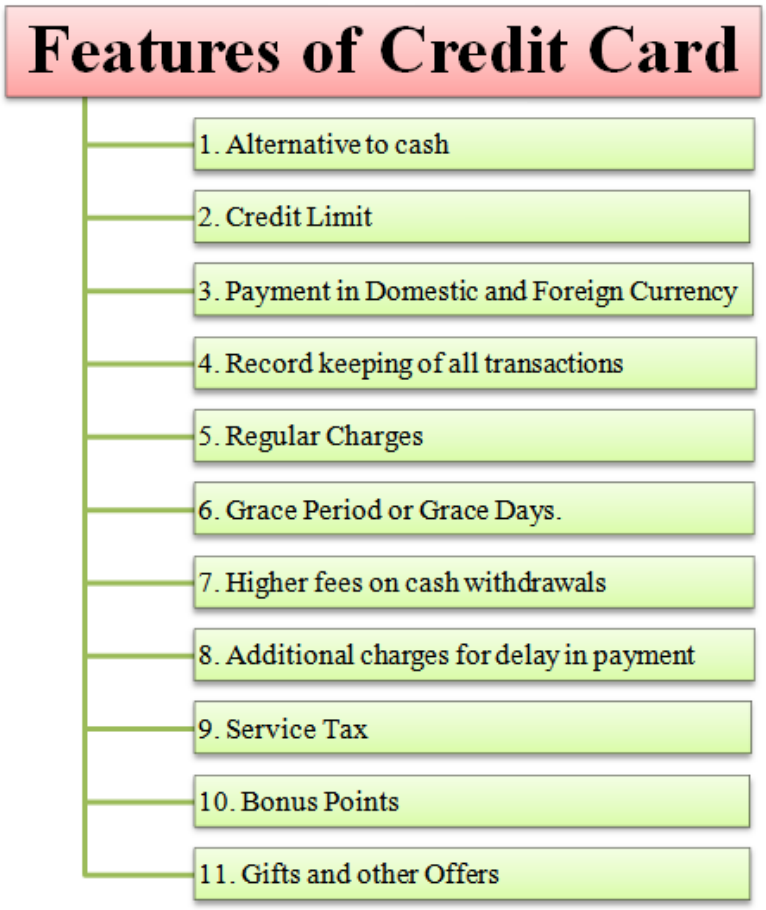


**Figure 2: Characteristics of Credit Card**

Though frauds can be online or offline [4], Credit card frauds are categorized in various ways but split into application frauds and behavioral frauds. In application frauds, the fraudsters apply for a credit card with a false ID whereas in behavioral frauds, the fraudsters find a way to obtain the cardholder's credential in other to use a pre-existing credit card. The fraudulent transactions are into six categories with respect to the fraudulent process:

1. Frauds from lost or stolen cards
2. Frauds from counterfeit cards
3. Online frauds
4. Bankruptcy frauds
5. Merchant frauds
6. Frauds from cards that got stolen during the expedition process.

However, an online fraud is committed through the internet, phone, shopping, web or in the absence of the card while an offline fraud is committed using a physical stolen credit card [4]. Tools have been developed to generate card numbers as shown in figure 3. These numbers are then used for "credit master attack", which consists in brute force attacking the merchant website with lots of possible cards [5].

