Springboard Capstone Project 1 Proposal

Project Title: *Porto Seguro's Safe Driver Prediction*

Project Objective: Predicting the probability that a driver will initiate an auto insurance in the next year

Client Background: Porto Seguro, one of Brazil's largest auto and homeowner insurance companies, completely agrees. Inaccuracies in car insurance company's claim predictions raise the cost of insurance for good drivers and reduce the price for bad ones.

Client expectations:

1. A more accurate prediction model that will allow them to further tailor their prices
2. Make auto insurance coverage more accessible to more drivers

Reference: https://www.kaggle.com/c/porto-seguro-safe-driver-prediction

Data: The data used for the project is from taken Kaggle platform. The data contains variables whose features label have no description. Features contain types,

1. Ordinal/Continuous variables
2. Binary variables
3. Categorical variables

Approach: The proposed broad approach to achieve the objective:

1. Understanding data and Data wrangling [1 week]
    a. Missing values
    b. Manipulating data into a workable format
    c. Exploratory data analysis
2. Dimensionality reduction [1 week]
    a. Correlations
    b. PCA
3. Data storytelling and Project Milestone 1 [1 week]
    a. Present my finding
    b. Feedback from mentor and community
4. Machine learning model [2 to 3 weeks]
    a. Initial prediction model
    b. First kaggle submission
    c. Refining the model
5. Final Kaggle submission
6. Final Project report submission

Deliverables:

1. Provide the following on GitHub:
    a. Code
    b. Documentation
    c. Project report
    d. Presentation
2. Personal deliverable: Secure Kaggle top 100