Idea/Approach Details

Technology Bucket : Finance
Company Name-FIS SOLUTIONS(INDIA)
Team Leader Name : Sarthak Srivastava

Category: Software
Problem Code : AK2
College Code : U-0840

Prototype - Our solution to the given problem related to Income Tax Fraud Detection is based on supervised learning using Artificial neural network which simulates like human brain in order to produce a cognitive reasoning which in this case is to identify whether an individual is fraud or not using the dataset available to it .We propose the use of ANN prediction algorithms applied to fraud detection in money transaction data of different individuals. Our ANN makes the prediction and the real results of an individual's transaction are compared with the prediction . The problem in context of supervised learning is a classification problem which is solved using ANN classifiers.

These are our steps to approach the problem:-

- Upload the dataset using pandas library
- Normalization using Z-score for data smoothing as data may be noisy and inconsistent . In case of outliers, Z-Score proves to be better when compared to Min- Max Normalization. Z-score normalization is calculated using - Z = (X - Mean) / (Standard Deviation) where X is feature
- Train-Test split
- Set up ANN and initialized it with the activation function ,learning rate , MSE and no of epoch
- Normalized trained features are passed as input to neural network setup

- Our ANN classifier learns from inputs given to the model using continuous forward and backpropagation until a cost function of given hypothesis is minimized. Predictions are compared with real values using confusion matrices

Our model works wonderfully for credit-card fraud detection systems . Training of credit card fraud dataset was done through various models such as ensemble system , clustering ,auto-encoders , Gaussian Regression etc but ANN proved to produce better results.

The periodic fetching , updation of dataset and re-training of dataset will be done through our model.

At last when the fraudster is identified all the information related to that individual like account no will be notified to income tax department through an api .

Our model clearly spots-

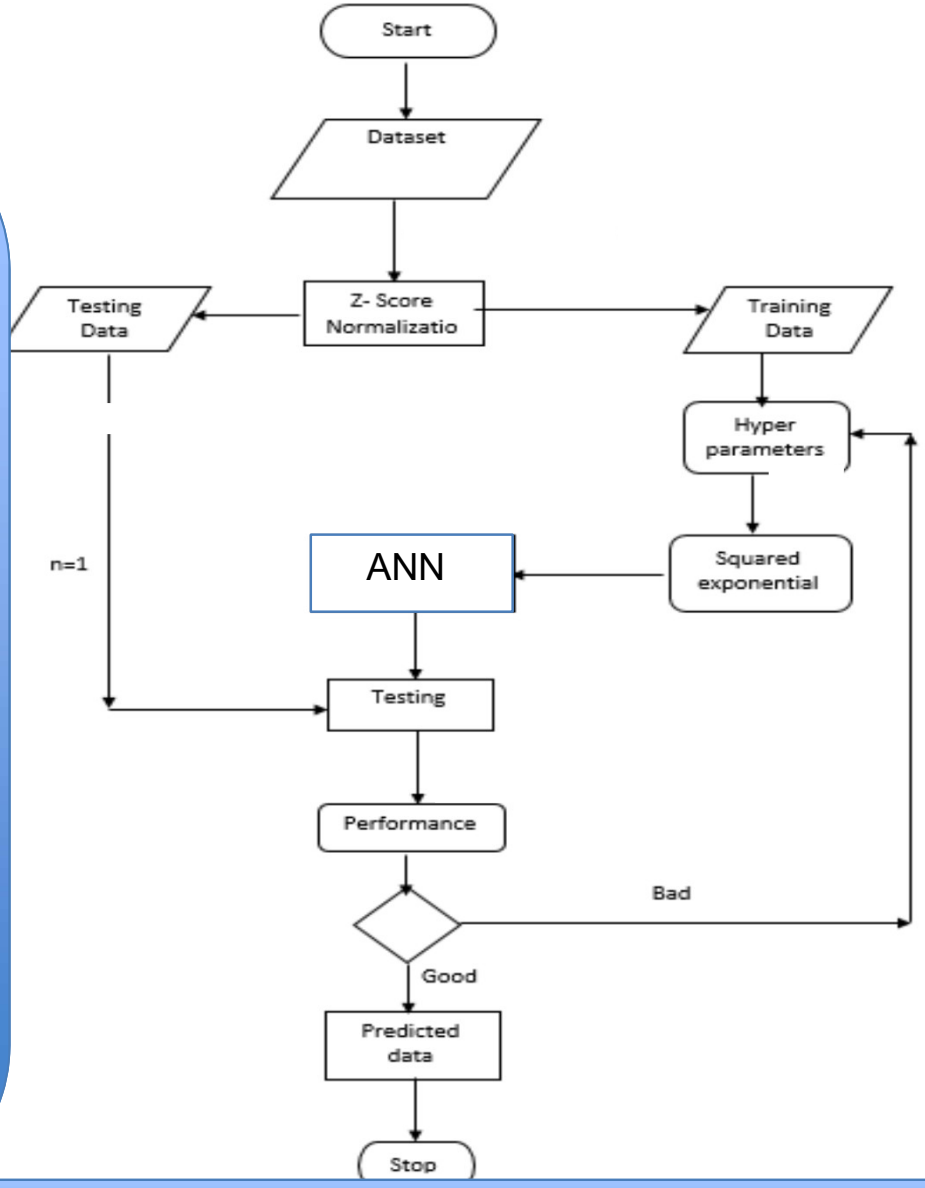
- Any suspicious transaction that gets flagged
- Identifies new criminal patterns
- Replace primitive Rule-based system

Technology stack-We are using Python as our preferred language and the neural network is developed using open source libraries Keras and Tensorflow along with machine learning libraries like Sklearn ,Scipy and Pandas . We are incorporating our project with Django for web development. Our proposed system has the ability to work directly on large databases. Our model is based on highly mathematical tools of linear algebra and produces highly accurate predictions when the sample is quite large .

Idea details

Use Case-Our model is beneficial for firms and organisation in detecting the fraudsters , who are hurdles for its development.

- Banking-In looking suspicious transaction and finding of fraudsters
- Insurance –Preventing false claiming of policies
- Retailers- In marketing it provides insight into customer purchasing behaviour and can detect fraud with accuracy and false positive
- Telephonic companies-to track suspicious fraudsters
- government agencies , e-commerce etc.



Dependencies-Suitable dataset that contains various features like transaction history , salary details ,total amount withdrawal and of account ,no of accounts linked ,account database etc. Python 3.5 version and above , Octave.