Airlines Customer Satisfaction-Classification

By :
Rana Alturki
Reem Binzeraiban

2R | DS
Data Science

Introduction

- **The objective of performing this project:** Assist an airline firm in determining the key aspects that influence customer satisfaction.

- **Data Description:** The data is obtain form “Analyze the data airline customer satisfaction” dataset in Kaggle.

100000 Record 23 Columns

Target : satisfaction

1: dissatisfied
0: satisfied

Methodology

01

Gathering Data

02

Exploratory Data Analysis

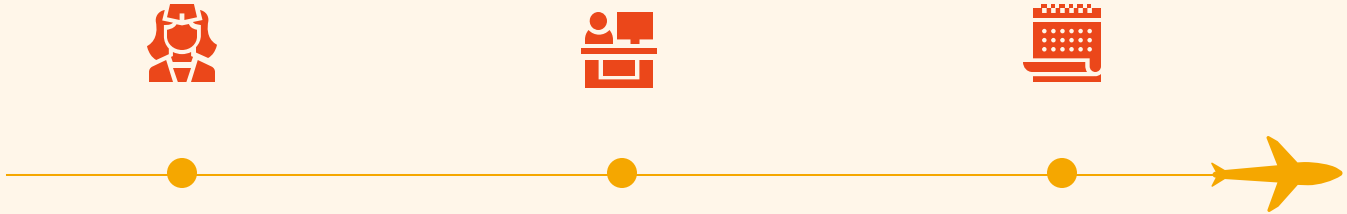
03

Classification Models

04

Evaluations

Data Preparation



Data Cleaning

- Find and drop null values

- No duplicated

Splitting Data

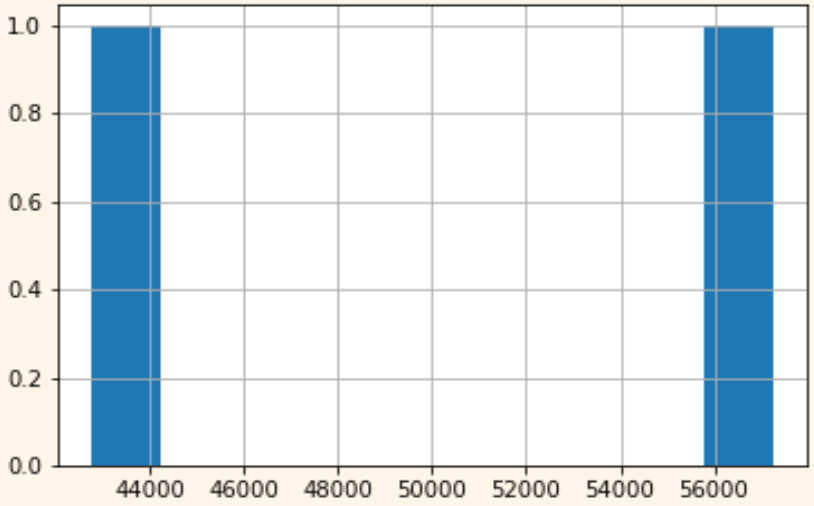
- Train : 80

- Test : 20

Feature Engineering

Dummy Variables

Data Balance



Logistic Regression:

