Data Science Framework Report

Based on Framework One

Kurt Lancaster

Define the Goal

- Goal: To build a predictive model that will allow Credit One to better predict what amount of credit to extend to the customer
  - Why does Credit One need this?
    - Customers have been defaulting as of late
  - What do we need from the company?
    - A dataset that contains customer demographics of people who have defaulted vs not defaulted
  - How will this project be delivered to the company?
    - A report will be located in a GitHub repository per the client's request

Collect & Manage Data

- What data is available?
  - A CSV file with customer 6 month credit history/demographics and default status
    - Features include: credit amt, sex, marital status, age, repayment status, bill and payment amt per month, default status (dependent variable)
  - Will the data help solve the problem? Is it enough?
    - Initial thoughts are that this CSV contains enough quality data to help improve Credit One's underwriting

Building the Model

- Extracting useful insights from the dataset
  - What techniques will be applied?
    - Extracting the data, cleaning the data, and performing initial assessments on what predictive model suits the given data
  - How many techniques will be applied?
    - After transforming the data, the techniques to build the model will be to train the data and use regression to structure the predictive model

Evaluate & Critique the Model

- Ensure that the model ultimately meets Credit One's expectations
  - Is the model accurate enough to satisfy the client?
    - Get the percentage threshold required from Credit One contact
  - Does the predictive model perform better than human intuition?
    - Analyze if the model's algorithm is predicting the obvious
  - Do the model's predictive results make real-world sense?
    - Look for examples of inefficient/garbage model behavior

Present Documents & Results

- A link will be provided to a GitHub repository with our findings
  - How to interpret?
    - Analyze what features have the greatest impact on default status
  - How confident should Credit One be in the model's predictions
    - Based on the accuracy of the model, Credit One will be able to judge how credit worthy it's customers are and how much credit they should extend
  - When should Credit One override the predictions?
    - If Credit One finds that the results are not improving the default rate, they should pause any underwriting based on the model and consider if more analysis is needed

Deploy & Maintain the Model

- Ensuring the model runs smoothly
  - How will the model be handed off to production?
    - After end testing and client satisfaction, the model will be placed into production
  - How will the model be maintained?
    - If the customer decides to run this model full-time, the development team will work in two week sprint increments and deliver a release post sprint to further optimize the model
    - If the model should need to be revised, the development team will roll back the changes and implement fixes to bugs