

Predicting Credit Card Approvals

Project Description

Commercial banks receive a lot of applications for credit cards. Many of them get rejected for many reasons, like high loan balances, low-income levels, or too many inquiries on an individual's credit report, for example. Manually analyzing these applications is mundane, error-prone, and time-consuming (and time is money!). Luckily, this task can be automated with the power of machine learning and pretty much every commercial bank does so nowadays. In this project, you will build an automatic credit card approval predictor using machine learning techniques, just like the real banks do.

Dataset description:

The dataset used in this project is the Credit Card Approval dataset from the UCI Machine Learning Repository.

<http://archive.ics.uci.edu/ml/datasets/credit+approval>

The features of this dataset have been anonymized to protect the privacy, but this blog http://rstudio-pubs-static.s3.amazonaws.com/73039_9946de135c0a49daa7a0a9eda4a67a72.html gives us a pretty good overview of the probable features. The probable features in a typical credit card application are *Gender, Age, Debt, Married, BankCustomer, EducationLevel, Ethnicity, YearsEmployed, PriorDefault, Employed, CreditScore, DriversLicense, Citizen, ZipCode, Income* and finally the *ApprovalStatus*. This gives us a pretty good starting point, and we can map these features with respect to the columns in the output.

Tasks:

1. First, we will start off by loading and viewing the dataset.
2. We will see that the dataset has a mixture of both numerical and non-numerical features, that it contains values from different ranges, plus that it contains several missing entries.
3. We will have to preprocess the dataset to ensure the machine learning model we choose can make good predictions.
4. After our data is in good shape, we will do some exploratory data analysis to build our intuitions.
5. We will apply all the classification methods that we learn by applying cross validation across them, then pick the best one.
6. Finally, we will build a machine learning model - that can predict if an individual's application for a credit card will be accepted- using the best algorithm and tune its hyperparameters to achieve the best performance.