

Thesis topic Proposal

Application of Machine Learning in Credit Card Fraud Detection

Credit card fraud is a growing problem in the finance world. However, to detect a fraud is not a such an easy task to perform. In order to tackle this problem, for the graduation thesis of my bachelor's degree in Information and Communication Technology, I have chosen the topic of **Application of Machine Learning in Credit Card Fraud Detection**. For this topic, I have chosen a dataset of transactions made by credit cards in September 2013 by European cardholders published on Kaggle, which is highly imbalanced. There are only 492 out of 284807 transactions are labelled as fraudulence. More specifically, I would like to look into how to process and deal with imbalance data using different techniques such as undersampling, oversampling and hybrid sampling. In additional, I would also use different predictive models such as logistic regression, random forest or boosting along with these processing techniques and compared the results between them. I would be using recall, precision, f1-score, ROC curve and precision-recall curve to evaluate each combination of the processing technique and predictive model.