Train : 0.77

Test: 0.77

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

Logistic Regression

Random Forest

Decision Tree

MODELLING

XGB Classifier

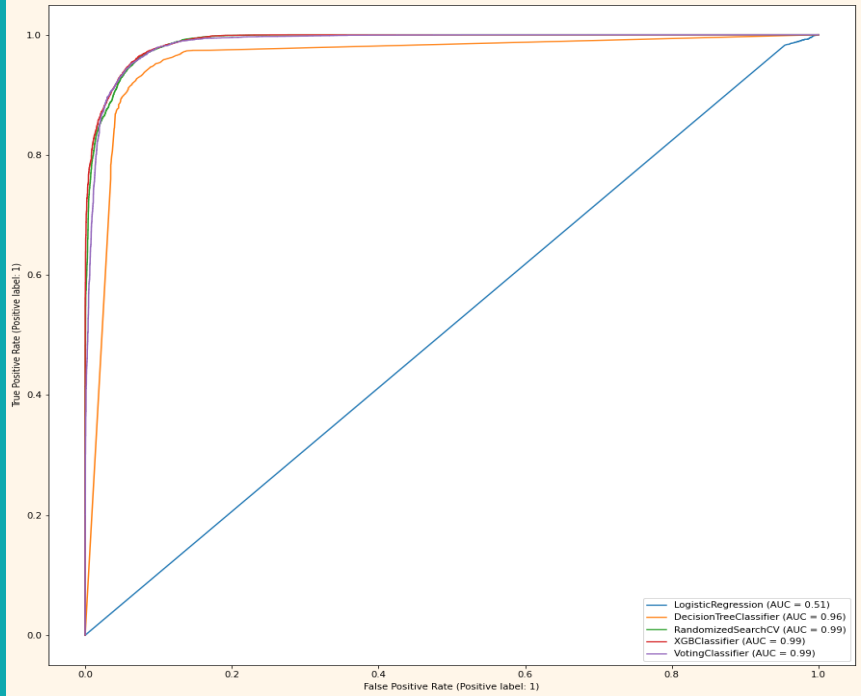
Voting

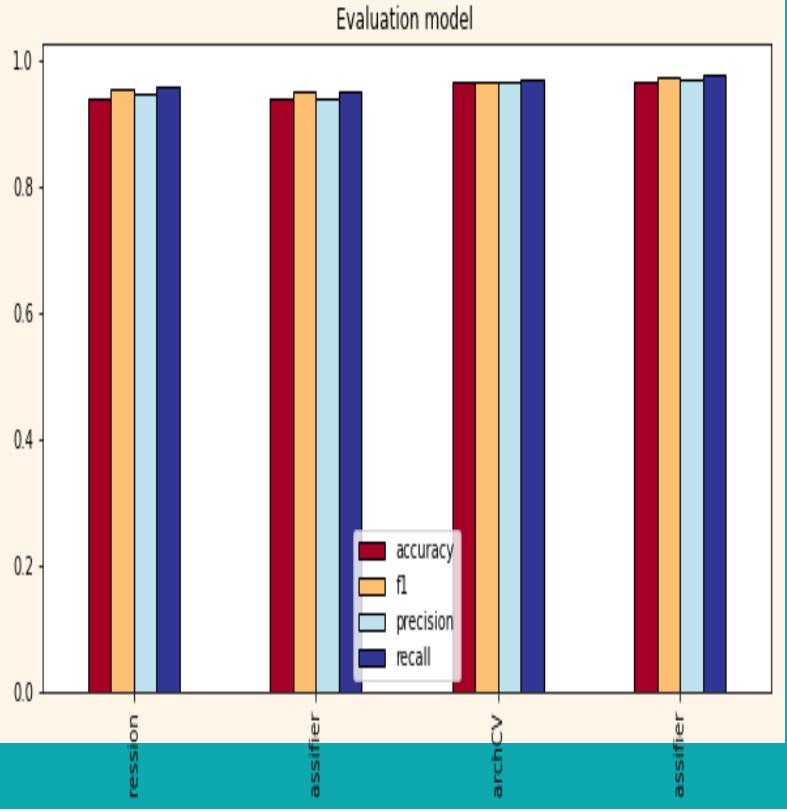
Stacking Classifier

EVALUATING

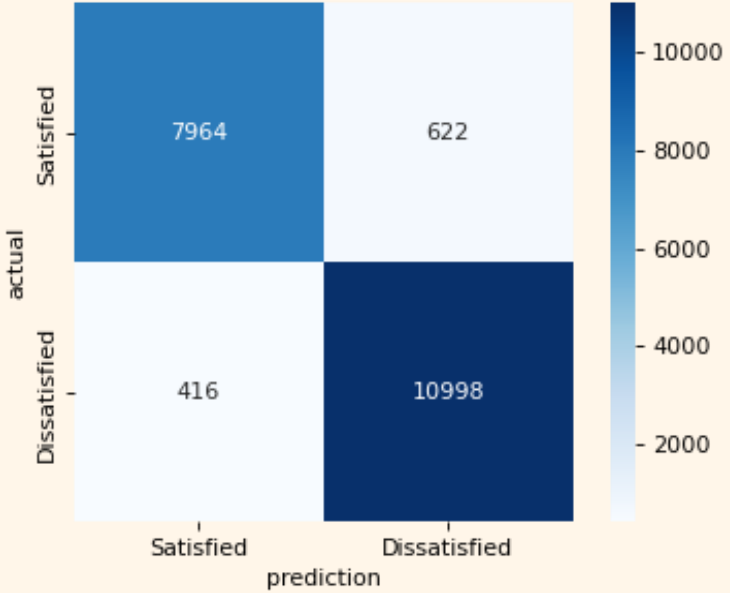
Model	Train	Test	Precision	Recal	Accuracy	F-1
Logistic Regression Basline	0.78	0.78	0.80	0.81	0.78	0.80
Standard Scaler	0.81	0.80	0.82	0.85	0.80	0.83
Decision Tree	0.96	0.93	0.93	0.94	0.93	0.93
Random Forest	0.96	0.94	0.93	0.94	0.93	0.93
