
# Flight Price Prediction

Submitted by:

Sagnik Das

# ACKNOWLEDGMENT

*I would like to express my special thanks of gratitude to (Datatrained) also to my SME (Mr Keshav Bansal). Who gave me the golden opportunity to do this wonderful project on the topic (Flight Price Prediction), which also helped me in doing a lot of Research and I came to know about so many new things I am really thankful to tall of them.*

# INTRODUCTION

- ## Business Problem Framing

  This business problem helps us to understand the trend of flight ticket price. Also, the factors on which the price depends.

- ## Conceptual Background of the Domain Problem

  This problem is divided into 4 steps:-

  a. Data scraping from the websites.
  b. EDA Analysis, on train and test dataset.
  c. Hyperparameter optimization tuning, on train dataset.
  d. Predicting the value on test dataset

- ## Review of Literature

  This problem helps us to understand the factors on which the flight price depends

- ## Motivation for the Problem Undertaken

  To find out the factors on which the value of a flight ticket can increase or decrease.

# Analytical Problem Framing

- Mathematical/ Analytical Modeling of the Problem

  Linear Regresion

$$Y = a + bX$$

- Data Sources and their formats

  Websites – Makemytrip, Yatra, Skyscanner, Goibibo, Via, etc.

- Data Preprocessing Done

- Web Scrapping
- Changed strings into integers
- Hyperparameter optimization using grid search cv
- Validate

- Data Inputs- Logic- Output Relationships

  Input – Website data

  Output – Flight price Prediction.

- State the set of assumptions (if any) related to the problem under consideration

  Here, you can describe any presumptions taken by you.

- Hardware and Software Requirements and Tools Used

- Hardware – Laptop
- Software – import pandas as pd, import os, import csv, import sklearn, import numpy as np, import matplotlib.pyplot as plt, import seaborn as sns, %matplotlib inline, Hyperparameter optimization using grid search cv.

# Model/s Development and Evaluation

- Identification of possible problem-solving approaches (methods)

  EDA Method, Linear Regression and prediction, Hyperparameter optimization using grid search cv.

- Testing of Identified Approaches (Algorithms)

  EDA Method, Hyperparameter optimization using grid search cv.

- Run and Evaluate selected models

  EDA Algorithm, Linear Regression, prediction, Hyperparameter optimization using grid search cv.

- Key Metrics for success in solving problem under consideration

  What were the key metrics used along with justification for using it? You may also include statistical metrics used if any.

- Visualizations

  Contour, Heat map, pair plot, box plot, histogram, gridsearchcv.

  If different platforms were used, mention that as well.

- Interpretation of the Results

  This problem helps us to identify that flight ticket price depends on dates, days (weekday or weekend), departure time, duration, layovers.

# CONCLUSION

- Key Findings and Conclusions of the Study
- This problem helps us to identify that flight ticket price depends on dates, days (weekday or weekend), departure time, duration, layovers, departure city, destination and also on the number of available flights.
- Learning Outcomes of the Study in respect of Data Science

  The graphs showed us the factors on which the flight price depends.

- Limitations of this work and Scope for Future Work

  More data required for better prediction.