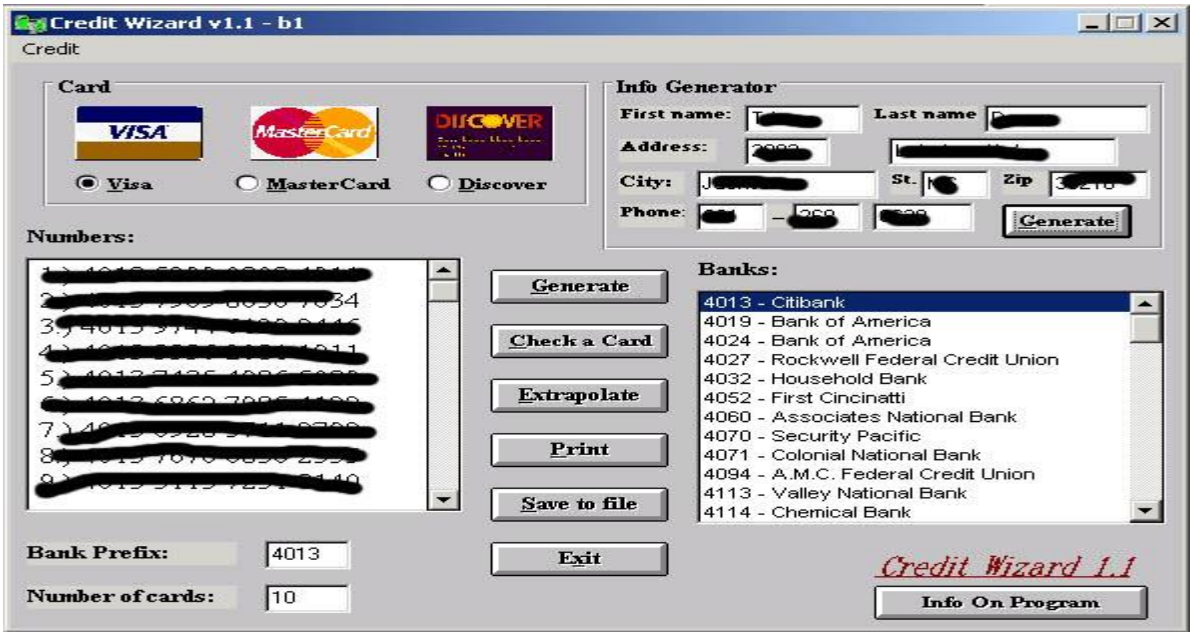


**Figure 3: Card Number Generator**

## Related Works

Fraud detection in Credit cards has drawn great attention from the scientific community, though countermeasures have been proposed. According to the most recent learning research, Machine Learning (ML), Support Vector Machines (SVM), Bayesian Networks (BN), Artificial Neural Networks (ANN), K-Nearest Neighbor (KNN), Hidden Markov (HM), Fuzzy Logic Systems (FLS) and Decision Trees (DT) are some of the techniques that are often employed for fraud detections [6].

In their research, the Support Vector Machines (SVM), Decision Trees, Logistic Regression and K-Nearest Neighbor algorithms offer medium accuracy while the two methods with the lowest accuracy are Fuzzy Logic and Logistic Regression. The Neural Networks, Naive Bayes, Fuzzy Systems, and KNN are of high detection rate. Meanwhile, ANN and Nave Bayesian Networks are two algorithms that consistently outperform each other. The drawback of these algorithms is that they do not always give the same result across different contexts. They perform better with some datasets while performing worse with others. While raw, unprocessed data yields decent accuracy with techniques like fuzzy logic systems and logistic regression, small datasets yield outstanding results with algorithms like KNN and SVM.

Another study was based on the transactions and data mining techniques, it examined data mining approaches such as Bayesian networks, Bayes Minimum Risk, evolutionary algorithms, Hidden Markov Models (HMM), and ontologies with the aim of lowering credit card fraud. Next, it focused on improving the efficiency of the techniques used to detect fraud by improving the prediction of fraudulent accounts. As a result of their findings, it was possible to detect fraud more effectively by combining a learning strategy with a widely used technique [7].

A hybrid feature selection and genetic algorithm for e-banking fraud detection was also suggested [8]. The method was created using the reinforcement learning component of the neural network, however, Whale algorithm was also studied. The whale algorithm and the recommended approach were compared; hence the results demonstrated that the suggested technique was extremely effective in identifying e-banking fraud.

However, in 2016, another method was devised for detecting credit card fraud using unsupervised algorithms [9]. A model for detecting credit card fraud was released to meet the demands of computation simplicity and operation transparency. Two unsupervised techniques, Principal Component Analysis (PCA) and SIMPLEKMEANS, were developed to account for the geographical location of both transactions and clients. The recommended method, according to the authors, may detect fresh instances of fraud and directly and precisely identify the transactions. PCA offered a flexible and comprehensive picture of the connections between multiple features.

Navanshu *et al* described logistic regression, decision trees, random forests, and SVM in their research. They worked with a dataset that had a significant skew. The standards used to assess performance include sensitivity, specificity, accuracy and precision. The outcome showed that the accuracy of Logistic Regression is 97.7%, that of Decision Trees is 95.5%, that of Random Forest is 98.6%, and that of SVM classifier is 97.5%. They determined that the Random Forest algorithm is the most accurate algorithm out of all algorithms and is the best algorithm for detecting fraud. The data imbalance problem led them to the additional conclusion that the SVM algorithm does not perform any better in terms of detecting credit card fraud [10].

