

Credit Card Fraud Detection Using Deep Learning

**Prof. S.C. Deshmukh¹, Mr. Rokade Dhananjay², Ms. Dumbre Rutuja³,
Ms. Patil Apeksha⁴, Ms. Pandharpote Anushka⁵**

Professor, Department of Information Technology¹

Student, Department of Information Technology²

Amrutvahini College of Engineering, College, Maharashtra, India³

Abstract: Credit cards offer an effective and convenient method for doing online transactions. Credit card abuse is becoming more likely as a result of increased credit card use. Both the owners of credit cards and financial institutions suffer large financial losses as a result of credit card theft. The major goal of this study is to identify such frauds, which may be done by looking at factors including the availability of public data, data with severe class imbalances, changes in the form of fraud, and high rates of false alarm. Applying the LSTM and RNN model's Deep Learning algorithm and adding extra layers for feature extraction and the categorization of credit card transactions as fraudulent or not is the major goal. The primary objective is to identify fraudulent credit card transactions with the use of deep learning algorithms. The most important elements from the CCF transaction dataset are ranked using feature selection techniques to aid with class label predictions. The credit card fraud detection dataset is utilised to extract the features and classification for the deep learning model by adding a number of additional layers. Apply alternative layer architectures to analyse the performance of both models. Text processing and the baseline model are linked to DL approaches. These techniques outperform the conventional algorithm for the identification of credit cards.

Keywords: DL:-Deep Learning, RNN:-*Recurrent Neural Network*, SMOTE:-*Synthetic Minority Oversampling*, IP:- *Internet Protocol*.

I. INTRODUCTION

Nowadays The use of credit cards has significantly expanded globally, and many increasingly believe in being cashless and are entirely reliant on online purchases. The credit card has facilitated and widened access to digital transactions. Each year, fraudulent credit card transactions result in enormous financial losses. Fraud has existed since the beginning of time and may take countless different forms. According to the 2017 PwC global economic crime study, there were 48 organisations affected by economic crime. Therefore, it is imperative to resolve the issue of credit card fraud detection. Additionally, the development of new technology offers additional possibilities for crooks to con people. In today's culture, the usage of credit cards is widespread, and in recent years, credit card fraud has continued to rise. Huge financial losses have been caused by fraud and have affected credit users personally as well as businesses and banks. Fraud may result in non-financial damages by harming a business's brand and image. For instance, if a cardholder experiences fraud with a certain company, he may lose faith in them and select a rival. Fraud detection is the practise of keeping track of a cardholder's transaction patterns to determine if an incoming transaction is legitimate and authorised or not. If not, the transaction will be flagged as fraudulent. We developed a protocol or model for this proposed project to identify fraud activities in credit card transactions. The majority of the key characteristics necessary to distinguish between genuine and fraudulent transactions may be provided by this system. It is harder to trace the modelling and pattern of fraudulent transactions as technology advances. With the development of artificial intelligence, machine learning, and other pertinent information technology sectors, it is now possible to automate this procedure and reduce the amount of labour-intensive work involved in identifying credit card fraud.

In today's culture, the usage of credit cards is widespread, and in recent years, credit card fraud has continued to rise. Huge financial losses have been caused by fraud and have affected credit users personally as well as businesses and banks. Fraud may result in non-financial damages by harming a business's brand and image. create and implement a Python-based system to anticipate credit card fraud using machine learning and deep learning models.

The creation of a technique or model to identify fraud in credit card transactions. The system is able to deliver the majority of the crucial characteristics needed to distinguish between genuine and fraudulent transactions. Tracking the modelling and pattern of fraudulent transactions is more challenging as technology advances. Seven chapters make up the study, which contains information on "Credit Card Fraud Detection Using Deep Learning."

II. LITERATURE SURVEY

The phrase "Fraud Detection Techniques for Credit Card Transactions,"

Modelling previous credit card transactions using information about cardholders who have been found to be fraudsters is an issue when trying to identify a credit score. The version is then employed to determine if recently initiated transactions are fraudulent or not. Our aim is to minimise fraudulent misclassification and detect all fraudulent transactions. One popular class model is the detection of credit card fraud. Using PCA-converted credit card transaction information, we concentrated on the examination and preparation of a number of anomaly detection techniques and record sets, including "neighbour outliers" and "forest zone isolation" procedures.

