**Credit Card Fraud Detection Using**

**State-of-the-Art Machine Learning**

**and Deep Learning Algorithms**

**Alternative Title:**

Credit card Fraud detection using Machine learning and Deep learning Algorithms.

**Aim:**

People can use credit cards for online transactions as it provides an efficient and easy-to-use facility.With the increase in usage of credit cards, the capacity of credit card misuse has also enhanced. Credit card frauds cause significant financial losses for both credit card holders and financial companies.the main aim is to detect fraudulent transactions using credit cards with the help of ML algorithms and deep learning algorithms.

**Abstract:**

The main focus has been to apply the recent development of deep learning algorithms for this purpose. Comparative analysis of both machine learning and deep learning algorithms was performed to find efficient outcomes. The detailed empirical analysis is carried out using the European card benchmark dataset for fraud detection.

**Synopsis:**

A machine learning algorithm was first applied to the dataset, which improved the accuracy of detection of the frauds to some extent. Later, three architectures based on a convolutional neural network are applied to improve fraud detection performance. Further addition of layers further increased the accuracy of detection. A comprehensive empirical analysis has been carried out by applying variations in the number of hidden layers, epochs and applying the latest models. The evaluation of research work shows the improved results achieved, such as accuracy, f1-score, precision having best optimized values respectively. The proposed model outperforms the state-of-the-art machine learning and deep learning algorithms for credit card detection problems. In addition, we have performed experiments by balancing the data and applying deep learning algorithms to minimize the false negative rate. The proposed approaches can be implemented effectively for the real-world detection of credit card fraud.

**Existing System:**

The relevant literature present many machines learning based approaches for credit card detection, such as Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression and XG Boost, which results low accuracy.In 2020, there were 393,207 cases of CCF out of approximately 1.4 million total reports of identity theft . CCF is now the second most prevalent sort of identity theft recorded as of this year, only following government documents and benefits fraud . In 2020, there were 365,597 incidences of fraud perpetrated using new credit card accounts. The number of identity theft complaints has climbed by 113% from 2019 to 2020, with credit card identity theft reports increasing by 44.6%. Payment card theft cost the global economy $24.26 billion last year. With 38.6% of reported card fraud losses in 2018, the United States is the most vulnerable country to credit theft.

**Problem Definition:**

By proposing Machine learning Algorithms, based approaches for credit card detection, such as Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression and XG Boost .The model results leds to low accraucy.

**Proposed System:**

Deep learning (DL) algorithms applied applications in computer network, intrusion detection, banking, insurance, mobile cellular networks, health care fraud detection, medical and malware detection, detection for video surveillance, location tracking, Android malware detection, home automation, and heart disease prediction.we explore DL Algorithms to identify credit card thefts in the banking industry in this model.It uses a number of deep learning algorithms for detecting CCF. However, in this model,we choose the CNN model and its layers to determine if the original fraud is the normal transaction of qualified datasets.

**Advantage:**

The imbalanced CCF dataset is transformed into a balanced dataset by removing non fraudulent transactions from the dataset. In a real-world transaction, fraudulent and non- fraudulent classes are not balanced due to the nature of the problem. For instance, if one million transactions are per- formed in a day, only a few can be fraudulent. The convolutional neural network model with layers architecture is applied to the balanced dataset to validate the proposed model. The model is trained over 100 epochs. The CNN layers architecture obtained above 90.00 % training and validation accuracy respectively. The accuracy and loss of CNN model using the balanced CCF dataset.

**Modules:**

* Dataset Collection
* Algorithm
* Detection

**Dataset Collection:**

Collect the dataset from Kaggle.com.

**Data Preprocess:**

We propose to alter the DL algorithm of the CNN model by adding the additional layers for features extraction and the classification of credit card transactions as fraudulent or otherwise. In this model , the main aim is to detect fraudulent transactions using credit cards with the help of deep learning algorithms.In First process,the imbalanced CCF dataset is transformed into a balanced dataset by removing non fraudulent transactions from the dataset. In a real-world transaction, fraudulent and nonfraudulent classes are not balanced due to the nature of the problem. For instance, if one million transactions are performed in a day, only a few can be fraudulent. The convolutional neural network (CNN) model with 14 layers architecture is applied to the balanced dataset to validate the proposed model. The model is trained over 100 epochs.

**Algorithm:**

Convolutional neural network (CNN): In deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet.It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other. Object detection is widely used for image processing and classification, estimating time series and detecting differences.Layers in the CNN Model: Here are six distinct layers in the CNN models are Inputlayer,Convo layer (Convo C ReLU), Pooling layer,Fully connected layer (FC),SoftMax/logistic layer, Output layer. Creation of the Model, The pipeline ofCNNmodel over keras includes conv layer, max pooling layer, dropout layer, convolayer, max pooling layer, dropout layer along with two fully connected layers sequentially, depicts input neuralnetwork and output of dropout layer. Compile the Model: Categorical Cross-Entropy: We buildbinary cross-entropy at prior portions and in ML.At that time, we used definite cross-entropy.PERFORMANCE-EVALUTION MEASURES: Traditional methods of estimating ML classifiers can use confusion metrics,precision,recall and F1 score relating to the difference between the rock bottom dataset truth and the model's prediction where TP, TN, FP, and FN denote true positive, true negative, falsepositive and false negative, respectively.

**Prediction:**

The imbalanced CCF dataset is transformed into a balanced dataset by removing non fraudulent transactions from the dataset. In a real-world transaction, fraudulent and nonfraudulent classes are not balanced due to the nature of the problem. For instance, if one million transactions are performed in a day, only a few can be fraudulent.

**Hardware Requirements:**

* Hard Disk : 500GB and Above
* RAM : 4GB and Above
* Processor : I3 and Above

**Software Requirements:**

* Operating System : Windows 10 (64 bit)
* Software : Python

Tools : Anaconda

**Archiecture diagram:**

Pre-ProcessDataset

Convolutional Neural Network

Feature Extraction

Real time

User input

User interface

Prediction

Model compile

 Image

Trained Model

**Conclusion:**

Detecting CCF, the number of features, number of transactions, and correlation between

the features are essential factors in determining the model’s performance. DL methods, such as CNNs and their layers, are associated with the processing of text and the baseline model. Using these methods for the detection of credit cards yields better performance than traditional algorithms.

**Future Enchanments:**

To increase the performance of existing examples, but they significantly decrease on the unseen data. The performance on unseen data increased as the class imbalance increased. Future work associated may explore the use of more state of art deep learning methods to improve the performance of the model proposed in this study.