The difficulties associated with financial card fraud detection and associated mechanisms are also used to create a number of supervised algorithms. The authors classified the most important technique explored in Raj and Portia's analysis of numerous approaches. It also evaluates every methodology in light of particular design criteria. A few useful techniques for detecting credit card fraud were compared and evaluated in the study. It focused on methods for spotting credit card fraud, including Bayesian learning and Dempster Shafer Fusion. Any card-issuing bank must have a reliable system in place for identifying credit card fraud [11].

In the course of comparing Local Outlier Factor and Isolation Factor algorithms got an accuracy of 97% by the former and 76% accuracy by the later [12]. Credit card fraud was detected with various supervised machine learning algorithms with real-world datasets [9]. Ensemble learning techniques are dependent and independent variables in order to improve the accuracy of identifying credit card fraud. Various supervised learning-based techniques are compared and discussed.

Although we expect that if these algorithms are trained with more real-world data, the effectiveness and prediction will improve [13]. All the techniques showed minimal variance in performance. In conclusion, fraudulent transactions which constitute significant loss for both clients and financial institutions have been tackled with various approaches with subpar results. There is therefore, greater need to create improved fraud detection technique using deep learning implementation. This will give an improved technique in financial card detection.

## EXPERIMENTAL RESULT AND ANALYSIS

The diagrammatic sequence of events is represented in Figure 4 below:

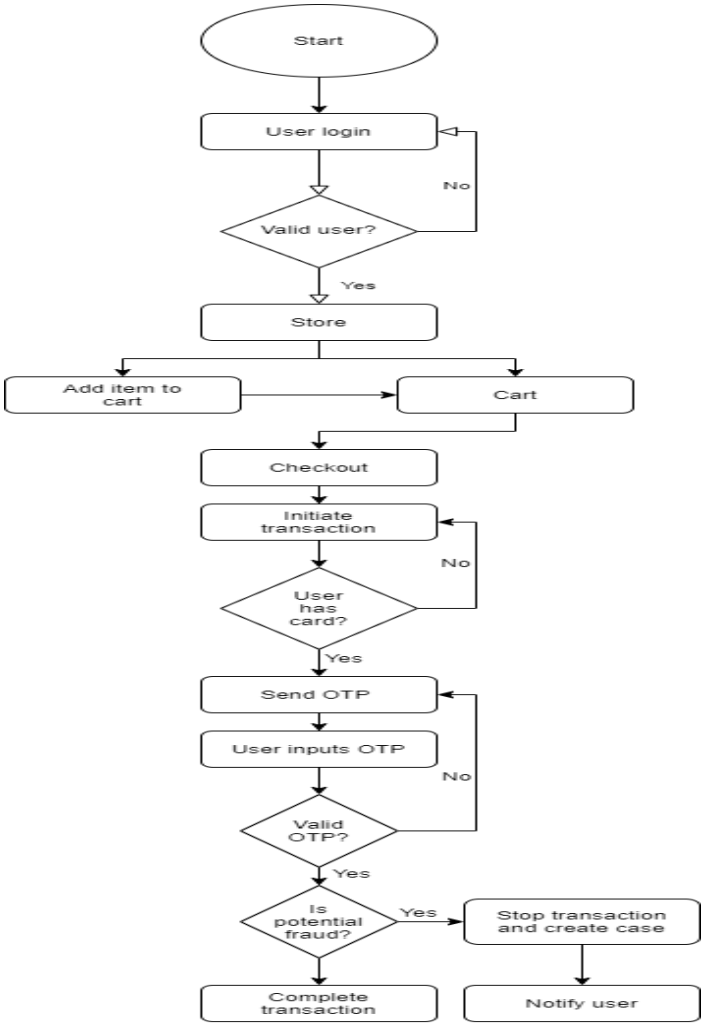


**Figure 4: Flowchart representing the workflow processes**

Authentication is first done by requesting and confirming the username and password after which the user initiates transactions with the credit card. The credit card about to be used is then validated by the system to ensure it belongs to the user before payment is approved. Payment validation is achieved by sending a secure code to the user to notify him of the pending transaction and the need to approve it by inputting the secure code. Lastly, a system check is done to prevent/detect any fraud attempt. In a case of no fraud, it completes the transaction otherwise it halts the transaction, creates a case and notifies the user. Figure 5 shows how data is fed through input devices such as the keyboard and is displayed on the monitor. Here, the input specification for the implementation of the fraud detection system used in an e-commerce site is shown below

