iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** II **Month of publication:** February 2024

DOI: https://doi.org/10.22214/ijraset.2024.58284

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

Credit Card Fraud Detection using ANN

S Atchaya
Bangalore Institute of Technology, K. R. Road, V. V. Pura, Bengaluru – 560004, Karnataka

Abstract: ***Frauds in credit card transactions are common today as most of us are using the credit card payment methods more frequently. This is due to the advancement of Technology and increase in online transaction resulting in frauds causing huge financial loss. Therefore, there is need for effective methods to reduce the loss. In addition, fraudsters find ways to steal the credit card information of the user by sending fake SMS and calls, also through masquerading attack, phishing attack and so on. This paper aims in using the multiple algorithms of Machine***
learning such as support vector machine (SVM), k-nearest neighbor (Knn) and artificial neural network (ANN) in predicting the occurrence of the fraud. Further, we conduct a differentiation of the accomplished supervised machine learning and deep learning techniques to differentiate between fraud and non-fraud transactions.
Keywords: Artificial neural network, credit card, fraud, k-nearest neighbor, machine learning and support vector machine.

I. INTRODUCTION

In recent years, as there is advancement of technology, most of them are using credit card for buying their needs so the frauds associated with it is also rising gradually. In the present world, almost all the enterprises from small to big industries are using the credit card as mode of payment.

Credit card fraud is happening in all organization such as appliances industry, automobile industry, banks and so on. Many of the process like data mining, machine learning algorithmic approaches are applied to identify the fraud in the credit card transactions but did not get considerable result. Hence, there is a need of effective and efficient algorithms to be developed that works significantly.

We try to avoid the fraudster using our credit card before the transaction gets approved by using artificial neural network algorithm and compared with few other machine learning algorithms.

II. LITERATURE SURVEY

1) *A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine*

This paper introduces an intelligent fraud detection approach for credit card transactions using an optimized light gradient boosting machine (OLightGBM). It integrates a Bayesian-based hyperparameter optimization algorithm to enhance the performance of the model. Experimental results on real-world datasets demonstrate superior accuracy (98.40%), AUC (92.88%), precision (97.34%), and F1-score (56.95%) compared to other approaches.

.

2) *S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. Hacid and H. Zeineddine, "An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection,"*

Credit card fraud poses a significant threat to financial institutions and individuals, with 33,305 cases reported in the first half of 2018. Despite extensive research on fraud detection solutions, imbalance classification remains a critical issue, leading to ineffective outcomes. Current approaches result in numerous false alarms, proving costly and potentially inaccurate, compromising the effectiveness of fraud prevention.

3) *C. Jiang, J. Song, G. Liu, L. Zheng and W. Luan, "Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism," in IEEE Internet of Things Journal, vol. 5, no. 5, pp. 3637-3647, Oct. 2*

This paper introduces a novel fraud detection method for the increasing number of credit card transactions in e-commerce. The approach involves grouping cardholders based on similar transaction behaviors, utilizing a window-sliding strategy, extracting behavioral patterns, training classifiers for each group, and implementing a feedback mechanism to address concept drift. Experimental results demonstrate the method's superior performance compared to alternatives in detecting online transaction fraud.

4) *Ishan Sohony, Rameshwar Pratap, and Ullas Nambiar. 2018. Ensemble learning for credit card fraud detection. In Proceedings of the ACM India Joint International Conference on Data Science and Management of Data Association for Computing Machinery, New York*

Timely detection of fraudulent credit card transactions is a business critical and challenging problem in Financial Industry. Specifically, we must deal with the highly skewed nature of the dataset, that is, the ratio of fraud to normal transactions is very small. In this work, we present an ensemble machine learning approach as a possible solution to this problem.

Our observation is that *Random Forest* is more accurate in detecting normal instances, and *Neural Network* is for detecting fraud instances.

We present an ensemble method - based on a combination of random forest and neural network - which keeps the best of both worlds, and is able to predict with high accuracy and confidence the label of a new sample. We experimentally validate our observations on real world datasets.

5) *D. Prusti and S. K. Rath, "Web service based credit card fraud detection by applying machine learning techniques,"*

This study addresses credit card fraud in online transactions by employing a Deep Artificial Neural Network (DANN) on the Kaggle European cardholder dataset. Feature selection with Linear Discriminant Analysis and data balancing using SMOTE-ENN were applied. The proposed approach achieved impressive results with 98% accuracy, 100% precision, 98% recall, and a 99% f1-score.