XGB Classifier	0.95	0.94	0.94	0.96	0.94	**0.95**

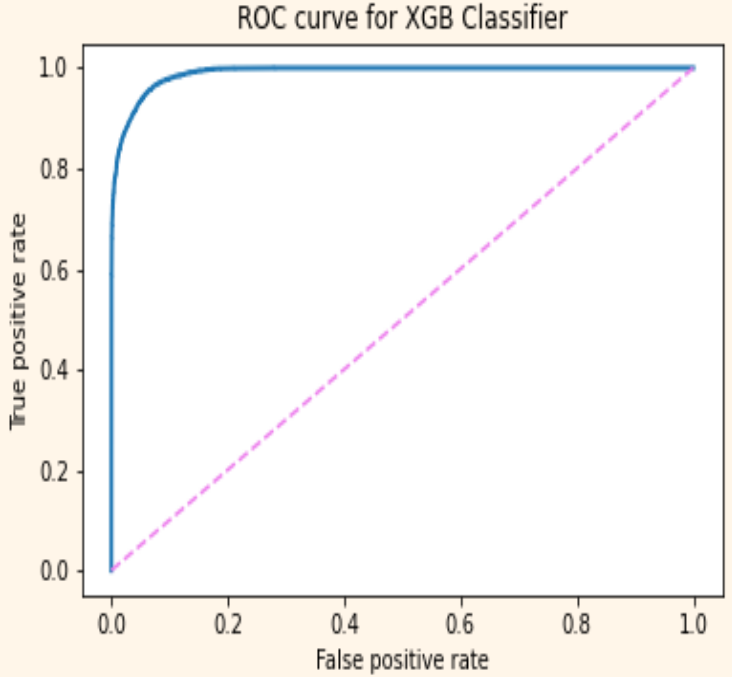
ROC CURVE


True Positive Rate (Positive label: 1)
False Positive Rate (Positive label: 1)
LogisticRegression (AUC = 0.51)
DecisionTreeClassifier (AUC = 0.96)
RandomizedSearchCV (AUC = 0.99)
XGBClassifier (AUC = 0.99)
VotingClassifier (AUC = 0.99)


Evaluation model
accuracy
f1
precision
recall
ression
assifier
archCV
assifier

Confusion matrix





Conclusion

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

Thank you!

We hope you enjoy it !