International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

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

J.Chandra babu[1], J.Pujitha[2], K.Likitha[3], R.Lavanya[4], P.Munipavan[5]
Assistant professor, student, Department of CSE, AITS, Tirupati

Abstract
Credit cards offer an effective and user-friendly facility, so people may use them for online purchases. The goal of this research study is to identify and prevent credit card fraud, which can result from high-class data imbalance, data accessibility, changes in fraud nature, and high false alarm rates. To achieve this goal, several machine learning-based strategies, including the Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, and XG Boost, have been presented in prior studies. However, the low accuracy of these methods necessitates the use of modern deep learning algorithms to prevent fraud losses. The study focuses on conducting a thorough empirical analysis of a dataset to identify fraud. The dataset was first subjected to a machine learning technique, which almost increased the accuracy of fraud detection. Variations in the number of hidden layers, epochs, and the use of the most recent models have also been subjected to a thorough empirical analysis.
Keywords: Data imbalance, Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, LogisticRegression, and XG Boost deep learning algorithms, dataset, empirical analysis.

1. Introduction

The use of credit cards and digital payment methods is becoming increasingly common, with a growing need for effective cybersecurity measures to combat fraud. Credit card fraud is a major issue, with global losses estimated to reach $43 billion in the next five years. Unfortunately, credit card users are increasingly being victimized, with more than 44 percent reporting two or more fraudulent charges in 2022 compared to 35 percent in 2021. To combat credit card fraud, most fraud detection systems use artificial intelligence, Meta-learning, and pattern matching. Hackers are increasingly using end-to-end technology and exploiting human vulnerabilities to access secure information. This has led to a rapid increase in the risk of network information insecurity. To address these challenges, fraud detection models must be able to process enormous amounts of data quickly and accurately. The models used must be simple and fast enough to detect anomalies and classify them as fraudulent transactions as quickly as possible. Additionally, to protect the privacy of users, data dimensionality can be reduced. Overall, the main objective of credit card fraud detection is to prevent
Online fraud when using online financial transactions. By leveraging advanced technologies such as AI and machine learning, fraud analysts can more effectively detect and prevent fraudulent transactions, reducing losses for individuals and businesses alike.

2. Related Work

Machine learning and deep learning techniques have shown great promise in detecting credit card fraud by recognizing unusual transactions. The first step in implementing such techniques is to collect and sort data, which is then used to train the model to predict the probability of fraud. In a recent research study conducted by Aditi Singh and others, different machine learning algorithms were tested for their effectiveness in detecting credit card fraud. The study examined recent developments and applications in the area of machine learning-based credit card fraud detection. The authors found that the Cat boost algorithm performed the best, with an accuracy of 99.87 percent, in detecting credit

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

card fraud. This algorithm uses gradient boosting and can handle categorical variables well, making it a promising technique for credit card fraud detection.

1. Feature Selection Methods

With the massive quantity of credit card data available, it becomes essential to develop effective feature selection approaches to reduce computation complexity and improve classification accuracy. G. K. Arun proposed a novel feature selection approach with deep learning to address this issue. The approach is called the binary emperor penguin optimization (BEPO) based feature selection with optimal gated recurrent unit (OGRU), or the BEPO-OGRU technique for credit card fraud detection. The primary objective of the BEPO-OGRU technique is to detect and classify possible credit card frauds using deep learning and feature selection. The BEPO-OGRU technique involves selecting the most relevant features from the credit card data and then using the OGRU model for fraud detection. The study conducted a wide range of simulations to evaluate the effectiveness of the BEPO-OGRU technique. The results showed that the proposed technique outperformed other existing methods with an improved accuracy of 94.78% and 94.16% on the German Credit dataset and Credit Card Fraud Detection dataset, respectively.

