Interpretable Models for Risk Modeling

Problem Statement

Square is a San Francisco based company that makes software and hardware payments products. The company has expanded into small business services with the creation of a lending arm, Square Capital. Essential to the operations of Square Capital is the extraction of meaningful patterns from financial data. In particular, Square is interested in creating fair and robust models for predicting loan default probabilities based on applicant financial history.

Students are tasked with using the Lending Club data, and supplementary data of choice, to build a model for predicting the loan default probability of a given loan application. The model should be both predictively robust as well as interpretable and fair, wherein interpretability and fairness is defined to be compliance with the letter and spirit of the US Equal Credit Opportunity Act. Specifically, loan decisions based on the model's recommendations must be deemed justifiable and defensible under an FTC audit.

Data Resources Students will use Lending Club data and other appropriate publicly accessible data.

Project Goals

Students will learn how to produce client-facing computational frameworks, pipelines and products. Here, the client perspectives under consideration are four-fold:

- legal and regulatory bodies
- lenders (banks)
- loan applicants (people applying for loans)
- future engineers and data analysts charged with maintaining and updating the model

A good data scientist needs to think about the "interpretability" of their modeling and implementation decisions along all these different axes.

The students will produce:

- A pipeline that takes the Lending Club data set in addition to supplementary data and produce an appropriately formatted output of default probabilities for each application as well as an "explanation" of the model decision in each case.
A scientific report documenting, justifying and analyzing theoretical underpinning of their model, implementation as well as model evaluation. The evaluation of the model must include:

- Results showing performance comparison of the model with reasonable benchmark models.
- Results demonstrating the "fairness" of their model on benchmark "fairness" datasets (TBA). I.e. test the chosen model on data sets wherein one control for everything except for say race, and prove that the model decisions are equitable and that the model "explanations" are sensible.