FLIGHT PRICE PREDICTION

DETAILED PROJECT REPORT

1. DATA COLLECTION
2. DATA CLEANING
3. FEATURE ENGINEERING
4. FEATURE SELECTION
5. HANDLING CATEGORICAL VARIABLES
6. TRAIN-TEST-SPLIT
7. FEATURE SCALING
8. MODEL TRAINING
9. MODEL TESTING
10. MODEL DEPLOYEMENT

DETAILED PROJECT DESCRIPTION

➤ **DATA COLLECTION :** The data is collected as per our requirement.The model is built based on the data choosed.Data is collected and stored for further exploration.

➤ **DATA CLEANING:** The gathered data is not clean for model building, so the data is to be cleaned i.e,various operations are to be done on the gathered data to get insights from it.

➤ **FEATURE ENGINEERING:** All the features have different properties with respect to data.The feature engineering step deals with different operations like finding null values,dealing with outliers etc.. This makes the data more flexible to our model building.

➤ **FEATURE SELECTION:** Among all the features available in the data, only certain features are required for the data.Those features are selected and rest are ignored in feature selection step.

➤ **HANDLING CATEGORICAL VARIABLES:** All the features selected may not be numerical variables.So the categorical variables are converted to numerical variables using diffrent techniques like one hot encoding,label encoding etc..

➢ **TRAIN-TEST SPLIT:** Nowthat all the variables are numerical our dataset is divided to train and test datasets.The training is given to our model and the model is tested with our test data set.

➢ **FEATURE SCALING:** All the variables we have now are of different scales.in feature scaling all the features are scaled down to single range.

➢ **MODEL TRAINING:** Now that our model is trained on the given data.Based on machine learning algorithm we train our model.

➢ **MODEL TESTING :** The trained model is tested and performance metrics is checked for the best performance.

➢ **DEPLOYEMENT:** The final model is deployed on cloud platforms as per our requirements.

➢First we cleaned our dataset properly by removing all null values and duplicate value present in dataset.

➢Then we performed feature extraction, in which I extracted journey date, month and departure and arrival hour,minutes in new separate column.

➢After that I performed feature engineering step in which I created one new feature "Total_Duration". In this feature

➢what I have done is , I converted total time in minutes.

➢Then I performed feature selection step in which I dropped some feature like(Route, Date_of_journey,

➢Departure_time, arrival_time and Additional_Info).

➢Then I handled categorical variable by performing One-Hot encoding.

➢Then I split the hole data set train-test split. After that I performed scaling on X_train and X_test.

Training and Testing :

➢ In this step, I trained my dataset on different Regression Learning algorithm(Linear, Random-Forest, K-NN,DecisionTreeRegression, SVR, Ridge, Lasso and Elastic net).

➢ After that I applied hyper-parameter tuning on all model which I have described above.

Deployement :

➢ First i have saved my model which we have trained and tested above , in pickle file format for model deployment.

➢ For API building we are using flask in backend.

➢ We are deploying this model on various cloud platform like Azure, AWS and heroku.

➢ We also Dockerized this model on Docker.

➢ For version control, we are using

OUTPUT

FLIGHT PRICE PREDICTION

Departure Date

dd-mm-yyyy --:-- --

Arrival Date

dd-mm-yyyy --:-- --

Source

Delhi ▾

Destination

Cochin ▾

Stopage

Non-Stop ▾

Airlines

Jet Airways ▾

Submit