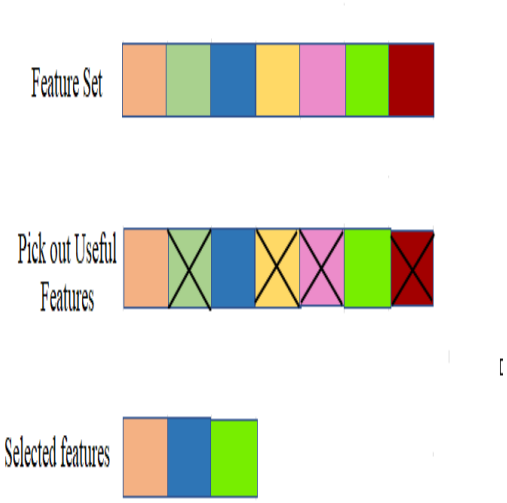

Fig 2.1.1: Feature Selection

2. Decision Tree and Random Forest Algorithms

Due to this the industry and customers who are using credit cards are facing a huge loss. There is a deficiency of investigation lessons on examining practical credit card figures in arrears to privacy issues. In this regard M R Dileep implemented two algorithms that are used viz Fraud Detection in credit card using Decision Tree and Fraud Detection using Random Forest. The efficiency of the model can be decided by using some public data as sample. The significance of the methods used in the paper is the first method constructs a tree against the activities performed by the user and using this tree scams will be suspected. In the second method a user activity based forest will have constructed and using this forest an attempt will be made in identifying the suspect. The investigational outcomes absolutely show that the mainstream elective technique attains decent precision degrees in sensing scam circumstances in credit cards.

3. Comparative Analysis of Supervised Classifiers

Credit Card Fraud Detection System detects given a transaction whether it is fraud or not. Shilpa. C presented comparative analysis for credit card fraud detection. That paper mainly focuses on how frauds can be analyzed using Machine Learning approaches. The main objective of this paper focused

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

on methods of handling imbalance in dataset of Credit Card transactions. Later, Credit Card transactions dataset is used on XGBoost and Random Forest Classifier.

4. Convolutional Neural Network

Credit Card Fraud (CCF) strategies continue to evolve with the business transformation, causing customers as well as the financial institutions to lose billions of dollars annually. Hence, Muhammad Liman Gambo implemented a Convolutional Neural Network (CNN) model for credit card fraud detection is proposed in using Adaptive Synthetic (ADASYN) sampling technique to address the imbalance dataset. The proposed model has achieved 0.9982, 0.9965, and 0.9999, accuracy, precision, and recall, respectively compared to other existing studies.

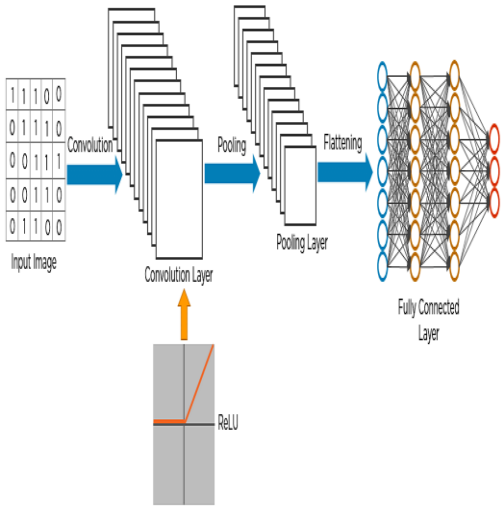

Fig 2.4.1 : Convolutional Neural Network

5. Evaluation Metrics

Evaluation metrics are used to measure the performance of a model or algorithm in terms of its accuracy, precision, recall, F1-score, and other relevant measures. It is crucial to evaluate a model to ensure that it is performing well on the given task and to identify areas for improvement.

Confusion Matrix

A confusion matrix is a table used to evaluate the performance of a classification model by comparing the predicted class labels with the actual class labels of a set of test data. The predicted values are usually shown in columns, while the actual values are shown in rows.

