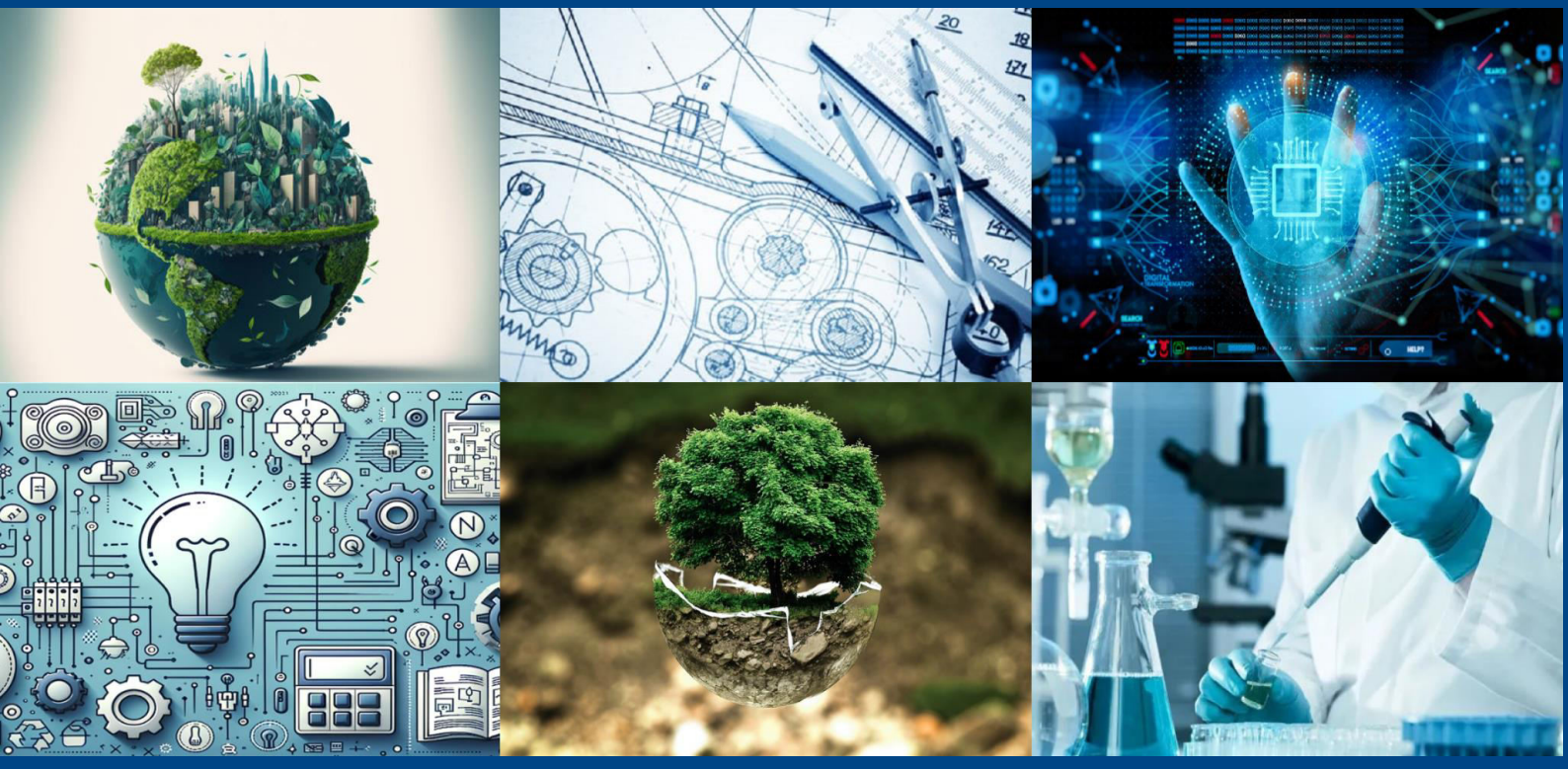
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 8, Issue 5, May 2025**

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805135

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Smart Credit Card Fraud Detection Using Artificial Intelligence

D. Prabhakaran, Dr. T. Geetha, V. Abirami,

Assistant professor, Department of Master of Computer Applications, Gnanamani College of Technology, Namakkal, Tamil Nadu, India