A study titled "Fraud Detection in Credit Card Transactions Using SVM and Random Forest Algorithms,"

All credit card-issuing institutions must work to lower the cost of putting in place an efficient fraud detection system. The fact that a credit card transaction does not require the card or the cardholder to be completed is one of the trickiest problems. As a result, the vendor is unable to confirm if the consumer making the purchase is an authorised cardholder. This method, which combines the random forest, decision tree, and support vector machine techniques, improves the accuracy of fraud detection. A classification method for monitoring the data set and enhancing the accuracy of the output data is the random forest algorithm. The effectiveness of the procedures is assessed based on their precision, sensitivity, and accuracy.

The study, "Fraud Detection in Credit Card Data Using Unsupervised Machine Learning Based Scheme,"

Each year, there are more fraudulent credit card data transactions. Researchers are experimenting with cutting-edge methods to identify and stop such scams in this direction.

However, there will always be a need for certain methods that can quickly and accurately identify these frauds. In this study, a method for identifying credit card fraud using unsupervised learning based on neural networks (NN) is proposed. The suggested strategy works better than the currently used Auto Encoder (AE), Local Outlier Factor (LOF), Isolation Forest (IF), and K-Means clustering algorithms. The proposed NN-based fraud detection approach has a 99.87% accuracy rate, while the existing AE, IF, LOF, and K Means methods have accuracy rates of 97%, 98%, 98%, and 99.75%, respectively.

"Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection,"

Online purchases are frequently made using credit cards. There have been reports of scams utilising credit cards in recent years. The detection and prevention of credit card theft are exceedingly challenging. Many issues in science and engineering are solved using the Artificial Intelligence (AI) approach known as Machine Learning (ML). In this study, machine learning algorithms are used to analyse a data set of credit card frauds, and the abilities of three machine learning algorithms to identify credit card fraud are compared. When compared to Decision Tree and XGBOOST methods, the accuracy of the Random Forest machine learning algorithm is the highest.

"Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms,"

The relevant literature presents many machines learning based approaches for credit card detection, such as Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression and XG Boost. However, due to low accuracy, there is still a need to apply state of the art deep learning algorithms to reduce fraud losses.

Title of paper: Cyber law, Name of the author: Taraq Hussain Sheik.

The report that was given states that "Cyber space is an ever-changing process and it should adapt to the demands of time. With the fast expansion of the internet industry, the author advises that the cyber laws already in effect be promptly updated to reflect the most recent offences committed. The wisdom of judgement should be discovered in the current reality of cyberspace crime, and the liberal interpretation of the law should be grounded in practical experience. The legal infrastructure that supports the internet should be in place, and it should be up to date. The adoption of cyber

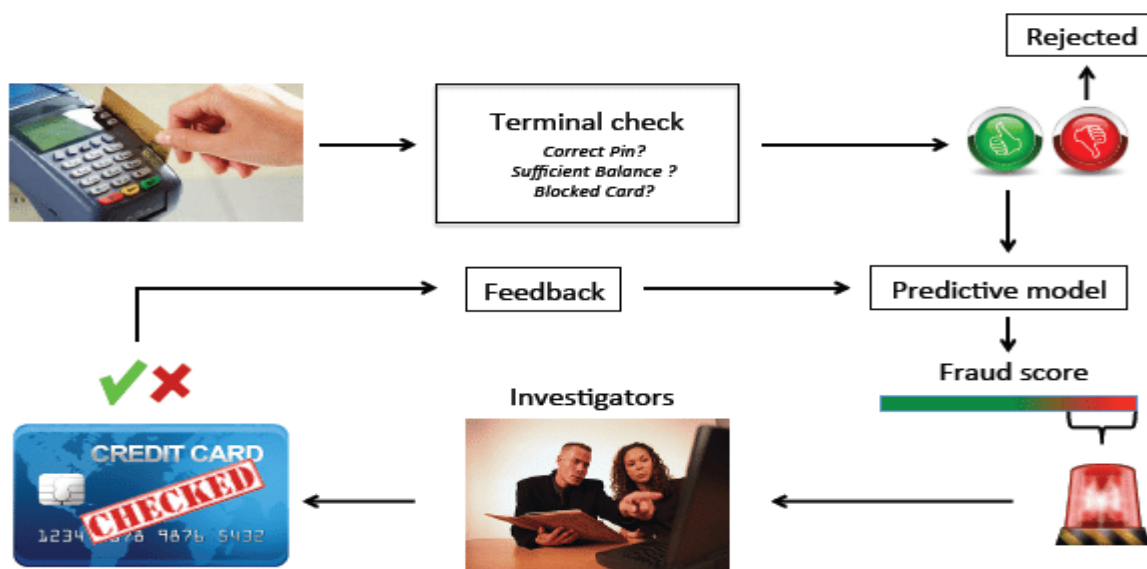
legislation is necessary in light of the existing infrastructure, but it has not been applied to e-commerce or to the other major aspect of the internet's future, and it should be updated in light of the industry's rapid expansion.