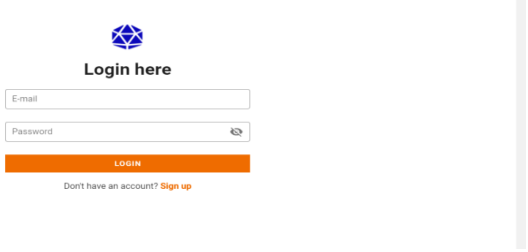


**Figure 5: User Login**

Figure 5 shows the user authentication process where the user is logged into the platform with a registered email address and password. The user email and password are then sent to the database to check for registered users with the exact credentials. If a matching user is found, the server returns the user's data along with a token for the user's logged in session. The user is then redirected to the home screen with lists of products as shown in figure 6

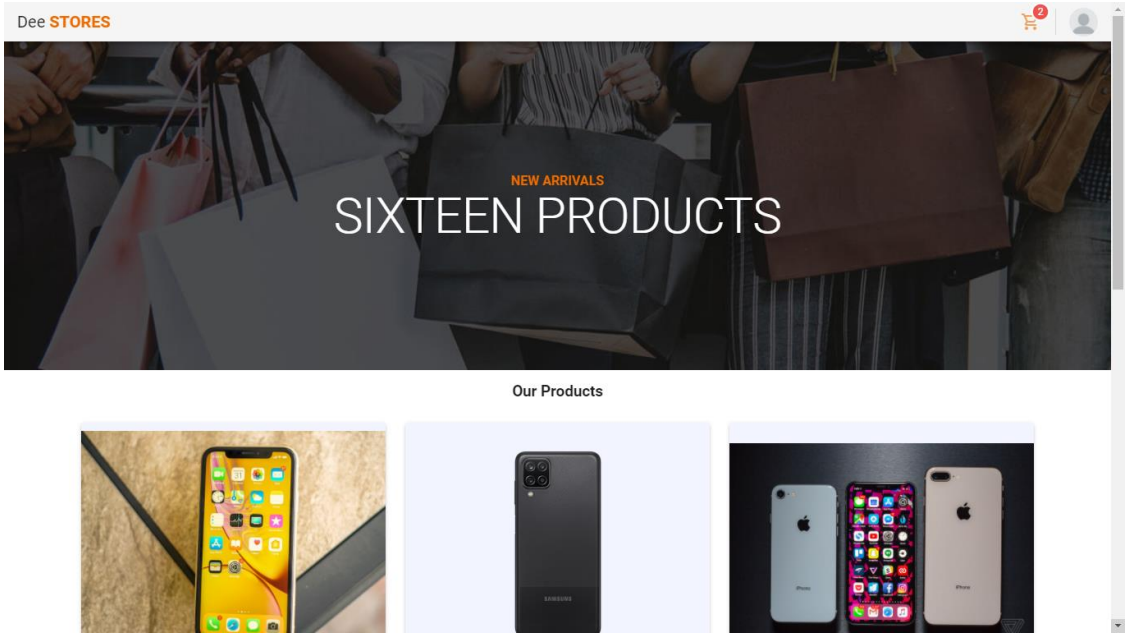


**Figure 6: Product's Home page**

Figure 6 shows the product's page where the user can easily add items to the cart with the click of a button. After all the selections have been made, the user then navigates to the cart page (Figure7) where the first stage of payment process begins.