III. SUMMARY OF LITERATURE REVIEW

This study proposes a Deep Artificial Neural Network (DANN) for credit card fraud detection using the Kaggle European cardholder dataset.

Employing Linear Discriminant Analysis and SMOTE-ENN for feature selection and data balancing, the DANN achieved 98% accuracy, 100% precision, 98% recall, and a 99% f1-score.

Another approach combines Random Forest and Neural Network in an ensemble method, addressing the skewed dataset issue, demonstrating high accuracy in detecting both normal and fraud instances. A novel fraud detection method groups cardholders, extracts behavioral patterns, and adapts to concept drift, outperforming alternatives.

An optimized LightGBM with Bayesian-based hyperparameter optimization attains superior accuracy, AUC, precision, and F1-score in real-world datasets.

IV. EXISTING SYSTEM

The existing systems are carried out by considering machine learning algorithms like Support Vector Machine, Naïve Bayes, K-Nearest Neighbor and so on and some of them used random dataset. Very few have used artificial neural network for credit card fraud detection.

V. PROPOSED SYSTEM

The Proposed system uses the Artificial Neural Network to find the fraud in the credit card transactions. Performance is measured and accuracy is calculated based on prediction.

And also classification algorithms such as Support vector machine and k-Nearest Neighbor are used to build a credit card fraud detection model.

We compare all the three algorithms used in the experiment and made a decision that artificial neural networks predicts well than system developed using support vector machine and k-nearest neighbor algorithms. The dataset used in the experiment consist of 31 attributes out of which 30 attributes consist of information related to name, age, account information and so on and last attribute give the outcome of the transaction in either 0 or 1.

VI. ARCHITECTURE

ANN is biologically inspired by human brain. The neurons are interconnected in the human brain like the same nodes are interconnected in artificial neural network.

Figure 1 depicts the structure of ANN with input, output and hidden layers. Inputs are x1, x2…Xn and output is y. w1…wn are the weights associated with inputs x1…xn respectively. There are 15 hidden layers used in this neural network. The activation function used in our credit card fraud detection model is RELU.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue II Feb 2024- Available at www.ijraset.com

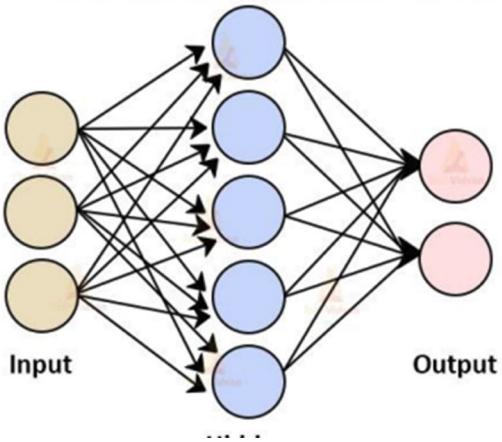


Fig. 1: Architecture of Artificial neural network

VII. METHODS

This section explains about the implementation, which includes the algorithm used for implementation of proposed system. In this paper, Implementations starts from loading the dataset. Than data pre-processing carried out that includes data cleansing and normalizing the data. Dataset is splitted into two dataset as train data and test data and model is trained and tested. Finally, system predicts whether transaction is fraud or non-fraud.

A. Programming Language Used

In the implementation of proposed system, we used python as programming language. Python is beginner's language, which provides various applications. In recent years, python had set the new trend because it is easy to use, interpreted, object-oriented, high-level, scripting language. Python is one of the best languages for the implementing machine learning. It provides rich packages and libraries that used in machine learning.

B. Packages and Libraries used

Some of the python library and packages used in proposed system are as follows:

1) Numpy Numpy is a python library. Abbreviation of Numpy is numerical python library. Numpy package is used for multidimensional arrays and linear algebraic operations.

2) Pandas Pandas is a python library. Pandas is used for data analysis and data manipulation tool. It is used to read the dataset and load the dataset. It is fast, flexible when working with data.

3) Scikitlearn A python package which is suitable for statistical model and machine learning models. A best suited python package for machine learning modeling.

4) Keras Keras is advanced stage of neural network application programming interface (API). It is able of run on top of tensor flow. Keras is mainly used while implementing deep learning algorithms such as CNN, RNN because its user friendly, modularity, and easy to extensibility. It runs on both CPU and GPU. In the experiment of finding the fraud or non fraud credit card transaction we had used Keras along with backend running tensor flow. This Keras along with tenor flow backend makes excellent choice for training neural network architecture.