Title of paper: Legal Implication of the Cyber Crimes. Name of the author: Nikita Barman

The information act, or the IT industry, separates cybercrimes into cyber contraventions and cybercrimes, claims the study. Intelligent individuals and advanced systems are needed to make the cyberspace network more safe. According to social standards, the hazy and confusing laws should be generously construed. The tools that can be used to hack, spam, and breach secure networks should be tracked down and rendered worthless. Limited versions of the knowledge and information about the cyberspace network should be made available.

III. METHODOLOGY

We developed a protocol or model for this proposed project to identify fraud activities in credit card transactions. The majority of the key characteristics necessary to distinguish between genuine and fraudulent transactions may be provided by this system. It is harder to trace the modelling and pattern of fraudulent transactions as technology advances. With the development of artificial intelligence, machine learning, and other pertinent information technology sectors, it is now possible to automate this procedure and reduce the amount of labor-intensive work involved in identifying credit card fraud.



Step1:Preprocessing & Dataset: The ready dataset is used to detect credit card fraud.a method that is utilised to format the raw data into something meaningful and effective.

Step2:The necessary characteristics are retrieved in order to train the model.

Step3: Model and Classifier: The model is classed after being trained on data.

Step 4:Test Data: Following model construction, testing data confirms that the model is capable of making precise predictions. Testing data should not be labelled if training and validation data include labels to track the model's performance metrics. To ensure that the ML algorithm was successfully trained, test data offers a last-minute, real-world verification of an unknown dataset.

IV. PROPOSED WORK

1. Logistic Regression Algorithm: The logit model, often known as the logistic regression method, is a popular model for dichotomous output variables that was developed for the prediction of illness categorization. Assume there are p

input variables, each of whose values is represented by an x_1, x_2, \dots, x_p . Let z represent the likelihood that an event will occur and $1-z$ represent the likelihood that it won't. Given by is the logistic regression model.

$$\log\left(\frac{z}{1-z}\right) = \text{logit}(z) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

or can be written by

$$Z = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}$$

where β_0 is the intercept and $\beta_0, \beta_1, \dots, \beta_p$ are the regression coefficients.

2. Recurrent Neural Network(RNN): Recurrent neural networks (RNNs) are a subclass of neural networks that may be applied to a variety of sequence modelling applications, including voice recognition, machine translation, language modelling, and more. RNNs, in contrast to conventional feedforward neural networks, feature a feedback loop that enables them to process sequential input by transmitting data from one time step to the following. The network can detect both short-term and long-term dependencies in the sequence thanks to this feedback loop. Each of the cells that make up an RNN has an input, a hidden state, and an output. Each cell's secret state is supplied back into the network as input for the following cell in the chain. Each time step's input and previous hidden state are merged to create a new hidden state, which is then utilised to create an output. RNNs come in a variety of forms, such as the fundamental RNN, LSTM, and GRU (Gated Recurrent Unit). The issue of vanishing gradients, which may happen in conventional RNNs when the network tries to propagate mistakes back through several time steps, is addressed by the widely used RNN variants LSTM and GRU. RNNs have demonstrated their efficacy for a variety of sequence modelling problems, and many applications in voice recognition and natural language processing now turn on RNN variants.

3. SMOTE Algorithm:

SMOTE is an algorithm that adds artificial data points to the actual data points to accomplish data augmentation. SMOTE can be viewed as an improved form of oversampling or as a particular data augmentation procedure. With SMOTE, you avoid producing duplicate data points and instead produce synthetic data points that are marginally different from the original data points.

The **SMOTE algorithm** works as follows:

You choose a random representative from the minority group.

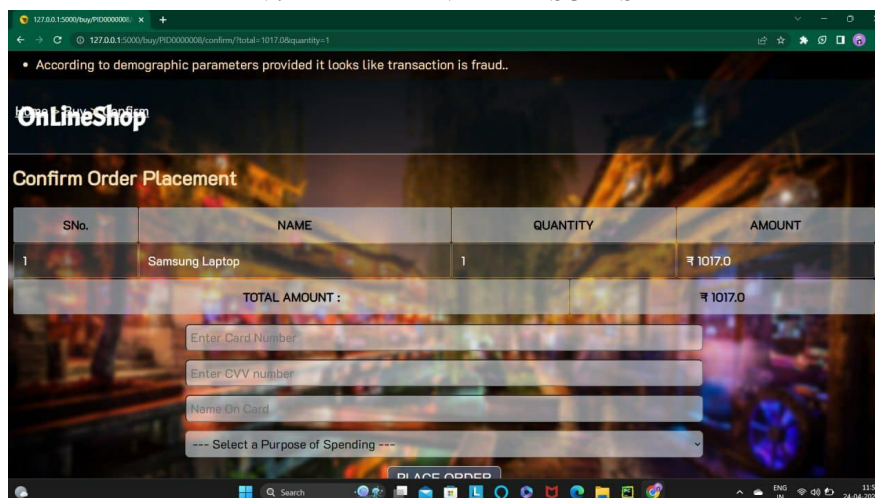
Find the k closest neighbours for the observations in this sample.