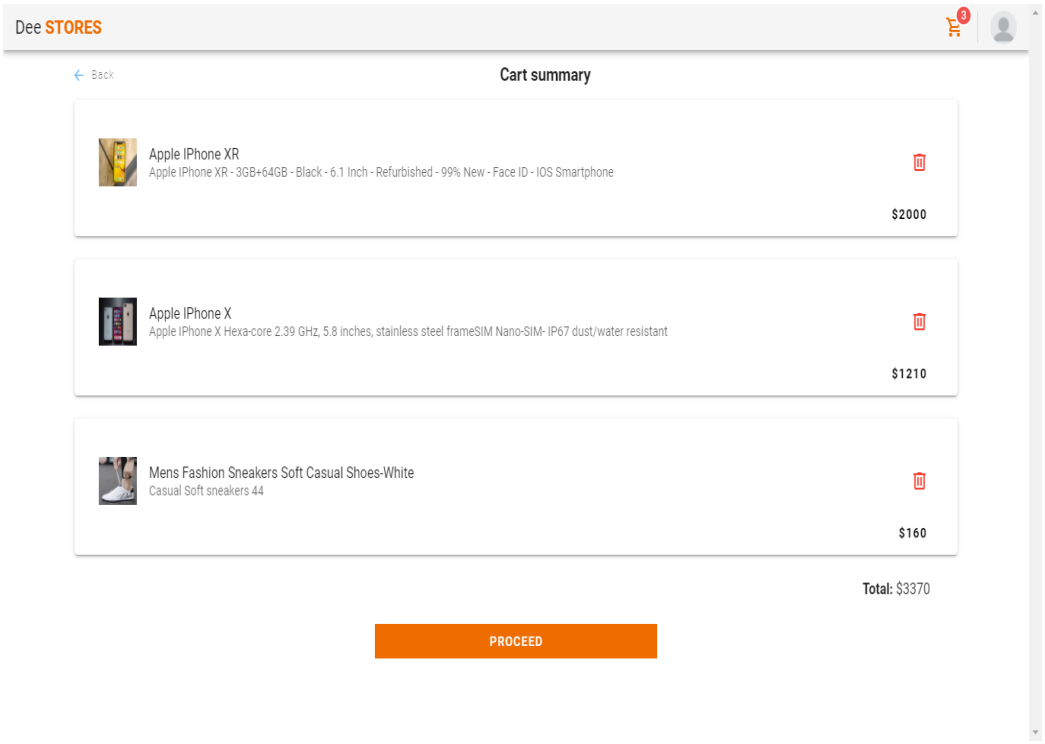


**Figure 7: Items in cart**

If the user decides to proceed to make a payment, the system will automatically send a secure OTP code to the user's email address as seen in Figure 8.

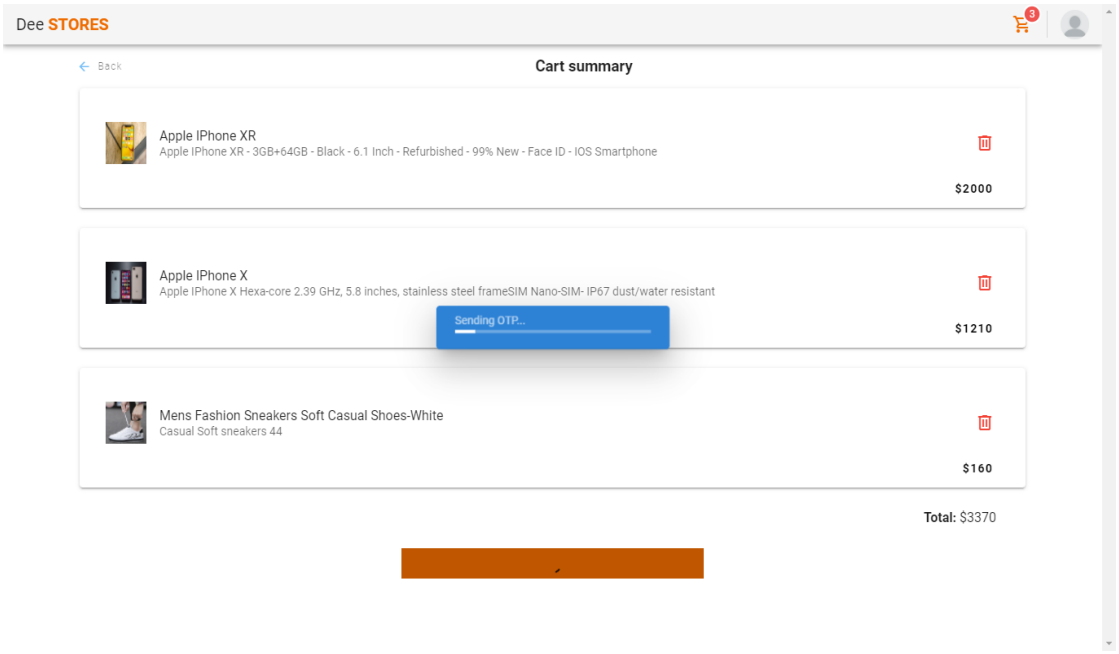


**Figure 8: Secure OTP code sent to user's email**

If a matching user is found, the server returns the user's data along with a token as shown in Figure 8 for the user's logged in session. The user is then redirected to the home screen with lists of products

