THE FINAL REPORT

Done by :Sara Alqhtani

Submit to: Dr.Essam Al Daoud

Abstract

Data consists of the details of customers in an airline company who have already flown with them. The main goal of this dataset is to forecast whether a future customer will be satisfied with their service and which aspects of their services should be highlighted more to increase customer satisfaction. Data consists of total 129880 observations (train data:103904, test data:25976) and 25 columns. By building the logistic regression model we find the model fit the data correctly and can predict the future customers satisfaction.

Design

The dataset from https://www.kaggle.com/yagmuruzun/airline-passenger-satisfaction-eda-ml/data for airline company customer satisfaction regarding the services provided. Logistic regression model used to forecast the future customer satisfaction in case of improving their services.

Data

Data consists of total 129880 observations (train data:103904, test data:25976) and 25 columns. Data contain numeric and object type features. The target column consists of two categories “satisfied” and “neutral or dissatisfied”.

Algorithms

- Converting categorical features to numeric variables.
- Delete the gender, customer type, type of travel, class.
- Using logistic regression model.

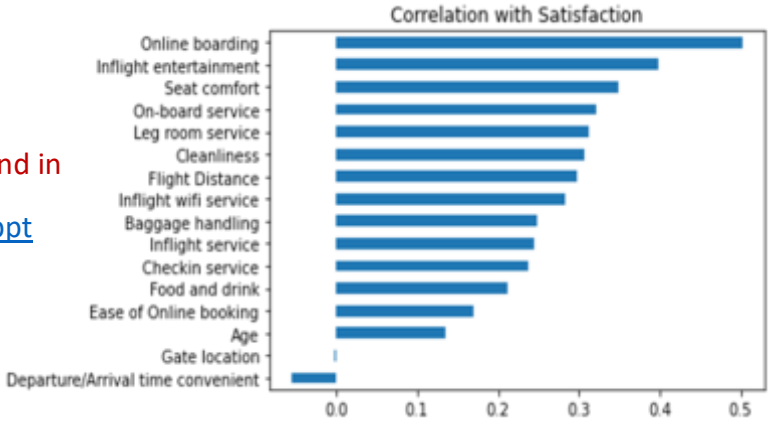
Tools

- Numpy
- Pandas
- Matplotlib
- Seaborn
- Sklearn

Communication

The presentation of the project can be found in

https://github.com/SaraAlghtani/project-ppt



	Mean_CV	Std_CV	Train_Score	Test_Score	Precision_Score	Recall_Score	F1_Score
Logistic Regression	0.872872	0.00206	0.872894	0.873755	0.868855	0.836679	0.852464