

## Credit Card Fraud Detection by using ANN and Decision Tree

*Jasmine A Hudali\*, Kamalakshi, K P Mahalaxmi, Namita S Magadum,  
Prof. Sudhir Belagali*

*Department of Computer Science,  
KLE College of Engineering and Technology, Karnataka, India.*

*\*Corresponding Author*

*E-mail Id:-jashudali@gmail.com*

### ABSTRACT

*Credit card fraud detection is one of the biggest ethical issues. Credit card fraud detection has different types of frauds. Based on those types of fraud it will apply the fraud detection techniques. There are several techniques for detecting the frauds. Those techniques take the input data and tell the user to that particular input has fraud or not. If the transaction is fraud it will takes some actions like sending the message to owner and credit card industries.*

**Keywords:** *Credit card fraud detection, Artificial neural network, Decision tree.*

### INTRODUCTION

The credit card fraud is a biggest problem. Fraud is one of the major 16 digit number is not enough for secure transactions. The main aim of this project is to identify the different types of credit card fraud and to review the alternative techniques that have been used in fraud detection. The main goal of credit card fraud detection is majority of users could not be fraudulent; we need to protect the user privacy. For implementing this we are using different techniques like deep learning, sampling techniques and real dataset. The types of deep learning algorithms like supervised, unsupervised and semi supervised algorithms.

We will use [3] supervised learning algorithm to detect the fraud that algorithm will takes the input and gives some expected output. In our case input will be credit card transaction and output will be the information that says, that input could response to fraud detection or not. The perceptron is an algorithm for supervised learning of binary classifier perceptron uses the feature extraction to solve our classification problem. Neural network problem is used to detect the fraud, it has 3

layers, input layer, hidden layer and output layer. The output given by these layers that are compared to the negative transactions we have in the dataset.

Credit cards are one of the most targets of fraud but not the only one. Fraud can occur with any types of credit products such as personal loans, home loans, and retail. For this credit card fraud detection we are going to use the transactions of the credit card as dataset. From a perspective it can be argued that banks and credit card companies should attempt to detect all fraudulent cases. Credit card fraud detection is divided into two categories. First one is Fraud Analysis is deals with the [3] supervised technique it takes the training data with desired output and second one is user behavior analysis that deals with the unsupervised method it takes only training data without desired output.

### EXISTING SYSTEM

If fraudster can steal our card, card number, or PIN the card owner may face many problems to block the card. Fraudster can identify card number or PIN through the help of cyber crime to

investigate the fraud. To avoid the entire disadvantages of credit card fraud we proposed some detection techniques.

### PROPOSED SYSTEM

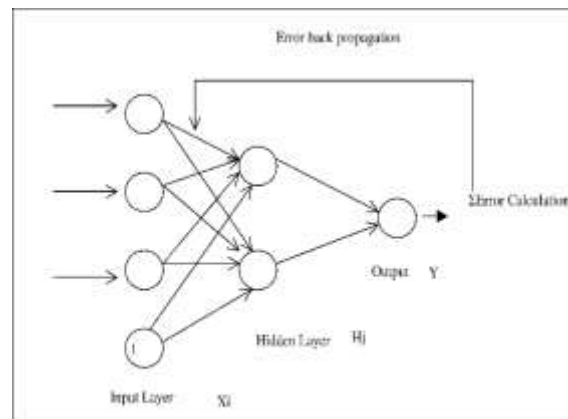
We proposed some detection technique to avoid the frauds is artificial neural network and decision tree. Here we need to deal with unbalanced dataset and the system will tell about the fraudulent over non fraudulent transactions. It gives the better performance to detect the fraud.

### DETECTION TECHNIQUES

**Neural networks:** Artificial neural networks provide a general, practical method for learning real-valued, discrete-valued and vector valued functions. In neural network there are three layers namely input layer, hidden layer and output layer. The input layer takes the input and multiplied by their some weights and sends it into next layer that is hidden layer. The hidden layer calculates the sum and that is send to output layer. The job of output layer compares the output which is given by the hidden layer and targeted output .if that are same that output will be updated. If there is an error found that will back propagate the output. One type of ANN system is based on a unit called a perceptron. Perceptron takes a vector valued of real-valued inputs, calculates the linear combination of these inputs.[5]

**Back propagation:** The back propagation algorithms learn the weights for a multilayer network, given a network with a fixed set of units and interconnection. It employs gradient descent to attempt to minimize the squared error between the network output values and the target values for these outputs. In the back propagation we first gives our input than compares our predicted input with the expected output after this will check for the errors if there is the presence of errors then it propagates the entire network for the correction of errors, then we need to

identify the weight once we get the weight then we have to update the weights and this process continues till we get the accuracy.