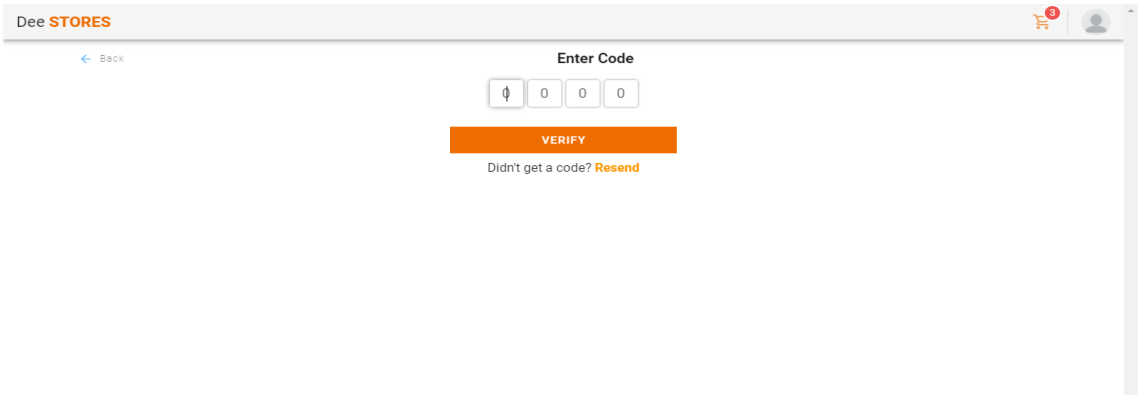


**Figure 9:User enters secure code**

Figure 9 shows the inputs provided by the system to the user for validating the OTP code sent. If the user enters the wrong code, the transaction process terminates. If the inputted code matches, the system then checks for a potential fraud case. Para venture, a case is found as in Figure 10, the transaction process is halted and a prompt email sent to the user.