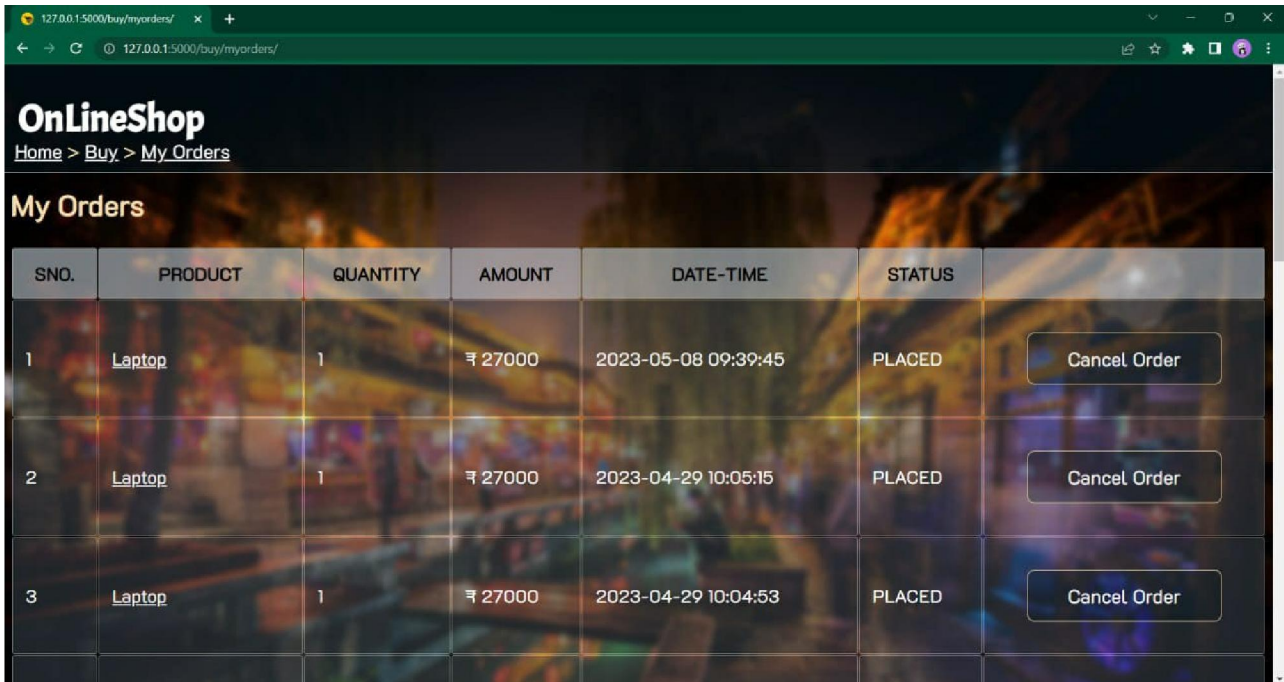
The vector between the current data point and one of those neighbours will then be determined using that neighbour.

The vector is multiplied by a chance number between 0 and 1.

You combine this with the existing data point to get the synthetic data point.