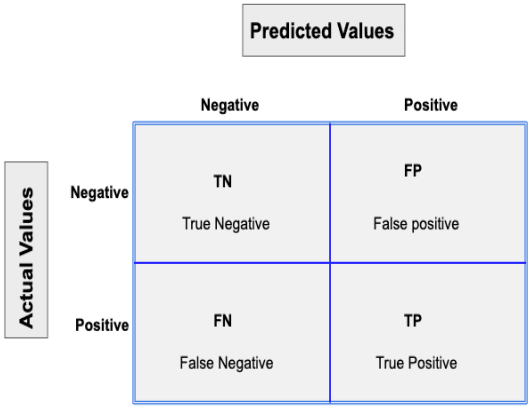

Fig 2.5.1 : Confusion Matrix

Each prediction can be one of the four outcomes, based on how it matches up to the actual value:

DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

- **True Positive (TP):** Predicted True and True in reality.
- **True Negative (TN):** Predicted False and False in reality.
- **False Positive (FP):** Predicted True and False in reality.
- **False Negative (FN):** Predicted False and True in reality.

There is also a list of rates that are often computed from a confusion matrix for a binary classifier:

- **Accuracy:** It is a commonly used evaluation metric for binary classifiers when the classes are roughly equal in size. It measures the proportion of true positives and true negatives among all the samples. It is calculated as follows:

Accuracy = (TP+TN)/Total Predictions

Accuracy can be misleading if the classes are imbalanced, meaning that one class has significantly more samples than the other. In such cases, a classifier that always predicts the majority class will have a high accuracy but may not be useful in practical applications. In such cases, other evaluation metrics such as precision, recall, and F1 score may be more appropriate.

- **Precision:** The precision metric is used to overcome the limitation of Accuracy. The precision determines the proportion of positive prediction that was actually correct. It can be calculated as the True Positive or predictions that are actually true to the total positive predictions

Precision = TP/(TP+FP)

- **Recall or Sensitivity:** It is also similar to the Precision metric; however, it aims to calculate the proportion of actual positive that was identified incorrectly.

Recall = TP/(TP+FN)

- **F1- Scores**: F1 Score is a metric to evaluate a binary classification model on the basis of predictions that are made for the positive class. F1 Score is useful in cases where we have imbalanced classes, where one class is much more frequent than the other. F1 Score takes into account both precision and recall, which are important for evaluating the performance of a binary classification model on imbalanced data.

F1 Score = (2*(precision*recall)) / (precision+recall)

3. DESIGN

The module simply means the software components that are been created by dividing the software. The software is divided into various components that work together to form a single functioning item but sometimes they can perform as a complete function if not connected with each other. In this proposed system, there are two modules. They are:

1. Service Provider
2. Remote User

1. Service Provider

The Functionalities of this Modules are as Follows:

1. Login
2. Browse & Train & Test Credit card Data Sets
3. View trained & Tested Data Sets Accuracy In Bar Chart
4. View trained & Tested Data Sets accuracy results
5. View Prediction of credit card fraud detection
6. View Prediction of credit card fraud detection
7. Download Prediction data sets
8. View Credit card fraud detection ratio results
9. View all Remote Users
10. Logout

2. Remote User

The Functionalities of this Modules are as Follows:

