**Flight Fare Prediction**

1. **About the Project:**

The name of the project is Flight Fare Prediction.

The goal is to build a Machine Learning model which predicts Flight Fare based on the features.

Features are : Airline, Date of Journey, Source, Destination, Route, Departure Time, Arrival Time, Duration, Total Stops, Additional Info and Price.

1. **Size of the Data:**

This Dataset has 10 thousand records and 11 Features.

1. **Data Type:**

This dataset has 10 object type datatype and the last column (‘Price’) is of integer data type.

1. **Techniques used for Data Pre-Processing:**

* Null values present in the columns were dropped.
* Unecessary columns like ‘Route’ and ‘Additional Info’ are dropped as they didn’t contribute to the Prediction.
* Combined a the same type of Airlines in the Airline column.
* As ML models are not capable of understanding object datatype, we created new features from the Date\_of\_Journey column – Journey\_Date, Journey\_Month, Journey \_year.
* ‘Toatal\_Stops’ column is also og object datatype, so performed label encoding on this columns. Before Label Encoding – ‘non-stop’, ‘2 stops’, ‘1 stops’, ‘3 stops’ and ‘4 stops’. After Label Encoding – 4, 1, 0, 2, 3.
* ‘Dept\_Time’ is of object data type, so we create 2 new integer columns from it – ‘Dep\_Hour’ and ‘Dep\_Min’.
* Similarly, ‘Arrival\_Time’ is of object data type, so we create 2 new integer columns from it – ‘Arrival\_Hour’ and ‘Arrival\_Min’.
* On categorical columns, OneHotEncoding is performed.

1. **Model Training**

* The Model is trained using RandomForestRegressor and XGBoostRegressor.
* Hyperparameter Tuning is also performed on these two Regressors.
* XgboostRegressor comes out as the winner with a accuracy of 77%.
* After HyperParameter Tuning, accuracy is increased by 5% (i.e., 82%).

1. **Deployment:**

* The Front-End of the Model is done using Stramlit, HTML and CSS.
* The model is successfully deployed on Heroku.