INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.165

Credit Card Fraud Detection using Machine Learning

Dharma.P[1], Vinoth Kumar.G[2], Hemavathy.P.V[3]., Dr.Preetha.M[4]

UG Student, Dept. of I.T., Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai,

Tamil Nadu, India [1]

UG Student, Dept. of I.T., Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai,

Tamil Nadu, India [2]

Assistant Professor, Dept., of I.T Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai,

Tamil Nadu, India[3]

Professor, Dept. of I.T., Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai,

Tamil Nadu, India[4]

ABSTRACT: A credit card is issued by a bank or financial services company that allows cardholderstoborrowfundswithwhichtopayforgoodsandserviceswithmerchants that accept cards for payment. Nowadays as everything is made cyber so there is a chance of misuse of cards and the account holder can lose the money so 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. This type of problems can be solved through data science by applying machine learning techniques. It deals with modeling of the dataset using machine learning with Credit Card Fraud Detection. In machine learning the main key is the data so modeling the past credit card transactions with the data of the ones that turned out to be fraud. The built model is then used to recognize whether a new transaction is fraudulentornot.. Thefirststepinvolves analyzingandpre-processingdataandthenapplyingmachine learningalgorithmonthecreditcarddataset andfindtheparametersofthealgorithm and calculate their performance metrics.

KEYWORDS: Machin Learning Algorithm; Data Pre Processing; Modules; Analysing the modules ; Dataset;

I. INTRODUCTION

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains.Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.Natural language processing (NLP) allows machines to read and understand human language. Machine learning is to predict the future from past data. Machine learning (ML) is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed

II. RELATED WORK

They proposed a method and named it as Information-Utilization- Method INUM it was first designed and the accuracy and convergence of an information vector generated by INUM are analyzed. The novelty of INUM is illustrated by comparing it with other methods. Two D-vectors (i.e., feature subsets)aandb,whereAiistheitchfeatureinadataset,aredissimilarindecision space, but correspond to the same O-vector y in objective space. Assume that only a is provided to decision-makers, but a becomes inapplicable due to an accident or other reasons (e.g., difficulty to extract from the data set). Then, decision-makers are in trouble. On the other hand, if all two feature subsets are providedtothem,theycanhaveotherchoicestoservetheirbestinterest.Inother words, obtaining more equivalent D-vectors in the decision space can provide more chances for decision-makers to ensure that their interests are best served. Therefore,itisofgreatsignificanceandimportancetosolveMMOPswithagood Pareto front

approximation and also the largest number of D-vectors given each O-vector. The Disadvantages is They had proposed a mathematical model and machine learning algorithms is notused and Class Imbalance problem was not addressed and the proper measure were nottaken

III. PROPOSED ALGORITHM

The proposed model is to build a classification model to classify whether its fraud or not. The dataset of previous credit card cases are collected where it is used to make the machine to learn about the problem. The first step for involves the analysis of data where each and every column is analyzed and the necessary measurements are taken for missing values and other forms of data. Outliers and other values which are not much impact is dealt. Then preprocessed data is used to build the classification model where the data will be split into two parts one is for training and remaining data for testing purpose. Machine learning algorithms are applied on the training data where the model learns the pattern from the data and the model will deal with test data or new data and classify whether its fraud ornot. The algorithms are compared and the performance metric of the algorithms are calculated. The Advantages is Performance and accuracy of the algorithms can be calculated and compared.Class imbalance can be dealt with machine learningapproaches
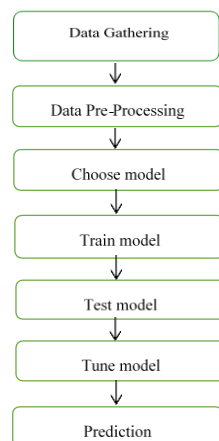

Fig.3.1. Process of Dataflow Diagram

IV. PSEUDO CODE

Step 1: Collect all the previous dataset of the Credit card users.
Step 2: Use machine learning with data science technology to find this.
Step 3: Then pre process and analyse the dataset with different types of modules.
Step 4: Use the four types of modules.
Step 5: Enter the data in each modules.
Step 6: Select the best accuracy result module.
Step 8: Then you get accuracy result and end.

V. SIMULATION RESULTS

In the credit card fraud detection, it is used by the machine learning technique (Artificial Intelligence) by with Data Science which is founded the output by using the four types of algorithm of Logistic Regression, Random Forest, Decision Tree Classifier, Naive Bayes with modules of Data Pre processing, Data Visualisation, Data Validation, Comparing algorithms and deployment and founded with the one algorithm which is the accuracy one is Logistic Regression and output is verified and runned successfully.

The Accuracy one of Logisytics regression processes, Data visualization,
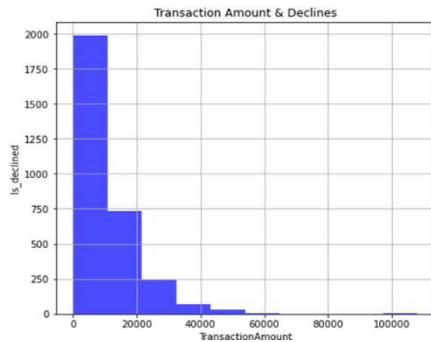
Fig.5.1. Data Analyzing of Visualization 1
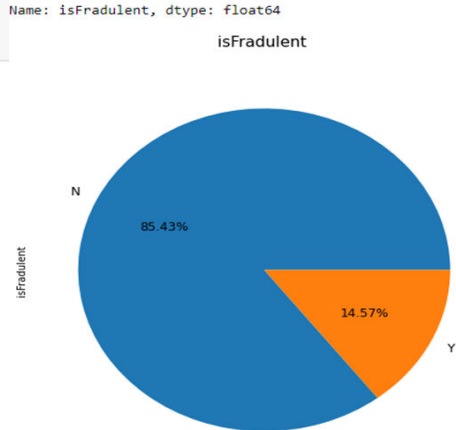
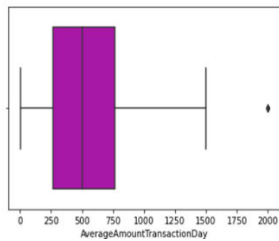Fig. Fig.5.2. Data Analyzing of Visualization 2
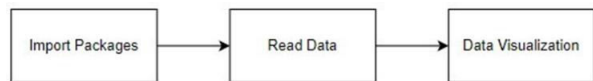

Fig.5.3. Data Analyzing of Visualization 3

Fig.5.4. Data Analyzing of Visualization 4

Logistic Regression is the accuracy, About of this module, It is a statistical method for analysing a data set in which there are one or moreindependentvariablesthatdetermineanoutcome.Theoutcomeismeasured withadichotomousvariable(inwhichthereareonlytwopossibleoutcomes).The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of interest (dependent variable = response or outcome variable) and a set of independent (predictor or explanatory) variable.
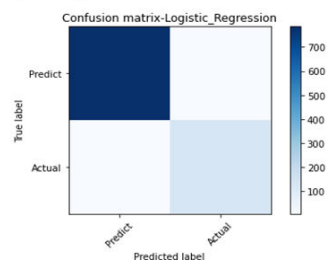

Fig.5.5. LogisticRegression 1
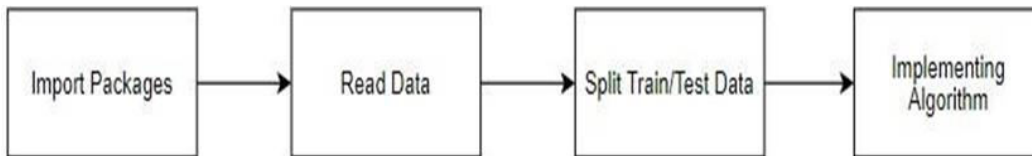

Fig.5.6.LogisticRegression 2

Fig.5.7. LogisticRegression 3

VI. CONCLUSION AND FUTURE WORK

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set is higher accuracy score will be find out. This application can help to find the Prediction of credit card fraud or not and the future work is credit card fraud prediction to connect with cloud model and to optimize the work to implement in Artificial Intelligence environment.

REFERENCES

[1]   Q. Wu, M. Zhou, Q. Zhu, Y. Xia, and J. Wen, "MOELS: Multiobjective evolutionary list scheduling for cloud workflows," IEEE Trans. Autom. Sci. Eng., vol. 17, no. 1, pp. 166–176, Jan. 2020.

[2]   L. Huang, M. Zhou, and K. Hao, "Non-dominated immune-endocrine short feedback algorithm for multi-robot maritime patrolling," IEEE Trans. Intell. Transp. Syst., vol. 21, no. 1, pp. 362–373, Jan. 2020.

[3]   X. Wang, K. Xing, C.-B. Yan, and M. Zhou, "A novel MOEA/D for multiobjective scheduling of flexible manufacturing systems," Complexity,vol. 2019, pp. 1–14, Jun. 2019.

[4]   J. J. Liang, C. T. Yue, and B. Y. Qu, "Multimodal multi-objective optimization: A preliminary study," in Proc. IEEE Congr. Evol. Comput. (CEC), Jul. 2016, pp. 2454–2461.

[5]   P. S. Oliveto, D. Sudholt, and C. Zarges, "On the benefits and risks of using fitness sharing for multimodal optimisation," Theor. Comput. Sci.,vol. 773, pp. 53–70, Jun. 2019. J. J. Liang, S. T. Ma, B. Y. Qu, and B. Niu, "Strategy adaptative memetic crowding differential evolution for multimodal optimization," in Proc IEEE Congr. Evol. Comput., Jun. 2012, pp. 1–7.

[6]   M. M. H. Ellabaan and Y. S. Ong, "Valley-adaptive clearing scheme for multimodal optimization evolutionary search," in Proc. 9th Int. Conf.Intell. Syst. Design Appl., Nov. 2009, pp. 1–6.

[7]   X. Li, "Efficient differential evolution using speciation for multimodal function optimization," in Proc. Conf. Genetic Evol. Comput. - GECCO,Jun. 2005, pp. 873–880.

[8]   B. Y. Qu, G. S. Li, Q. Q. Guo, L. Yan, X. Z. Chai, and Z. Q. Guo, "A niching multi-objective harmony search algorithm for multimodal multi-objective problems," in Proc. IEEE Congr. Evol. Comput. (CEC), Jun. 2019, pp. 1267–1274.

[9]   C. Yue, B. Qu, and J. Liang, "A multiobjective particle swarm optimizer using ring topology for solving multimodal multiobjective problems,"IEEE Trans. Evol. Comput., vol. 22, no. 5, pp. 805–817, Oct. 2018.

[10]   L. Yan, G. S. Li, Y. C. Jiao, B. Y. Qu, C. T. Yue, and S. K. Qu, "A performance enhanced niching multi-objective bat algorithm for multimodal multi-objective problems," in Proc. IEEE Congr. Evol. Comput. (CEC), Jun. 2019, pp. 1275–1282.

[11]   Y. Liu, H. Ishibuchi, G. G. Yen, Y. Nojima, and N. Masuyama, "Handling imbalance between convergence and diversity in the decision space in evolutionary multi-modal multi-objective optimization," IEEE Trans. Evol. Comput., vol. 24, no. 3, pp. 551–565, Jun. 2020.

[12]   E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," IEEE Trans. Evol. Comput., vol. 7, no. 2, pp. 117–132, Apr. 2003.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 9940 572 462 🟢 6381 907 438 ✉ ijircce@gmail.com

Scan to save the contact details