Credit card Fraud Detection

Name: Omar Tahir | Goldsmiths, University of London | Department of Computer Science

Supervisor: Tim Blackwell

Abstract

This research project implements machine learning. This includes five different algorithms that were used these include K-nearest neighbour, Decision Trees, Support Vector Machine, Random Forest and finally Convolutional Neural Network.

All algorithms were run twice using different balancing techniques. All models for credit card fraud detection are developed in Python using the TensorFlow and Keras libraries, once all the modules have been runned they have been compared between each other to ensure that the final model selected for use is the most effective one.

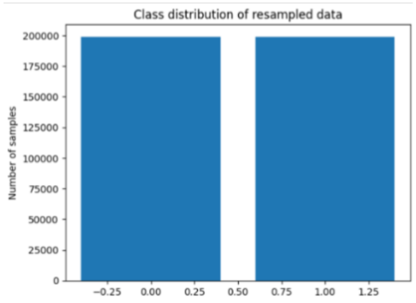
Introduction & Background

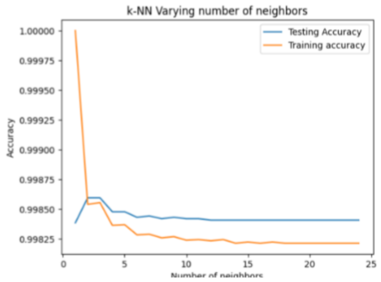
Deep learning for credit card fraud detection entails analysing credit card transactions and identifying possibly fraudulent behaviour using a type of artificial intelligence known as a neural network.

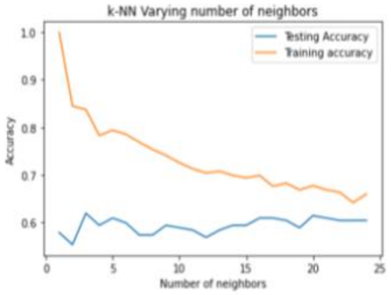
This can be accomplished in a variety of ways, including utilizing unsupervised learning algorithms to discover anomalous patterns in data or supervised learning algorithms to learn from labelled instances of fraudulent and non-fraudulent transactions.

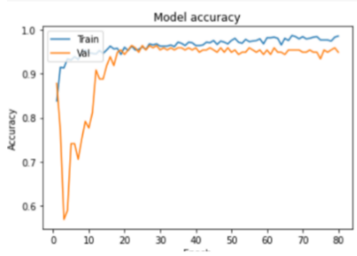
Once trained, the model can be used to automatically flag questionable transactions for additional inspection.

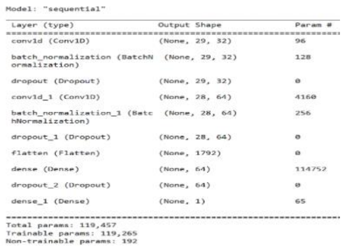
Diagrams











```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 29, 32)            96

 batch_normalization (BatchN  (None, 29, 32)           128
 ormalization)

 dropout (Dropout)           (None, 29, 32)            0

 conv1d_1 (Conv1D)           (None, 28, 64)            4160

 batch_normalization_1 (Batc  (None, 28, 64)           256
 hNormalization)

 dropout_1 (Dropout)         (None, 28, 64)            0

 flatten (Flatten)           (None, 1792)              0

 dense (Dense)               (None, 64)                114752

 dropout_2 (Dropout)         (None, 64)                0

 dense_1 (Dense)             (None, 1)                 65

=================================================================
Total params: 119,457
Trainable params: 119,265
Non-trainable params: 192
```

Methodology

- Downloading the dataset
- Sampling the dataset
- Balancing the data
- Finding the best Hyperparameter tunning
- Training data
- Using different machine learning algorithms
- Displaying results

All algorithms were run twice using different balancing techniques. All models for credit card fraud detection are developed in Python using the TensorFlow and Keras libraries, Once all the modules have been runned they have been compared between each other to ensure that the final model selected for use is the most effective one.

Results

Name	SMOTE accuracy	Undersampling accuracy	Highest Accuracy
K-nearest neighbor	0.998	0.619	SMOTE
Decision Trees	0.946	0.962	Undersampling
Support Vector Machine	0.998	0.883	SMOTE
Random forest	0.999	0.934	SMOTE
Convolutional Neural Network	0.995	0.986	SMOTE

This table clearly shows that SMOTE appears to be the most effective technique, as it produced the highest accuracy for all models. Random forest, which achieved an accuracy of 0.999 using SMOTE, was the model with the highest accuracy. An accuracy of 0.999 (99.9%) is considered very good in most machine learning applications. It means that the model predicted 999 out of 1000 instances correctly or has a 0.1% error rate.

Conclusion & Future Prospects

The primary goal of this project was to improve the accuracy of quickly identifying credit card fraud in order to reduce financial losses. Given that detecting credit card fraud is critical for banks to maintain customer confidence and satisfaction, as well as to mitigate the impact of fraudulent activity, an accuracy of 0.999 (99.9%) is generally regarded as adequate for detecting credit card fraud. However, it is important to note that accuracy alone may not be enough to evaluate the performance of a credit card fraud detection model, and additional metrics such as precision, recall, and F1 score should be considered when developing models in the future. Furthermore, fraud detection models must be constantly updated and improved to keep up with evolving fraud techniques. Overall, credit card fraud detection is a critical process in the financial industry for preventing financial losses and safeguarding customers against fraudulent activity.