鯨析

Porto Seguro Safe Driver Prediction

PREPARED FOR

MR. ADRIANOMOALA

APRIL 1, 2022

ABOUT

Nothing ruins the thrill of buying a brand new car more quickly than seeing your new insurance bill. The sting's even more painful when you know you're a good driver. It doesn't seem fair that you have to pay so much if you've been cautious on the road for years.

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.

In this competition, you're challenged to build a model that predicts the probability that a driver will initiate an auto insurance claim in the next year. While Porto Seguro has used machine learning for the past 20 years, they're looking to Kaggle's machine learning community to explore new, more powerful methods. A more accurate prediction will allow them to further tailor their prices, and hopefully make auto insurance coverage more accessible to more drivers.

Business Goal

We are assembled to solve this classical challenge for insurance.

Our goal has always been to provide affordable and accessible services for the drivers and company, and we believe in breaking down the barriers that make accurate insurance claim prediction prohibitive for any safe driver.

- **For the driver:** Help drivers reduce unnecessary insurance expenses and increase rational willingness to purchase insurance.
- **For the company:** Further tailor insurance prices to make auto insurance coverage more accessible to more drivers; greatly save data maintenance costs and accurately locate target customer groups.
- **For the future:** Narrow the company's research scope on customer characteristics, expand the pool of elite feature pool, and respond quickly to new customer claims.

Our Team

FOUNDERS PORTO-SEGURO | KAGGLE

SENIOR PARTNERS

Strategy Plan

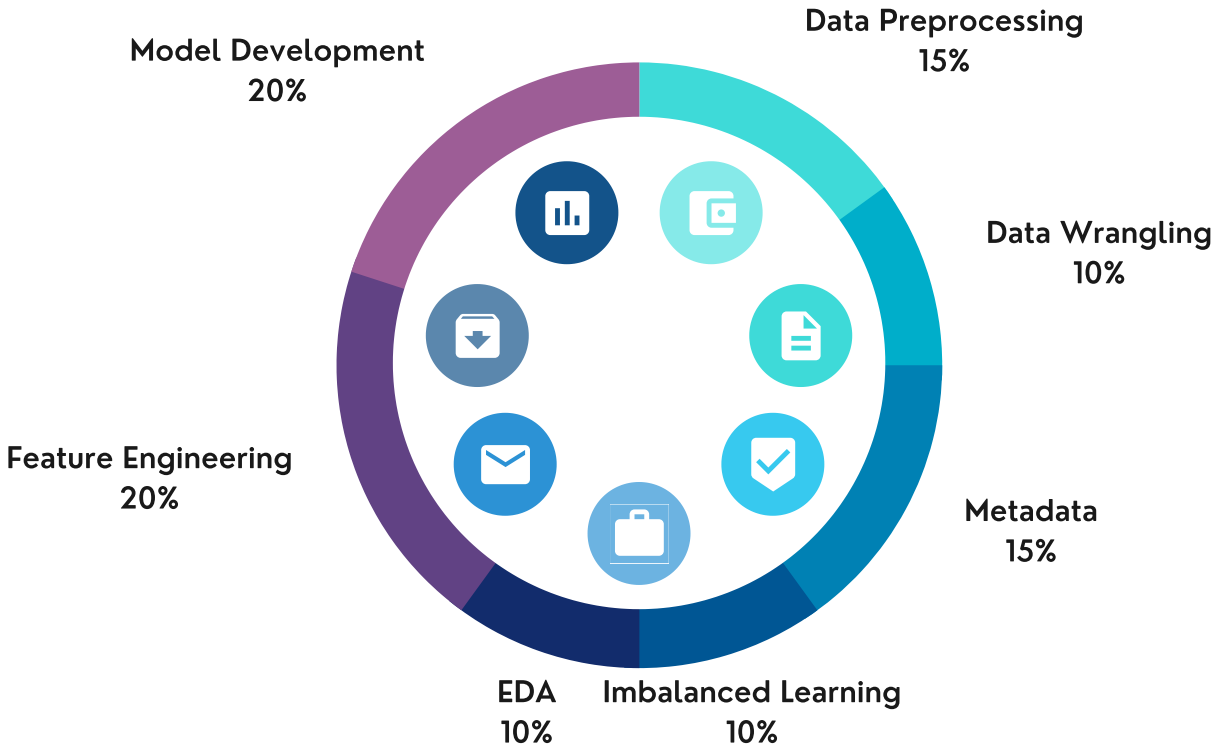
We are now looking towards building a model that predicts the probability that a driver will initiate an auto insurance claim in the next year, which will serve as an efficient and cost-effective way to access more target customers.

Aside from complex model construction, we will also provide exhaustive data preparation processes including but not limited to:

- **Data Preprocessing**
- **Data Wrangling/manipulation**
- **Metadata Framework Construction**
- **Imbalanced Learning Technique**
- **Exploratory Data Analysis**
- **Feature Engineering**

Breakdown

- Data Preprocessing
- Data Wrangling/manipulation
- Metadata Framework Construction
- Imbalanced Learning Technique
- Exploratory Data Analysis
- Feature Engineering
- Model Development



Evaluation

FINAL NORMALIZED GINI COEFFICENT

0.28566

10-FOLD CV

EPOCH 1	0.256629
EPOCH 2	0.305255
EPOCH 3	0.287905
EPOCH 4	0.285242
EPOCH 5	0.294075
EPOCH 6	0.294992
EPOCH 7	0.278625
EPOCH 8	0.277246
EPOCH 9	0.270604
EPOCH 10	0.307598