Airlines Customer Satisfaction-Classification

Abstract:

Customer satisfaction metric is a vital element for businesses. Airlines companies need to measure this level of satisfaction in order to control and improve it, thus airlines can achieve a higher quality level that meets customer expectations.

Introduction:

Our company is data Science (DS) is a data science company specializing in data analytics and data science. We provide breakthrough and effective solutions and help clients to achieve their own goals through the utilization of data to improving performance, building forecasting models for better decisions making and solutions.

Problem Statement:

An airline is a company that provides air transport services for traveling passengers and freight. In this project, DS company will analyze the data airline customer satisfaction data of the flights provide actionable insights that will enable them to improve satisfaction. The goal of this project is to determine what factors lead to customer satisfaction for an Airline.

Data Description

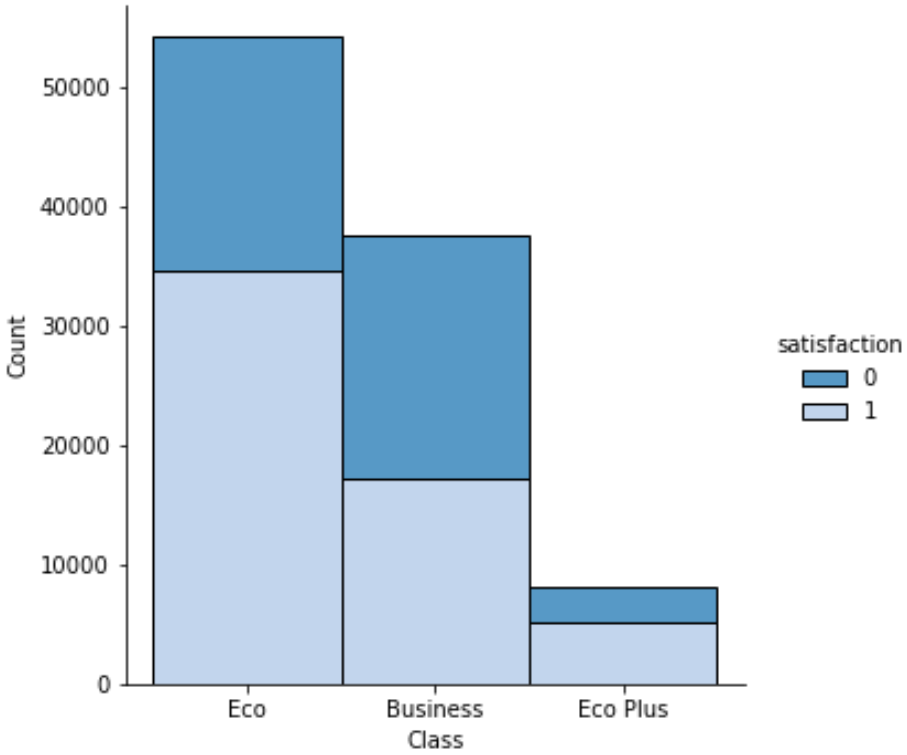
We will be using the " analyze the data airline customer satisfaction data " Dataset from Kaggle: the dataset has 129880 numbers of rows and 23 numbers of feature

Tools:

Technologies: python, Jupyter notebook

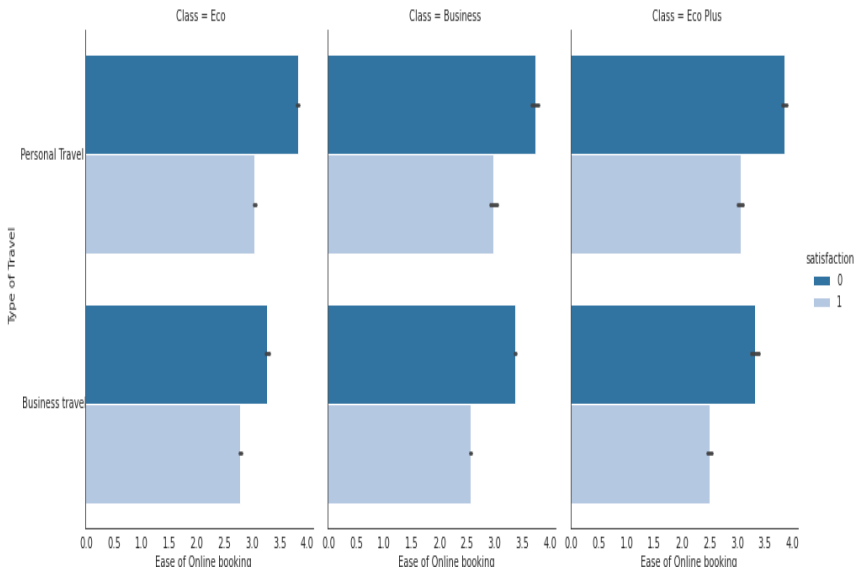
Libraries: pandas, numpy, matplotlib, seaborn, LogisticRegression , sklearn.linear_model, sklearn.model_selection, sklearn.preprocessing

1-Are most passengers satisfied, dissatisfied or neutral in general for the airline?

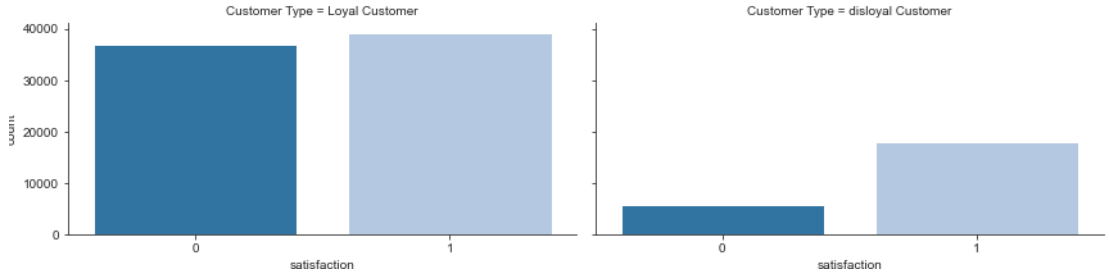


Eco class have higher dissatisfied than satisfied, and business class have higher satisfied than dissatisfied, and Eco plus class have higher dissatisfied than satisfied.

