CS6890: Fraud Analytics Using Predictive and Social Network Techniques
Assignment 4: Fraud Detection Using an Autoencoder and Variational Autoencoder

Archit Vivek Ganvir (CS21BTECH11005)
Maharshi Kadeval (CS21BTECH11027)
Abhinav Yadav (CS21BTECH11002)
Harsh Goyal (CS21BTECH11020)
Anshul Sangrame (CS21BTECH11004)

May 2024

1 Problem Statement

Develop an ML model to identify fraudulent transactions among a given set of credit card transactions.

2 Introduction

In this assignment, we use an autoencoder and a variational autoencoder to identify fraudulent transactions within a given dataset of legitimate and fraudulent credit card transactions.

An autoencoder is a type of artificial neural network used to learn efficient codings of unlabeled data (unsupervised learning). An autoencoder learns two functions: an encoding function that transforms the input data, and a decoding function that recreates the input data from the encoded representation.

In this assignment, we use an autoencoder for anomaly detection, wherein we train our autoencoder on only legitimate transactions. We then calculate a threshold value for the error, and test the model on a set of transactions (both legitimate and fraudulent). The transactions for which the error is above the threshold value are classified as fraudulent transactions.

Variational autoencoders are like autoencoders, but an input point is mapped to a distribution within the latent space, rather than to a single point in that space.

3 Description of Dataset

The transactions have two labels: “1” for fraudulent and “0” for normal transactions. The total number of transactions is 284,807. Only 492 (0.2 percent) of them are fraudulent. Each credit card transaction is represented by 30 features: 28 principal components extracted from the original credit card data, the time between each transaction and the first transaction in the data set, and the amount paid for each transaction.

4 Procedure

- The Time and Amount columns in the dataset are converted to log scale for dynamic range compression
- The values in each column are normalized (by MinMax scaling).
- The dataset is split into legitimate and fraudulent transactions.
- The legitimate transactions are divided into train and test sets in the ratio of 95:5.
- The autoencoder is designed with 30-15-7-15-30 units in each layer (input layer, 2 hidden layers for encoder, 1 hidden layer for decoder and 1 output layer) with ReLU as the activation function.
- The autoencoder is trained on the train set with the adam optimizer, MAE (Mean Absolute Error) as the loss function and a batch size of 64 for 100 epochs.
- The threshold value is calculated by reconstructing the train set.
- The validation set is created by combining the test set and the set of fraudulent transactions.
- The validation set is reconstructed, and the transactions with an error greater than the threshold value are classified as fraudulent transactions.
- Various metrics such as accuracy, precision, recall and F1 score are calculated.
- Similar steps are followed for the variational autoencoder, which is designed with 30-15-7-4-15-30 units in each layer (input layer, 3 hidden layers for encoder, 1 hidden layer for decoder and 1 output layer) with ReLU as the activation function.

5 Results

Following are the results with the autoencoder:
True Positives = 435
False Positives = 1617
True Negatives = 12599
False Negatives = 57
Accuracy = 0.8861843894479196
Precision = 0.21198830409356725
Recall = 0.8841463414634146
F1 Score = 0.34198113207547176

Following are the results with the variational autoencoder:
True Positives = 443
False Positives = 1872
True Negatives = 12344
False Negatives = 49
Accuracy = 0.8693908077236878
Precision = 0.1913606911447084
Recall = 0.9004065040650406
F1 Score = 0.3156394727467047

- The accuracy and recall values are sufficiently good with both models, since most of the transactions (both legitimate and fraudulent) are correctly classified.

- The precision is very low with both the models, since a lot of legitimate transactions are classified as fraudulent (false positives). This is unavoidable, since the dataset has a very small proportion of fraudulent transactions.

- The autoencoder is able to correctly classify 435 of the 492 fraudulent transactions (88.42%), and the variational autoencoder is able to correctly classify 443 of the 492 fraudulent transactions (90.04%).

- Increasing the threshold value increases the recall, but reduces accuracy and precision. This would help to identify a higher number fraudulent transactions, but would also greatly increase the number of false positives.

- Decreasing the threshold value increases precision and accuracy, but reduces recall. This would lead to a decrease in the number of true positives.