V. EXPERIMENTAL RESULTS

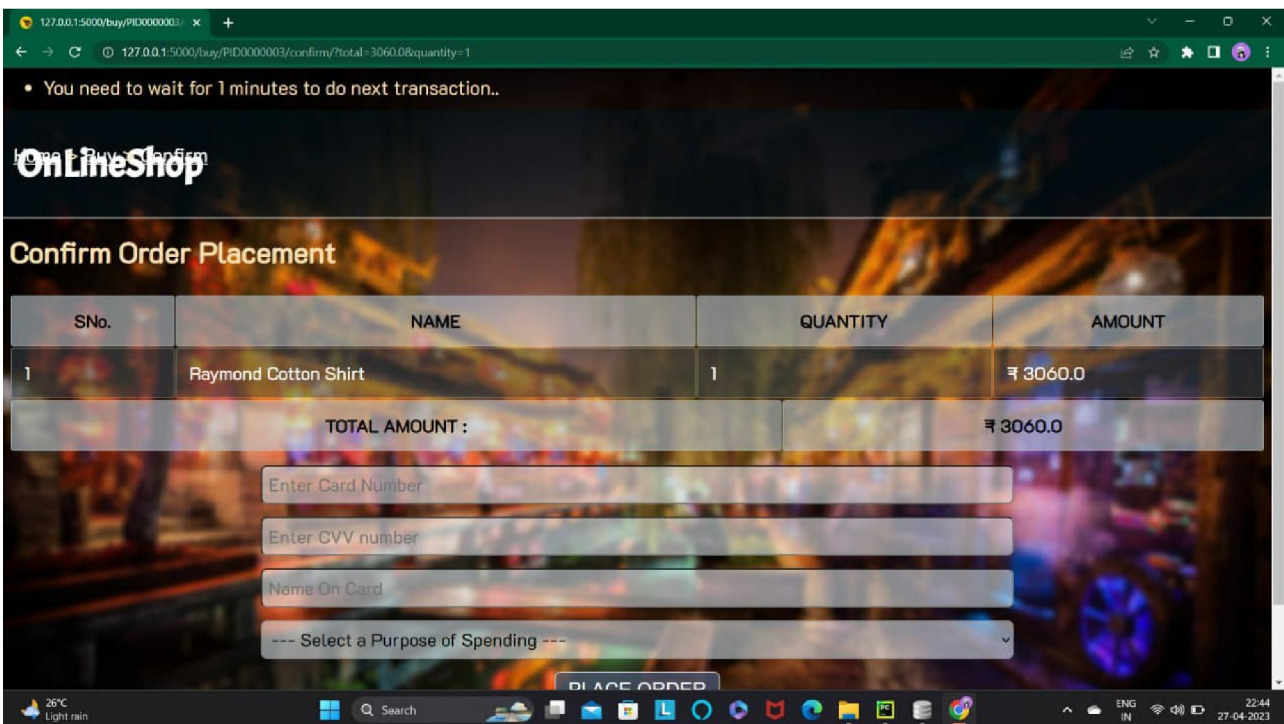




OnLineShop
Home > Buy > My Orders

My Orders

SNO.	PRODUCT	QUANTITY	AMOUNT	DATE-TIME	STATUS	
1	Laptop	1	₹ 27000	2023-05-08 09:39:45	PLACED	Cancel Order
2	Laptop	1	₹ 27000	2023-04-29 10:05:15	PLACED	Cancel Order
3	Laptop	1	₹ 27000	2023-04-29 10:04:53	PLACED	Cancel Order



OnLineShop
Home > Buy > Confirm

Confirm Order Placement

You need to wait for 1 minutes to do next transaction..

SNo.	NAME	QUANTITY	AMOUNT
1	Raymond Cotton Shirt	1	₹ 3060.0
TOTAL AMOUNT :			₹ 3060.0

Enter Card Number

Enter CVV number

Name On Card

--- Select a Purpose of Spending ---

PLACE ORDER

VI. CONCLUSION

Identification of credit card fraud is the system's primary goal. This system suggests a technique that uses stronger algorithms to identify fraud in addition to basic construction with associated features. The suggested application would raise people's knowledge of and ability to recognise phoney or counterfeit currency that is in circulation among those who frequently conduct cash transactions in their different enterprises.

REFERENCES

- [1] Y. Singh, K. Singh and V. Singh Chauhan, "Fraud Detection Techniques for Credit Card Transactions," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), 2022, pp. 821-824, doi: 10.1109/ICIEM54221.2022.9853183.
- [2] S. K. Saddam Hussain, E. Sai Charan Reddy, K. G. Akshay and T. Akanksha, "Fraud Detection in Credit Card Transactions Using SVM and Random Forest Algorithms," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2021, pp. 1013-1017, doi: 10.1109/I-SMAC52330.2021.9640631.
- [3] A. K. Rai and R. K. Dwivedi, "Fraud Detection in Credit Card Data using Unsupervised Machine Learning Based Scheme," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 421-426, doi: 10.1109/ICESC48915.2020.9155615.
- [4] V. Jain, M. Agrawal and A. Kumar, "Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 86-88, doi: 10.1109/ICRITO48877.2020.9197762.
- [5] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," in IEEE Access, vol. 10, pp. 39700-39715, 2022, doi: 10.1109/ACCESS.2022.3166891.(BASE Paper)