# Abstract

This project implies successful implementation of machine learning algorithms to make better predictions on public dataset. I possess business knowledge of finance, commercial banking, and currently working for loan department. In finance industry, retail & commercial banks have set their business goals to extend credit to their prospects and target high outstanding balances to hit their targets. In this project, bank’s dummy loan data found on Kaggle is used to train model on imbalanced dataset and balanced dataset. Using business expertise on dataset, we would be able to understand how accuracy and performance varies from imbalanced to balanced dataset.

In final report, I have conducted data cleaning, exploratory data analysis, followed by data modelling, evaluation, and its performance.

Keywords: Logistic Regression, Support Vector Machine, Random Forest, Artificial Neutral Network, SMOTE.