HOD, Department of Master of Computer Applications, Gnanamani college of Technology, Namakkal, Tamil Nadu India

PG Student, Department of Master of Computer Applications, Gnanamani College of Technology, Namakkal, Tamil Nadu, India

ABSTRACT: In our project, mainly focussed on credit card fraud detection for in real world. It is vital that credit card companies are able to identify fraudulent credit card transactions so that customers are not charged for items that they did not purchase. Such problems can be tackled with Data Science and its importance, along with Machine Learning, cannot be overstated. This project intends to illustrate the modelling of a data set using machine learning with Credit Card Fraud Detection. Collect the credit card datasets for trained dataset. Then will provide the user credit card queries for testing data set. After classification process of random forest algorithm using to the already analysing data set and user provide current dataset. Optimizing the accuracy of the result data. Then will apply the processing of some of the attributes provided can find affected fraud detection in viewing the graphical model visualization. The performance of the techniques is evaluated based on accuracy, sensitivity, and specificity, precision. The results indicate about the optimal accuracy for Random Forest is 98.6% respectively.

KEYWORDS: Credit Card, Pre-Processing, Random Forest Algorithm.

I. INTRODUCTION

'Fraud' in credit card transactions is unauthorized and unwanted usage of an account by someone other than the owner of that account. Necessary prevention measures can be taken to stop this abuse and the behaviour of such fraudulent practices can be studied to minimize it and protect against similar occurrences in the future. In other words, Credit Card Fraud can be defined as a case where a person uses someone else's credit card for personal reasons while the owner and the card issuing authorities are unaware of the fact that the card is being used. Fraud detection involves monitoring the activities of populations of users in order to estimate, perceive or avoid objectionable behaviour, which consist of fraud, intrusion, and defaulting. This is a very relevant problem that demands the attention of communities such as machine learning and data science where the solution to this problem can be automated. This problem is particularly challenging from the perspective of learning, as it is characterized by various factors such as class imbalance. The number of valid transactions far outnumber fraudulent ones. Also, the transaction patterns often change their statistical properties over the course of time. These are not the only challenges in the implementation of a real-world fraud detection system, however. In real world examples, the massive stream of payment requests is quickly scanned by automatic tools that determine which transactions to authorize. Machine learning algorithms are employed to analyse all the authorized transactions and report the suspicious ones. These reports are investigated by professionals who contact the cardholders to confirm if the transaction was genuine or fraudulent. The investigators provide a feedback to the automated system which is used to train and update the algorithm to eventually Improve the fraud-detection performance over time.

DATA SET

The datasets contains transactions made by credit cards. The dataset is highly unbalanced. It contains only numerical input variables which are the result of a PCA transformation. Cannot provide the original features and more background information about the data. Features V1, V2, ... V28 are the principal components obtained with PCA, the only features which have not been transformed with PCA are 'Time' and 'Amount'. Feature 'Time' contains the seconds

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

elapsed between each transaction and the first transaction in the dataset. The feature 'Amount' is the transaction Amount, this feature can be used for example-dependent cost-sensitive learning. Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise."

II. PRE-PROCESSING

Data preprocessing which mainly include data cleaning, integration, transformation and reduction, and obtains training sample data needed. It is a data mining technique that transforms raw data into an understandable format.
Steps in Data Preprocessing
1. Import libraries
2. Read data
3. Checking for missing values
4. Checking for categorical data
5. Standardize the data
6. PCA transformation
7. Data splitting

III. FEATURE EXTRACTION

Feature selection include reducing the computational costs, saving storage space, facilitating model selection procedures for accurate prediction, and interpreting complex dependencies between variables. The features that are well selected not only optimize the classification accuracy but also reduce the number of required data for achieving an optimum level of performance of the learning process. Feature selection methods usually include search strategy, assessment measure, stopping criterion, and validation of the results. ***Search strategy*** is a search method used for producing a subset of candidate features for assessment. ***An assessment measure*** is applied for evaluating the quality of the subset of candidate features. Validation is the study of validity of the selected features with the realworld datasets. Filter and Wrapper methods are the most important methods of feature selection.