1. Register

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

2. Login
3. Predict credit card fraud detection type
4. View Your Profile
5. Logout

3. Activity Diagram for Service Provider

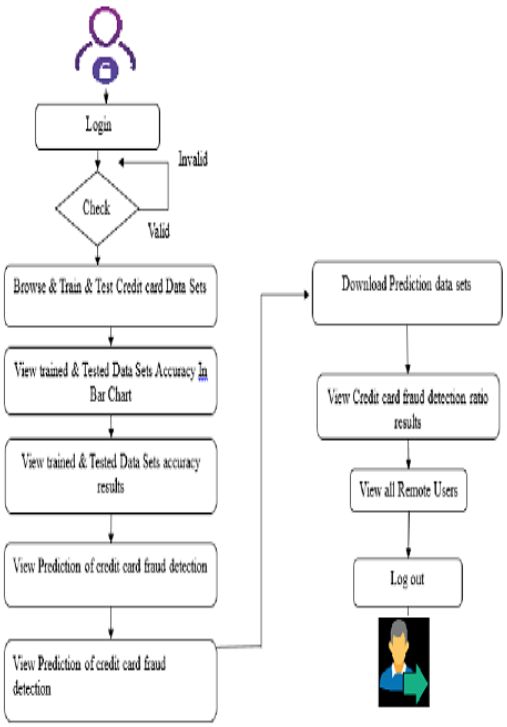

Fig 3.3.1: Activity Diagram for Service Provider

4. Activity Diagram for Remote User

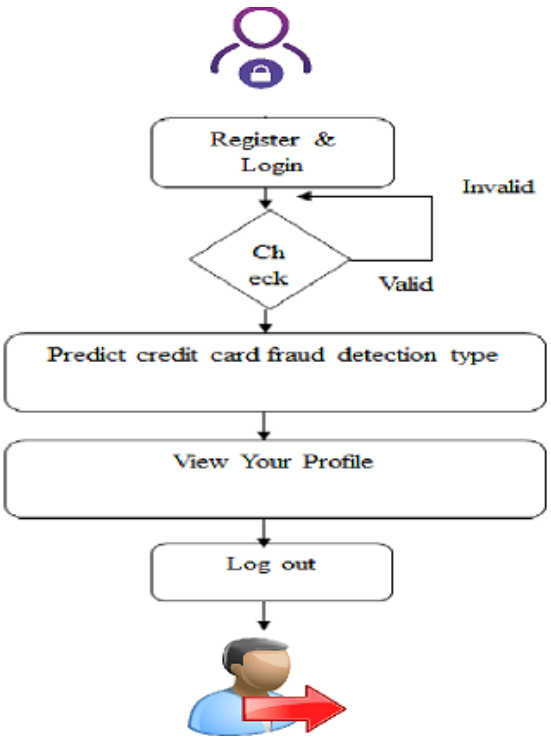

Fig 3.4.1: Activity Diagram for Remote User

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

4. EXPERIMENTAL RESULT

The use of feature selection helps to identify the most important features for predicting fraud, while deep learning models extract complex features and improve classification accuracy. Comparative analysis is performed to evaluate the effectiveness of the proposed approach against existing approaches, and performance evaluation metrics such as accuracy, precision, and recall are used to assess the quality of the classifiers. The experiments are conducted on the latest credit card dataset to evaluate the performance of the models.

1. The Accuracy of Deep Learning Algorithms

The training and validation accuracy of baseline CNN and proposed CNN algorithms are shown in the Table below.

Measurements	Epoch Size 35	Epoch Size 14
Total Fraudulent Transactions	107	107
Accuracy	0.999	0.998
Precision	0.932	0.568
Recall	0.775	0.813
TP	83	87
FP	6	66
TN	56849	56789
FN	24	20

Fig 4.1.1: Deep Learning Accuracy

2. Accuracy and F1 Score of ML algorithms

We construct five different types of classification models during this stage. We could solve categorization issues with a variety of alternative models. The models. The below table represents the accuracy and f1 score of the ML algorithms.

s.no	Algorithm Name	F1 score (%)	Accuracy (%)
1	Logistic Regression(LR)	73.6	99.9
2	SVM Algorithm(SVM)	77.7	99.9
3	Random Forest Tress (RF)	77.3	99.92
4	XG Boost (XG)	84.5	99.93
5	Decision Tree (DT)	81.1	99.94

Fig 4.2.1: F1 score and Accuracy of ML models

3. Machine Learning Algorithms Comparative Analysis

The below graph represents the comparative analysis of the f1 score and accuracy of Machine Learning algorithms. The below graph represents the comparative analysis of the f1 score and accuracy of Machine Learning algorithms.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Special Issue: 1 Volume No.7 April – 2023
DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