5) MySQL MySQL is database which is used for storage purpose. In the experiment of fraud identification in card transaction we had used MySQL for storing the user details namely user name, password, email-id and phone number. While entering into application, user needs to register by providing the credential. These credentials are stored in database. Thereafter, user needs to login by giving username and password. The application will validate the login and registered information than user is moved to next window.

6) Tkinter Tkinter is python library which is used for Graphical User interface (GUI). It can be used on both Unix and Windows platform. We can create it by importing Tkinter module then GUI is created and one or more widgets are added finally, called in loop.

C. Classification Techniques

List of algorithm used in implementation of our experiment are:

- Support Vector Machine
- K-nearest algorithm
- Artificial Neural Network

1) Support Vector Machine

Pseudocode:

● Importing the necessary packages Example: import pandas as pd
● def SVM
Step 1: START
Step 2: Reading the dataset. pd.read.csv (file name) # reads the dataset file
Step 3: Data cleaning and preprocessing of data
● Resampling the data as normal and fraud class i.e. normal = 0 and fraud =1 under
● Under sampling of data is done
● Data is scaled (if any null value then eliminated) and normalized.
● Dataset is splitted into two set as train data and test data using split () on training data is used to split the data.
Step 4:Training the data using the SVM algorithm SVM classifier is called as classifier.predict ( ) # which predicts whether transaction fraud or nonfraud.
Step 5: Calculating the fraud transactions and valid transactions, then calculating the recall, precision and accuracy and stored in the respective locations
Step 6: STOP

2) K-Nearest Neighbor Pseudocode

Step 1: START
Step 2: Loading of dataset pd.read.csv (csv file) # reads the csv file and loads
Step 3: Cleaning and normalization of data
● Normal = 0
● Fraud = 1 # resampling
● Data is scaled and normalized
● Train_test_split() # splitting of dataset into train and test data
Step 4: Train the model then fit the trained model
● Trained the data using
● Knn classifier KNeighborsClassifier() # knn classifier which does classification of transactions
Step 5: Calculating the number of fraud, valid transactions and recall, precision and accuracy calculated.
Step 6: STOP

3) Artificial Neural Network (ANN):

Pseudocode:

The ANN algorithm has two parts:
● Training part and testing part.
● Training part:
○ Def ANN
Step 1: START
Step 2: Loading and observing the dataset
● pd.read.csv(.csv) # reads the dataset
● resampling of data
● StandardScaler() #scaling and normalization of data
Step 3: Data pre-processing
● Train_test_split() #Splitting of data

Step 4: Training the model
- Dense() #Adding data to activation function

Step 5: Analyzing the model
- Prediction of fraud is made and this trained data is stored .it can used to test (training the model takes longer time so it is stored)

Step 6: STOP

Testing part: Def ANN It is carried out similar way only difference is that thestored trained model is used to test the data and classify it.

VIII. RESULT AND DISCUSSION

Dataset: The proposed system makes use of the dataset downloaded from this website: www.kaggle.com . Dataset used is the transactions made by customer in a European bank in the year 2013-14. It consist of 31 columns, in which 30 columns are the features and the one class is the target class which decides about whether the transaction is fraud or non-valid. 5

Evaluation measure: The end result is evaluated based on the confusion matrix and precision, recall and accuracy is calculated. It contains two classes: actual class and predicted class.

The confusion metrics depends on these features:

1) *True Positive:* In which both the values positive that is 1. True Negative: it is case where both values are negative that is 0.
2) *False Positive:* This is the case where true class is 0 and non-true class is 1.
3) *False Negative:* It is the case when actual class is 1 and non-true class is 0.

Precision defined as follows:
- Precision = true positive / Actual result
- Precision = true positive/(true positive + false positive)

Recall defined as follows:
- Recall = true positive / predicted result ○ Recall = true positive/(true positive + false negative)

Accuracy defined as:
- Accuracy = (true positive + true negative)/ total

A. Result

Table 1 Represents the Accuracy, Recall and Precision

Algorithms	Accuracy	Precision	Recall
SVM	0.9349	0.9743	0.8976
KNN	0.9982	0.7142	0.0393
ANN	0.9992	0.8115	0.7619