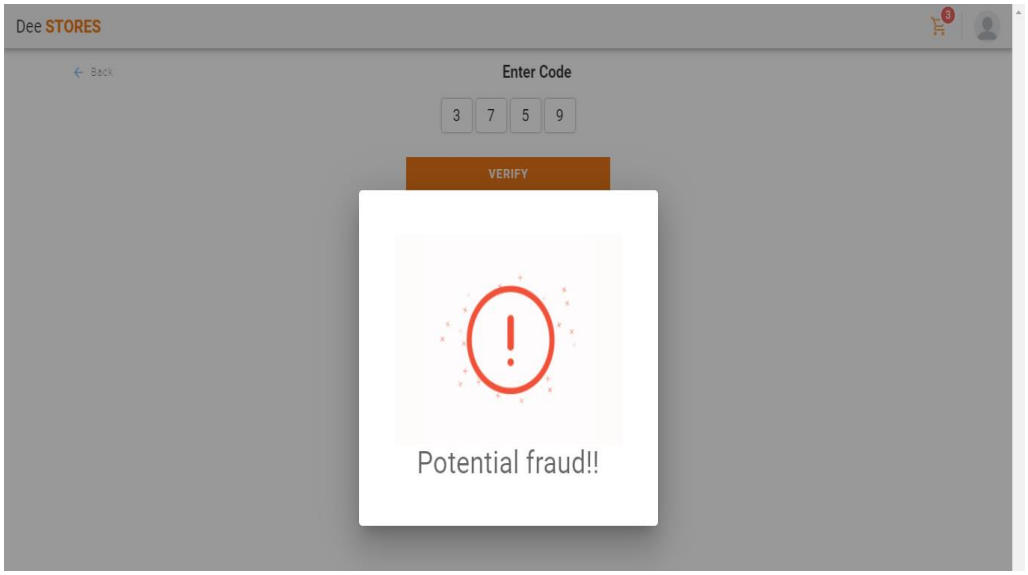


**Figure 10: System flags transaction as potential fraud**

## CONCLUSION

Recently, companies and individuals at large have been victims of online fraud. In this article, for the real-time detection of credit card fraud, a new approach was proposed for detecting online fraud according to research analysis and user requirements. The above enhanced model has shown a more promising result in the detection and prevention of credit card fraud as it recognizes when a transaction is fraudulent or not by employing multiple preventive fraud checks like User authentication, payment request validation (this confirms the user's authentication before transaction is initiated), payment confirmation(which further confirms the transaction by sending a secure code to the user notifying him of the pending transaction(s) and the need to approve it by inputting an already sent secure code), then finally, User System Check. The transaction is halted with a case created and sent to legal card owner notifying him of a fraudulent attempt. In a case of no fraud, the transaction process is completed. Hence, an improved and more secure transaction is ensured with reduced online transaction risks associated. However, this model can also be used in Banks, FinTech companies, Online Stores, Online Betting platforms and Mobile Wallets.

## REFERENCES

[1] Azeez, N. A., Idiakose, S. O., Onyema, C. J., & Vyver, C. Van Der. (2021). Cyberbullying Detection in Social Networks: Artificial Intelligence Approach. *Journal of Cyber Security* and Mobility, 10(4), 745–774. https://doi.org/10.13052/jcsm2245-1439.1046

[2] Andrew Bloomenthal (2021). Credit Card Definition

[3] Roman Chuprina (2021). Credit Card Fraud Detection Case Study: Improving Safety and Customer Satisfaction. https://spd.group/machine-learning/credit-card-fraud-detection-case-study/

[4] Khyati Chaudhary, Jyoti Yadav & Bhawna Mallick (2012), A review of Fraud Detection Techniques: Credit Card, International Journal of Computer Applications (0975 – 8887) Volume 45– No.1, May 2012

[5] Lucas, Y., Portier, P. E., Laporte, L., Calabretto, S., HeGuelton, L., Oblé, F., & Granitzer, M. (2019). Dataset shift quantification for credit card fraud detection. *Artificial Intelligence and Knowledge Engineering*.

[6] Yashvi Jain, Namrata Tiwari, Shripriya Dubey, Sarika Jain (2019), A Comparative Analysis of Various Credit Card Fraud Detection Techniques, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7 Issue-5S2, January 2019.

[7] Pushpalatha, B & Joseph, C. W. (2017) "Credit Card Fraud Detection Based on the Transaction by Using Data Mining Techniques," Vol. 5, No. 2, pp. 1785-1793, 2017

[8] Pouramirarsalani, A, Khalilian, M &Nikravanshalmani, A (2017). "Fraud Detection in Ebanking by using Hybrid Feature election and Evolutionary algorithms," International Journal of Computer and Network Security, Vol. 17, No. 8, pp. 271-279, 2017.

[9] Lepoivre, M. R., Avanzini, C. O., Guillaune Bignon, Legendre, L.& Piwele, A. K. (2016) "Credit Card Fraud Detection with Unsupervised Algorithms," Journal of Advances in Information, Vol. 7, No. 1, pp. 34-38, 2016.

[10] Navanshu Khare & Saad Yunus Sait (2018). Credit Card Fraud Detection Using Machine Learning Models and Collating Machine Learning Models, International Journal of Pure and Applied Mathematics Volume 118 No. 20 2018, 825-838 ISSN: 1314-3395

[11] Raj, S. B. E., & Portia, A. A. (2011). Analysis on credit card fraud detection methods. In Computer, Communication and Electrical Technology (ICCCET), 2011 International Conference on (pp. 152- 156).IEEE.

[12] Hyder John, Sameena Naaz (2019). Credit Card Fraud Detection using Local Outlier Factor and Isolation Forest, International Journal of Computer Sciences and Engineering. Vol.-7, Issue-4, April 2019; E-ISSN: 2347-2693

[13] Subramanian, R. R., Ramar, R. (2019) "Design of Offline and Online Writer Inference Technique", International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 2S2, Dec. 2019, ISSN: 2278-3075

[14] Richard J. Bolton & David J. Hand (2001). Unsupervised profiling methods for fraud detection. Credit Scoring and Credit Control VII, pages 235–255.