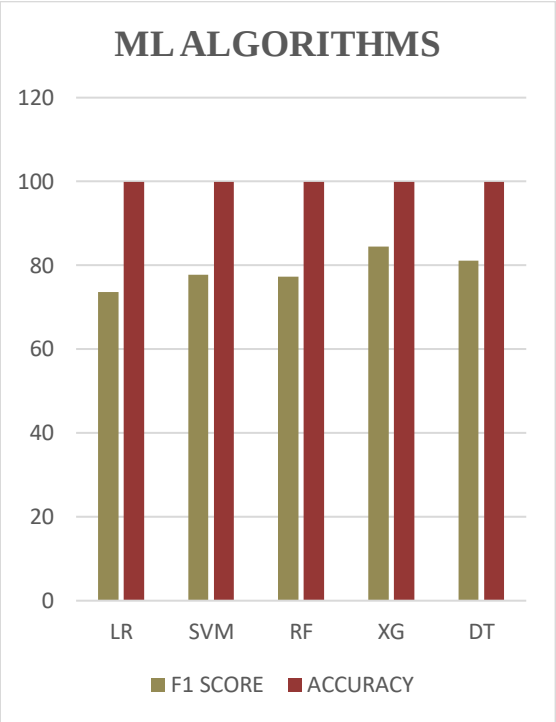

Fig 4.3.1: Graph of comparison of ML Algorithms

5. Conclusion:

Machine learning and deep learning models are better for fraud detection models. They can recognise thousands of patterns from large datasets. Those models offers an insight into how users behave by understanding their app usage, payments, and transactions methods. Some of the benefits of fraud detection using ML are Faster detection, Higher Accuracy, improved efficiency with data.

6. References:

[1]A. Singh, A. Singh, A. Aggarwal and A. Chauhan, "Design and Implementation of Different Machine Learning Algorithms for Credit Card Fraud

[2]G. K. Arun and P. Rajesh, "Design of Metaheuristic Feature Selection with Deep Learning Based Credit Card Fraud Detection Model," *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, Coimbatore, India, 2022, pp. 191-197.

[3]M. R. Dileep, A.V. Navaneeth and M. Abhishek, "A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms," *2021 Third Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV),* Tirunelveli, India, 2021, pp. 1025-1028.

[4]S. C and A. H. Shanthakumara, "A Comparative Analysis of Supervised Classifiers for DetectingCredit Card Frauds," *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2022, pp.1-6.

[5] M. L. Gambo, A. Zainal and M. N. Kassim, "A Convolutional Neural Network Model for Credit Card Fraud Detection," *2022International Conference on Data Science and Its Applications (ICoDSA),* Bandung, Indonesia, 2022, pp. 198-202.

[6]S. Mittal and S. Tyagi, "Performance Evaluation of Machine Learning Algorithms for Credit Card Fraud Detection," *20199th International Conference on Cloud Computing, Data Science &Engineering (Confluence)*, Noida, India, 2019, pp. 320-324.

DOI:10.46647/ijetms.2023.v07si01.070 ISSN: 2581-4621

[7] T. Jemima Jebaseeli, R.Venkatesan and K. Ramalakshmi, "Fraud detection for credit card transactions using random forest algorithm", *Adv. Intell. Syst. Comput*, vol. 1167, pp.189-197, 2021.
[8]C. Shilpa and S. A. H, "A Comparative Analysis of Supervised Classifiers for Detecting Credit Card Frauds", *2022 Int. Conf. Comput. Commun. Informatics*, no. 978, 2022.
[9] A. A. El Naby, E. El-Din Hemdan and A. El-Sayed,"Deep learning approach for credit card fraud detection", *ICEEM 2021 - 2nd IEEE Int. Conf. Electron. Eng*, pp. 0-4, 2021.