PES University, Bengaluru

UE20CS302-MACHINE INTELLIGENCE

TITLE : CREDIT CARD FRAUD DETECTION

NAME : SATISH
BATCH : 02
SRN :PES 2UG20CS550

Literature Review:

Enhanced Credit Card Fraud Detection Model Using Machine Learning

Author: Noor Saleh Alfaiz * and Suliman Mohamed Fati

The approach proposed in the above work is based on two stages of implementation. The first stage is to identify the top three machine learning algorithms from a pool of nine. The second stage seeks to combine the top three algorithms with nineteen different resampling strategies.

The Area under the Receiver Operating Characteristic Curve (AUC), Accuracy, Recall, Precision, and F1-Score are used to evaluate each model in both stages. Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), Nave Bayes (NB), Random Forest (RF), Gradient Boosting Machines (GBM), Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosting (XGBoost), and Category Boosting are the nine methods in the first stage (CatBoost). The 19 resampling strategies are divided as follows in the second stage: There are 11 undersampling techniques, 6 oversampling approaches, and 2 combinations of both undersampling and oversampling techniques.

The total number of models in both stages are 66, with their 330 evaluation metric values that took nearly one month to obtain. The best model out of all these is AllKNN along with CatBoost (AllKNN-CatBoost). Finally, AllKNN-CatBoost is compared with previous works with the same dataset and similar approaches. Indeed, AllKNN-CatBoost outperforms previous models in terms of AUC (97.94%), Recall (95.91%), and F1-Score (87.40%).

Another dataset and different optimization algorithms may be used in the study's future phases. These include the Earthworm Optimization Algorithm (EWA), Elephant Herding Optimization (EHO) [55], Moth Search (MS)

algorithm [56], Slime Mold Algorithm (SMA), and Harris Hawks Optimization (HHO), to name a few.

A Review of Machine Learning Applications for Credit Card Fraud Detection with A Case study

Author: Zahra Faraji

This paper discussed machine learning strategies for detecting credit card fraud. Fraud is a significant issue for financial organisations; consequently, implementing a model that can handle data quickly and efficiently is vital. Data balancing is critical to establishing a stable and generalised method. Knowing the different algorithms can help you make a better judgement on which technique to choose. In this research, five distinct supervised approaches were applied to public Kaggle data, and their performance was evaluated using imbalanced and balanced data. In addition, a new ensemble model was used to increase the performance of the individual classifiers.

This study has some shortcomings, one of which is the data. Because the data for this study was limited to one financial institution, the conclusions cannot be extended to all banks or financial institutions. Future research can look into ML approaches with larger datasets. Another disadvantage is that this study does not employ unsupervised approaches. The new ensemble model can be applied to a set of supervised and unsupervised procedures to evaluate if performance can be improved. Unsupervised machine learning algorithms can be compared to supervised techniques. This study's data consists solely of numerical characteristics. Other sorts of data, such as textual data, can help to improve the fraud detection process.

Fraud detection in credit card transaction using machine learning techniques

Author: Imane Sadgali, Nawal Saei, Fouzia Benabbou.

This study analyses the performance of four supervised machine-learning approaches for credit card fraud detection: decision trees, k-nearest neighbour, random forests, and support vector machines. In our evaluation, we used a single generic data set on credit card transactions to confirm or deny the conclusions of our state of the art, whereby most publications offer findings using a specific dataset. Following that, the performance of each technique was examined for comparison. Support vector machines have proven to be superior to the competition.

It is feasible that in the further this research will investigate the usage of neural networks, as well as other supervised, unsupervised, and reinforcement learning techniques. Our primary goal is to identify the strategies that produce the greatest results in order to incorporate them into our adaptive model of credit card fraud detection.