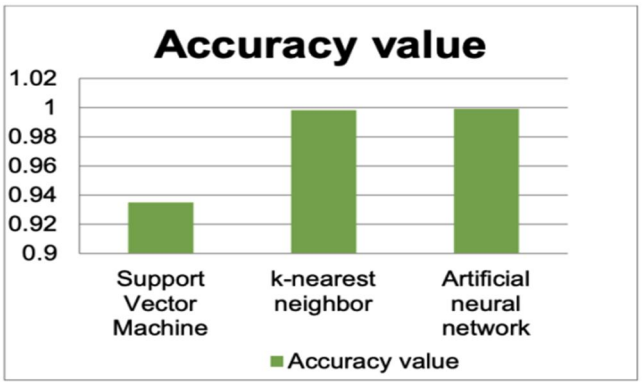


Fig. 2: represent the plot of accuracy obtained using SVM, KNN and ANN

ANN The figure 6 shows the accuracy performance measure of SVM, KNN and ANN algorithms ,This show that credit card fraud detection using artificial neural networks predicts at higher accuracy then Support vector machine and k-nearest neighbor algorithms for fraud detection in credit card transactions.

IX. CONCLUSION

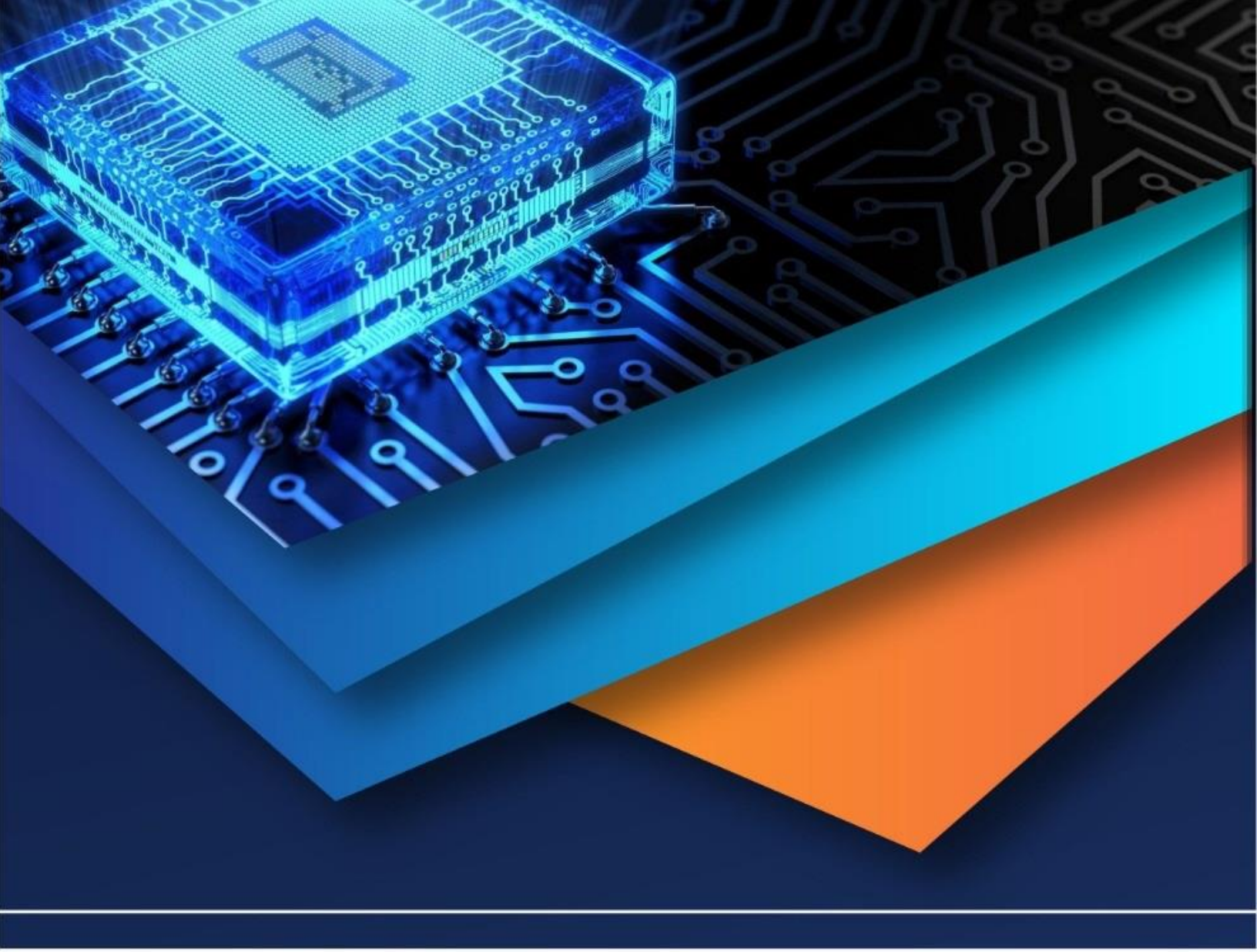
In this research, we have proposed a method to detect the fraud in credit card transactions that is based on deep learning. We first compare it with machine learning algorithms such as k-Nearest Neighbor, Support vector machine etc. Finally we have used the neural network, even though tough to train the model which would fit fine to model for detecting a fraud in credit card Transactions. In our model, by using an artificial neural network (ANN) which gives accuracy approximately equal to 100% is best suited for credit card fraud detection .It gives accuracy more than that of the unsupervised learning algorithms. In this research work, data pre-processing, normalization and under-sampling carried out to overcome the problems faced by using an imbalanced dataset.