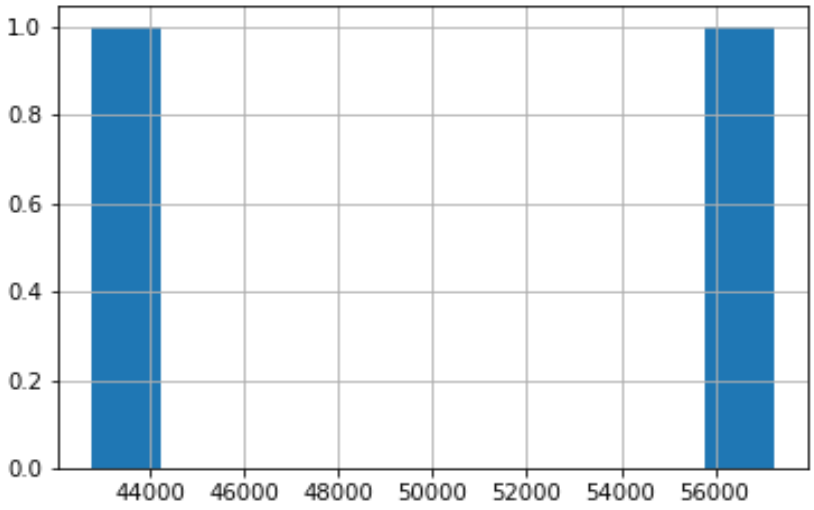
2- What type of class was more satisfied passengers?



satisfied passengers were more in which class loyal customers or disloyal customers?



Data Balance:

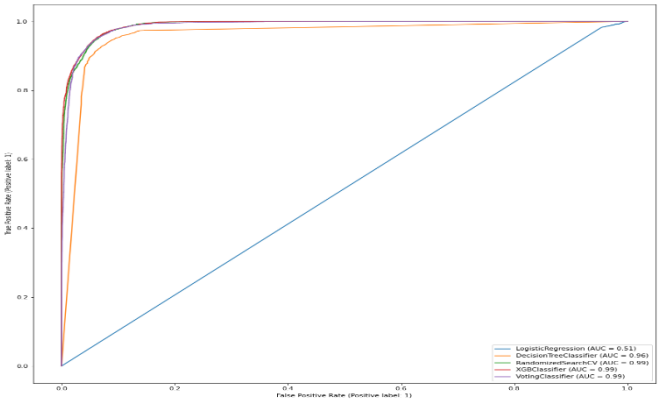


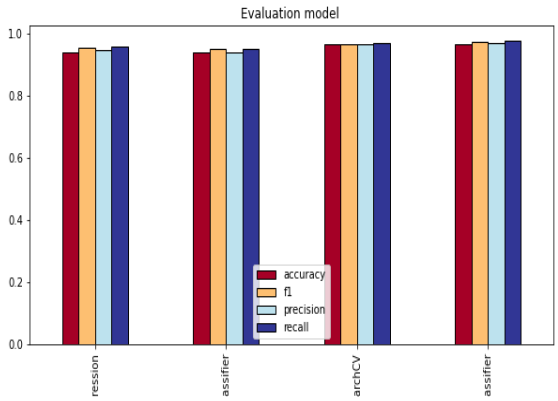
Logistic Regression:

- **Train : 0.77**
- **Test: 0.77**

This model isn't Oversampling or Under sampling since the accuracy between train and test data is just similar.

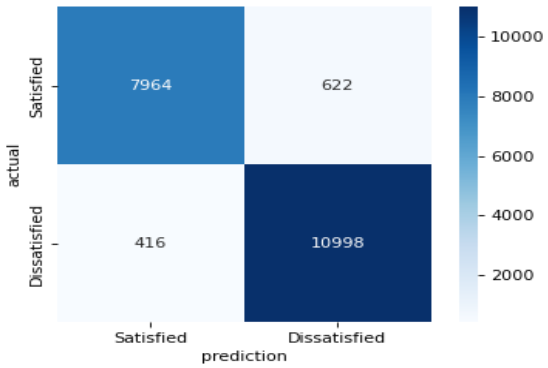
ROC CURVE:

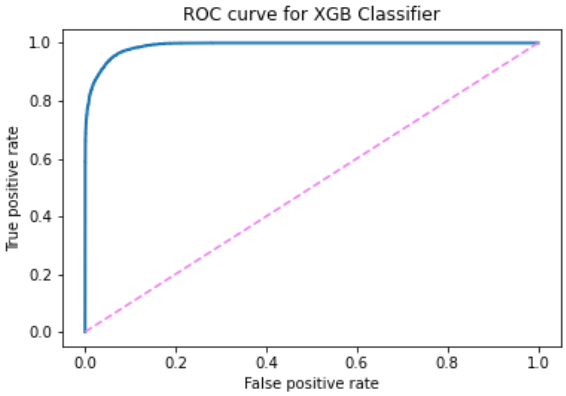




This graph explains all the model and choose the best model as a result the best model is XGB classifier.

Best Models:





Machine Learning to analyze customer satisfaction, we find that **XGB Classifier** is the best machine learning model to predict our customer satisfaction data.