Notes: The weight connecting node  $i$  in the input layer to node  $j$  in the hidden layer is denoted by  $W_{ij}$ , and the weight connecting node  $j$  in the hidden layer to the output node is represented by  $V_j$

**Fig.1:-Back Propagation**

**Decision tree:** Decision Tree is a classifier which is imported from the scikit library. Decision tree is a method for approximating discrete valued target functions, in which the learned function is represented by a decision tree. Decision tree provides the classification of the instances, each node in the tree specifies a test of some attribute of the instance, and each branch descending from that node corresponds to one of the possible values for this attribute. An instance is classified by starting at the root node of the tree, testing the attribute specified by the node, then moving down the tree branch corresponding to the value of the attribute in the given example. This process is then continued for the subtree rooted at the new node. For the decision tree calculation we split our dataset into training dataset and the testing dataset based on the dataset which we considered. We use prediction method to measure the performance in test dataset.[4]

Decision tree is a flowchart like structure each internal node denotes a test attribute which represents outcome of attribute. Decision tree can be used divisibly and

explicitly to represent decisions and decision making. Regression trees are used with dependent variable is continues. Classification trees are used when dependent variable is categorical. Confusion matrix will take the expected classes as arguments it is used to get the prediction by this will get to know about the fraudulent transactions.[2]

Positive (P): Observation is positive.

Negative (N): Observation is negative.

True Positive (TP): Observation is positive, and is predicted to be positive.

False Negative (FN): Observation is positive but is predicted negative.

True Negative (TN): Observation is negative and is predicted to be negative.

False Positive (FP): Observation is negative but is predicted positive.

**Accuracy:** It is the proportion of correct classifications from overall number of cases.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

**Recall:** It is the proportion of correct positive classifications from cases that are actually positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$



**Fig.2:-**Decision Tree for Play Tennis

**Precision:** It is the proportion of correct positive classification from cases that are predicted as positive.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

## ANALYSIS AND RESULT

In credit card fraud detection analysis the input data by using neural networks it's a practical method for learning real-valued, discrete-valued and vector-valued functions.

And use back propagation where it takes our input then compares our predicted input with the expected output and check for the errors if there is a presence of errors then it propagates entire network for the correction of errors and the back propagation provides the accuracy and decision tree works as a classifier and it classifies fraudulent and non fraudulent transaction. It shows the transaction is fraudulent or non fraudulent if the transaction is fraudulent then it identifies the fraud and reduces the chances of fraudulent transaction.

## LITERATURE SURVEY

1. <https://ieeexplore.ieee.org/xpl/conhome/5158913/proceeding>.
- This paper proposes a credit card fraud detection model using outlier detection based on distance sum according to the infrequency and conventionality of fraud in credit card transaction data, applying outlier mining into credit card fraud detection.
- This paper shows the model of feasible and accurate in detecting credit card fraud.
2. <https://ieeexplore.ieee.org/xpl/conhome/8784131/proceeding>.
- This paper presents a method to quantify day-by-day the dataset shift in our face-to-face credit card transactions dataset. Classify the days against each other and measure the efficiency of the classification.
- The more efficient the classification, the more different the buying the

behavior between two days, and vice versa.

3. <https://ieeexplore.ieee.org/xpl/conhome/7511893/proceeding>.
- This research paper focuses on the creation of a scorecard from relevant evaluation criteria, features, and capabilities of predictive analytics vendor solutions currently being used to detect credit card fraud.
- The scorecard provides a side-by-side comparison of five credit card predictive analytics vendor solutions adopted in Canada.
4. <https://ieeexplore.ieee.org/xpl/conhome/4280076/proceeding>.
- This study investigates the efficacy of applying classification models to credit card fraud detection problems.
- Classification method logistic regression is tested for their applicability in fraud detections.
- This paper provides a useful framework to choose the best model to recognize the credit card fraud risk.
5. <https://ieeexplore.ieee.org/xpl/conhome/8316803/proceeding>.
- This paper presents a survey of current techniques used in credit card fraud detection and telecommunication fraud.
- The goal of this paper is to provide a comprehensive review of different techniques to detect fraud.

## CONCLUSION

Credit card fraud detection defines fraudulent over non fraudulent transaction. There are 2 labels to define the fraud transaction that are 0 and 1. 0 indicates the non-fraudulent transaction where 1 indicates the fraudulent transaction. This paper defines; first it identifies the type of fraud based on the type of fraud applying alternative techniques to detect the fraud. Whenever a card gets fraud, the owner of the card has right to block the card, this will protect the user privacy.

## REFERENCES

1. Linda Delamaire, Hussein Abdou, John Pointon *Credit card fraud detection techniques: a review*. 2009.
2. Sotiris Kotsiantis, Dirnitris Kanellopouios, Panayiotis Pintelas. *Handling Imbalanced Datasets: A review*. 2006.
3. Suman. *Survey Paper on Credit Card Fraud Detection*. 2014.
4. Aihua Shen, Rencheng Tong, Yaochen Deng. *Application of Classification Models on Credit Card Fraud Detection*. 2007.
5. Y.Sahin, E. Duman. *Detecting Credit Card Fraud by ANN and Logistic Regression*. 2011.