X. ACKNOWLEDGMENT

The gratification and euphoria that come with the achievement of any work would be unfinished unless we mention the name of the people, who made it possible, whose relentless guidance and support served a beacon light and served our effort with success.
We express our sincere thanks and wholehearted credit to our internal guide Madhu H K, Professor, Department of MCA, BANGALORE INSTITUTE OF TECHNOLOGY ®, Bengaluru for his constant encouragement, support and guidance during the seminar work.

REFERENCES

[1] A. Taha and S. J. Malebary, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine," in IEEE Access, vol. 8, pp. 25579-25587, 2020, doi: 10.1109/ACCESS.2020.2971354.
[2] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. Hacid and H. Zeineddine, "An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection," in IEEE Access, vol. 7, pp. 93010-93022, 2019, doi: 10.1109/ACCESS.2019.2927266.
[3] C. Jiang, J. Song, G. Liu, L. Zheng and W. Luan, "Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism," in IEEE Internet of Things Journal, vol. 5, no. 5, pp. 3637-3647, Oct. 2018, doi: 10.1109/JIOT.2018.2816007.
[4] Ishan Sohony, Rameshwar Pratap, and Ullas Nambiar. 2018. Ensemble learning for credit card fraud detection. In Proceedings of the ACM India Joint International Conference on Data Science and Management of Data Association for Computing Machinery, New York, NY, USA, 289–294. DOI: https://doi.org/10.1145/3152494.3156815
[5] Phuong Hanh Tran, Kim Phuc Tran, Truong Thu Huong, Cédric Heuchenne, Phuong HienTran, and Thi Minh Huong Le. 2018. Real Time Data-Driven Approaches for Credit Card Fraud Detection. In Proceedings of the 2018 International Conference on E-Business and Application. Association for Computing Machinery, New York, NY, USA, 6–9. DOI: https://doi.org/10.1145/3194188.3194196
[6] Imane Sadgali, Nawal Sael, and Faouzia Benabbou. 2019. Fraud detection in credit card transaction using neural networks. In Proceedings of the 4th International Conference on Smart City Applications (SCA '19). Association for Computing Machinery, New York, NY, USA, Article 95, 1–4. DOI: https://doi.org/10.1145/3368756.3369082
[7] D. Prusti and S. K. Rath, "Web service based credit card fraud detection by applying machine learning techniques," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, pp. 492-497, doi: 10.1109/TENCON.2019.8929372.
[8] M. Zamini and G. Montazer, "Credit Card Fraud Detection using autoencoders based clustering," 2018 9th International Symposium on Telecommunications (IST), Tehran, Iran, 2018, pp. 486-491, doi: 10.1109/ISTEL.2018.8661129.
[9] S. Akila and U. S. Reddy, "Credit Card Fraud Detection Using Non-Overlapped Risk Based Bagging Ensemble (NRBE)," 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, 2017, pp. 1-4, doi: 10.1109/ICCIC.2017.8524418.
[10] M. S. Kumar, V. Soundarya, S. Kavitha, E. S. Keerthika and E. Aswini, "Credit Card Fraud Detection Using Random Forest Algorithm," 2019 3rd International Conference on Computing and Communications Technologies (ICCCT), Chennai, India, 2019, pp. 149-153, doi: 10.1109/ICCCT2.2019.8824930.
[11] Z. Li, G. Liu, S. Wang, S. Xuan and C. Jiang, "Credit Card Fraud Detection via Kernel-Based Supervised Hashing," 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/ SCI), Guangzhou, 2018, pp. 1249-1254, doi: 10.1109/SmartWorld.2018.00217.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY