##### Vemu logo VEMU INSTITUTE OF TECHNOLOGY

Approved by AICTE-New Delhi, Permanently Affiliated to JNTUA

Accredited by NBA (CSE, ECE & EEE) & NAAC, Recognized by 2(F) & 12(B) UGC

Act. An ISO 9001:2015 Certified Institution.

Tirupati – Chittoor Highway, P. Kotha Kota, Pakala – 517112.

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

ACADEMIC YEAR:2022-2023 MONTH: APRIL

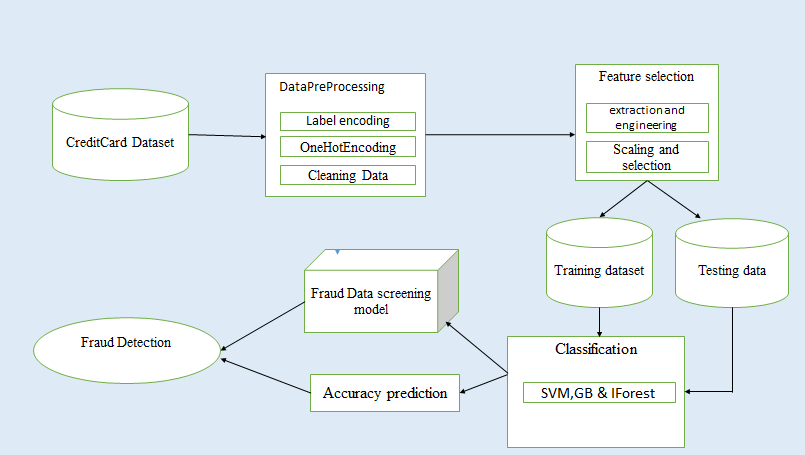
**SUPERVISED AND UNSUPERVISED MACHINE LEARNING ALGORITHMS FOR CREDIT CARD FRAUDULENT TRANSACTION DETECTION**

**ABSTRACT**

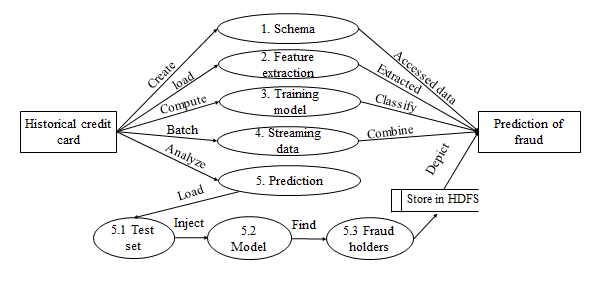
The goal of data analytics is to delineate hidden patterns and use them to support informed decisions in a variety of situations. Credit card fraud is escalating significantly with the advancement of the modernized technology and become an easy target for fraudulent. Credit card fraud is a severe problem in the financial service and costs billions of a dollar every year. The design of fraud detection algorithm is a challenging task with the lack of real-world transaction dataset because of confidentiality and the highly imbalanced publicly available datasets.

Apply different unsupervised machine learning algorithms to detect credit card fraudulent transaction using a real-world dataset. Furthermore, employ these algorithms to implement a super classifier using ensemble learning methods. Identify the most important variables that may lead to higher accuracy in credit card fraudulent transaction detection. Additionally, we compare and discuss the performance of various supervised machine learning algorithms exist in literature against the super classifier that can implemented.

**ARCHITECTURE**



**DESIGN**

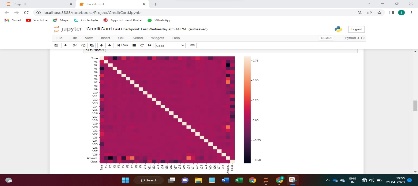
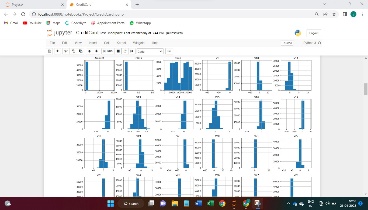


* DFDs are used to Specify Functions of the Information System and how data flow from function to function. It’s a collection of function that manipulates data. On a DFD, data items flow from an external data source or an internal data store to an internal data store or an external data sink, via an internal process. A DFD provides no information about the timing of processes, or about whether processes will operate in sequence or in parallel.
* DFD level 0 for credit card fraud detection explains the overall functionality of the entire project. The fig below shows the data flow diagram level 0 for credit card fraud detection.
* From the above diagram shows the overall representation of each module and their functionality in this project.
* In this module using catalyst optimization algorithm create data frame. Then catalogue the optimized logical path. Analyse the physical plans and store the data to the data base.
* This section extends the analysis of the techniques proposed previously for the big data forensic investigation. There are different factors that can influence the selection and performance of the forensic techniques. These factors should be analysed closely before carrying out the implementation. The following table has been populated based on the available resources and the sensitivity of the data to be used for investigation.

**ADVANTAGES**

* More efficient result for detect Fraud Detection discover.
* High accuracy
* Less detection times
* Get user based historical records.
* Avoid complex problem.

**SCREEN SHOTS**

**CONCLUSION**

Detection of credit card fraud is an intent part of testing for the researchers over a long time and will be an interesting part of testing in the coming time. We are introducing a fraud detection system for credit-cards by applying three different algorithms and training our machine using these algorithms with the transaction records we have. We used a supervised and unsupervised algorithm like IForest, SVM and GB to select the features from our transaction dataset which uses correlation and variance as parameters to select the features. We have set the summation of variance as 95% for selecting the features using the IForest, SVM and GB algorithm. We also applied the IForest, SVM and GB feature selection algorithm as there are no features which have high variance and correlation with the class column which determines the transaction as fraud or not. We have applied the three machine learning algorithms as stated in methodology and the models indicate a high accuracy score for each one of them. The scores of each model were 99.7%, 99.8%, 99.7% for the iforest, support vector machine and gradient boosting classifier algorithms respectively. As these models have high accuracy and get the best results with high precision in determining the fraud detections in credit card transaction records.

**Present By**

**G. JYOTHIKA (194M1A0530)**

**A.S. SRIDEVI (194M1A0505)**

**D.LIKHITHA (194M1A0523)**

**K. VIDYA (194M1A0551)**

**IV-CSE-A**

**Under the Guidance of**

**Dr. C. SIVA BALAJI YADAV, Ph.D., Associate Professor, Dept of CSE.**