IV. RANDOM FOREST ALGORITHM

It is the basic classifier and it establishes a large number of trees. Random forests is an effective prediction tool widely used in data mining. It constructs a series of classification trees which will be used to classify a new example. The idea used to create a classifier model is constructing multiple decision trees, each of which uses a subset of attributes randomly selected from the whole original set of attributes. Candidate split dimension a dimension along which a split may be made. Candidate split point one of the first m structure points to arrive in a leaf. Candidate split a combination of a candidate split dimension and a position along that dimension to split. These are formed by projecting each candidate split point into each candidate split dimension. Candidate children each candidate split in a leaf induces two candidate children for that leaf. These are also referred to as the left and right child of that split.

V. TRAINED DATA

The quality, variety, and quantity of your training data determine the success of your machine learning models. The form and content of the training data often referred to as labeled or human labeled data or ground truth dataset is designed for to train specific ML models with an end application in perspective.

VI. FRAUDULENT NOTIFICATION

Card issuer may be able to send credit card fraud alert notifications via text message to help you detect unauthorized charges quickly.

VII. RESULT AND DISCUSSION

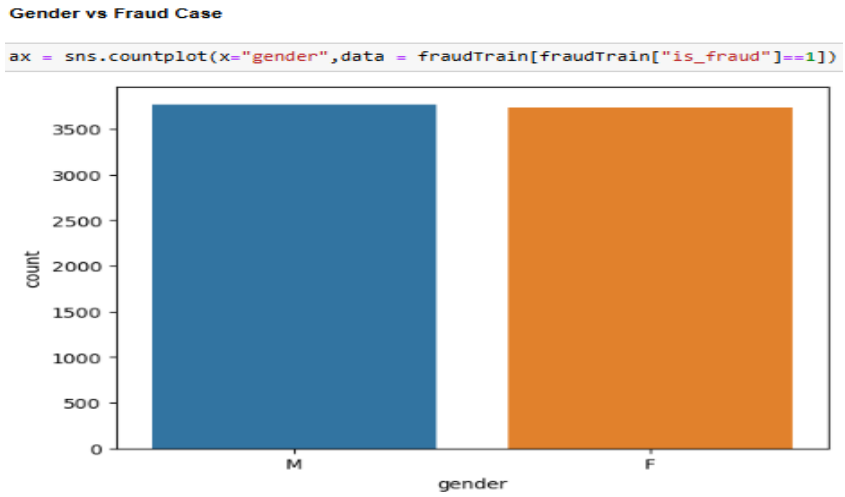
The AI-powered smart credit card fraud detection system demonstrated high accuracy in identifying fraudulent transactions by analyzing patterns in transaction data. Comparing this AI-based system to traditional rule-based fraud detection methods, it was evident that AI provided significant improvements in accuracy and efficiency.

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805135

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Graphical Representation of gender wise Fraud Detection

```
Score the X-train with Y-train is :  1.0
Score the X-test  with Y-test  is :  0.9958756133945393
Accuracy score  0.9958756133945393
```

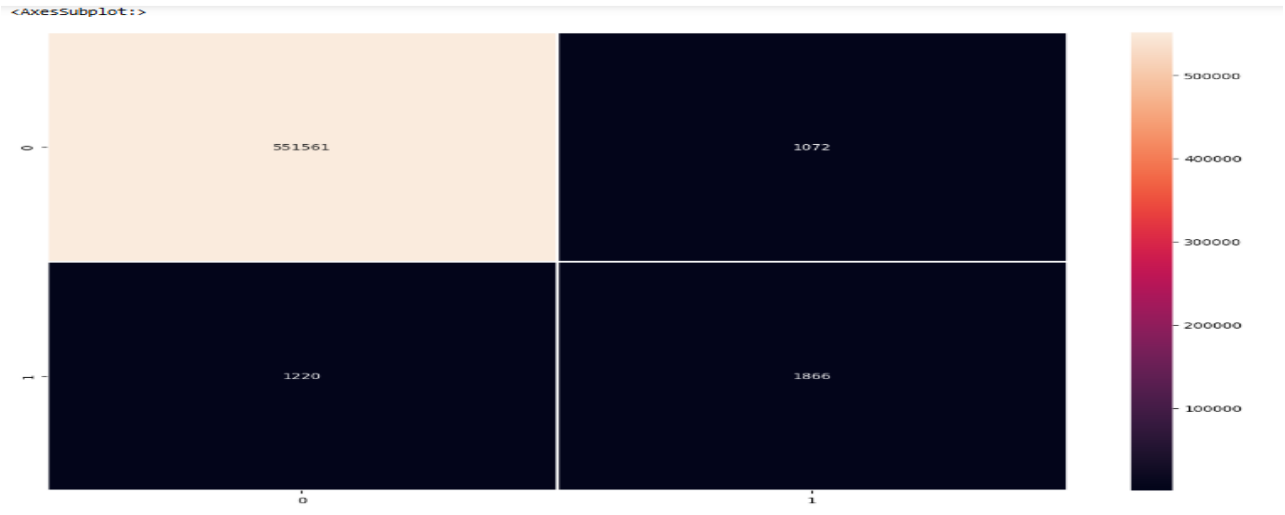
Accuracy Score

```
Classification report :
                precision    recall  f1-score   support

           1       0.60      0.64      0.62      2938
           0       1.00      1.00      1.00    552781

    accuracy                           1.00    555719
   macro avg       0.80      0.82      0.81    555719
weighted avg       1.00      1.00      1.00    555719
```

Classification report of credit card Fraud Detection



AxesSubplot

VIII. CONCLUSION

This Project has examined the performance of two kinds of random forest models. A real-life dataset on credit card transactions is used in our experiment. Although random forest obtains good results on small set data, there are still some problems such as imbalanced data. Our future work will focus on solving these problems. The algorithm of random forest itself should be improved. For example, the voting mechanism assumes that each of base classifiers has equal weight, but some of them may be more important than others. Therefore, we also try to make some improvement for this algorithm.

REFERENCES

[1] O. Adewumi and A. A. Akinyelu, "A survey of machine learning and nature-inspired based credit card fraud detect ion techniques," International Journal of System Assurance Engineering and Management, vol. 8, pp. 937–953, 2017J. Clerk Maxwell, A T reatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[2] A. Srivastava, A. Kundu, S. Sural, A. Majumdar, "Credit card fraud detect ion using hidden Markov model," IEEE Transact ions on Dependable and Secure Comput ing, vol. 5, no. 1, pp. 37–48, 2008K. Elissa, "Tit le of paper if known," unpublished.

[3] Bansal, J. C., Singh, P. K., Saraswat, M., Verma, A., Jadon, S. S., and Abraham, A. (2011). Inert ia weight strategies in part icle swarm optimization. In Nature and Biologically Inspired Computing (NaBIC), (Salamanca, Spain, October 19 - 21, 2011) IEEENaBIC'11,633--640.

[4] Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. Information Fusion. 28 (Mar. 2016), 45--59

[5] Bharill N., T iwari, A., and Malviya, A. (2016). Fuzzy Based Clustering Algorithms to Handle Big Data with Implementation on Apache Spark. In Proceedings of the IEEE 2nd Internat ional Conference on Big Data Comput ing Service and Applications, (Oxford, UK, March 29-April 01, 2016). IEEE BigDataService '16, 95104.

[6] Y. Sahin, S. Bulkan, and E. Duman, "A cost -sensitive decision tree approach for fraud detection," Expert Systems with Applications, vol. 40, no. 15, pp. 5916–5923, 2013.

[7] TheNilsonReport(October2016)[Online].Available:https://www.nilsonreport.com/upload/contentpromo/The_Nilson _Report_10-17-2016.pdf

[8] J. T. Quah, and M. Sriganesh, "Real-time credit card fraud detection using computational intelligence," Expert Systems with Applications, vol. 35, no. 4, pp. 1721–1732, 2008.

[9] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C., "Data mining for credit card fraud: A comparative study," Decision Support Systems, vol. 50, no. 3, pp. 602–613, 2011.

[10] S. Panigrahi, A. Kundu, S. Sural, and A. K Majumbar, "Use of Dempster-Shafer theory and Bayesian inferencing for fraud detection in communication networks", Lecture Notes in Computer Science, Spring Berlin/ Heidelberg, Vol. 4586,2007, p.446-460

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com