porto-seguro-safe-driver-prediction

**The abstract should make it clear what your goal is, what resources (data, software, papers) you will use, and what we should expect to see at the end of the project.**

Goal:

To build a model that predicts whether an individual will claim for an auto insurance in the next year

Resources:

# Data : Po Porto Seguro’s Safe Driver Prediction Dataset from Kaggle competition

Software : Python 3.6 using Spyder, Tableau for Data visualization

Papers:

Expected results : To achieve maximum possible accuracy in classifying the auto insurance unseen data.

**Project Flow:**

**Business understanding** – The goal of our project is to predict whether an individual will claim an auto insurance in the next year.

**Data understanding** – The aim here is to get a good insight of the data set provided. This involves visual inspection of data and designing metadata.

**Data preparation** – In this step, data quality checks are performed and exploratory data visualization are carried out.

**Modelling evaluation** – A suitable data classification algorithm is developed to build a binary classification model.

**Evaluation** – K-fold cross-validation is performed on the data set to finetune the model.