Technocolabs Raushan Kumar

Predicting Credit Card Defaults and Deriving Financial Insights

15th October 2020

Overview

This is a data science project that deals with data from a financial company that provide credits. The aim is to propose a Machine Learning model the effectively predicts the whether a customer seeking credit would default or not. The overall job is to effectively identify the credit defaulters and prevent losses to the company.

The data used in this python data science project has been collected by the finance company itself.

Tools & Technologies

Python Pandas Matplotlib Sciket-Learn Jupyter Notebook

Tasks

- A proper analysis of the available credit defaults data.
- Propose an effective ML model
- Summarize the cost caused by credit defaults
- Minimize the credit defaults cost through effective prediction
- Derive the various financial insights

Goals Achieved

- In this data science project, the credit defaults data was analysed.
- Used the various data science tools to manipulate and visualize the data collected.
- An effective ML model using “Random Forest” algorithm was designed.
- Derived financial insights about the company from the data available.
- The model is able to save almost 25-26 % cost of all defaults.

Outcome

- It was observed that at least 30% of the accounts that were in default last month tend to default again next month.
- “Random Forest” turns out to be the best model for our dataset and so used for modelling.
- It is estimated that the cost of credit defaults can be reduced by 25-26% making a huge profit for the company using this model.