PROJECT TITLE

Flight Fare Prediction Using Analysis of Flight Input Features through ML.

ABSTRACT

As domestic air travel in India is becoming increasingly popular with different air ticket booking channels coming online these days, passengers are trying to understand how these airline companies make decisions over time about ticket prices. Therefore, many methods are ready to provide the proper time to do so. The customer who buys an air ticket by estimating the price of the airfare is recently proposed. The majority of these strategies make use of sophisticated Computational Intelligence Prediction Models an area of science known as Machine Learning (ML), thus here we will go through how we can train a model that will help us to predict prices of air tickets.

INTRODUCTION

These days, domestic air travel is becoming more and more common in India. Travelers are trying to learn how these airline companies make choices over time about ticket prices with multiple air travel booking outlets coming online. For a passenger, it is a time-consuming method to search websites for deals and offers. The cost can therefore depend on various variables. This venture uses AI to show the types of flight tickets after some time to estimate the costs.
The ultimate goal of the airways is to achieve profit, while the customer is looking for the minimum cost. Usually, consumers try to book the ticket well in advance of the departure date to prevent airfare hikes as the date gets closer. But that's not the truth, really. By giving more than they should for the same seat, the customer can finish up.
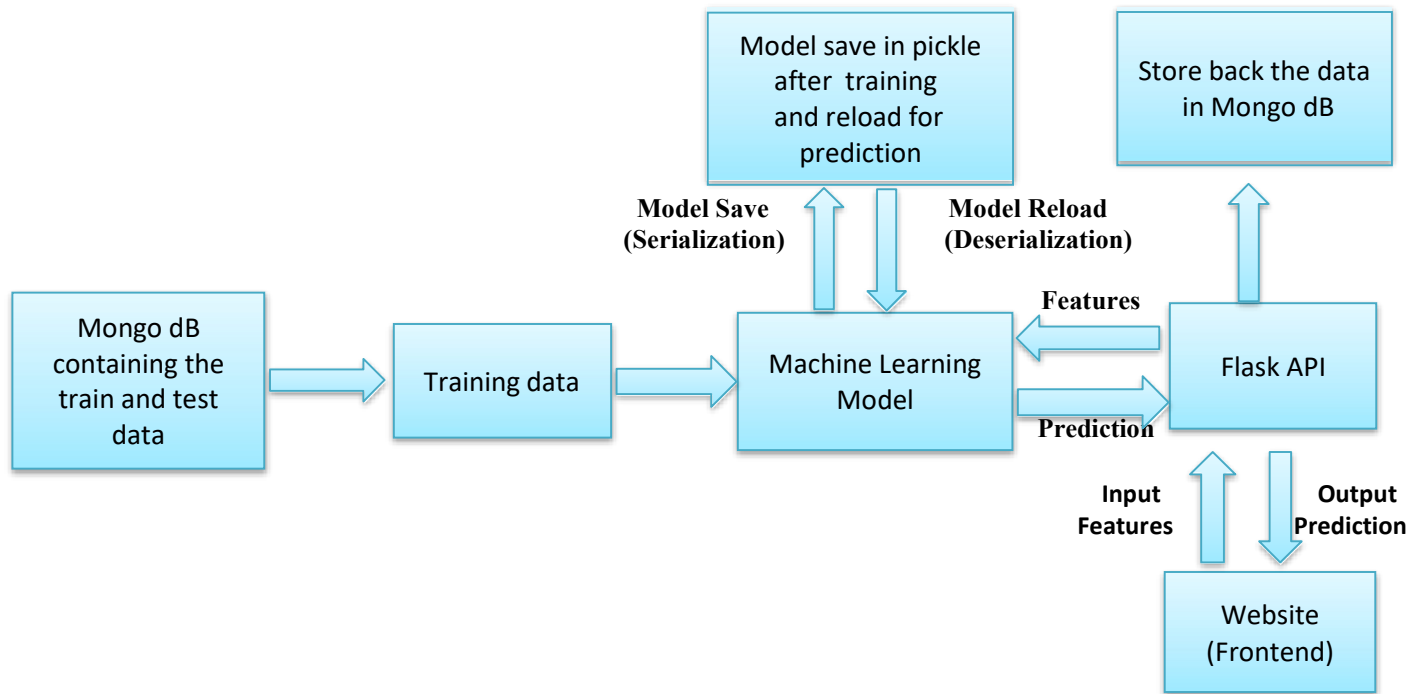
Black Box Diagram:-

Fig No:1 -> The above diagram depicts the block diagram of the entire workflow of the flight fare Prediction using the ML model end to end.

OBJECTIVES

Flight ticket prices can be something hard to guess. We have been provided with prices of flight tickets for various airlines between the months of March and June of 2019 and between various cities, using which we aim to build a model which predicts the prices of the flights using various input features.

Also, we will be storing our flight datasets in the MongoDB which is a non-relational document database that provides support for the JSON-like storage. It is one of the best databases for machine learning for several reasons. It provides flexible data models, powerful query language, stores and retrieves the trained models as JSON documents etc.

Our objective is also to deploy our project on live server (cloud) by which any user can access this Service from anywhere. To fulfill the objective, we are doing following things:

- To build a prediction model we are using "Machine Learning (ML) algorithm on dataset".

- To process a new data in dataset we are using "Unsupervised Machine Learning".
- To provide user access we are creating UI by using "Django, Bootstrap".
- Finally, we are storing our predicted model in the MongoDB.

DATA COLLECTION

Here, we collected the dataset on flight fare prediction from Kaggle. Below is the link for the same:

https://www.kaggle.com/nikhilmittal/flight-fare-prediction-mh/

We have 2 datasets here — training set and test set. The training set contains the features, along with the prices of the flights. It contains 10683 records, 10 input features and 1 output column — 'Price'.
The test set contains 2671 records and 10 input features. The output 'Price' column needs to be predicted in this set. We will use Regression techniques here, since the predicted output will be a continuous value.
Following is the features available in the dataset –
- Airline
- Date_of_Journey
- Source
- Destination
- Route
- Dep_Time
- Arrival_Time
- Duration
- Total_Stops
- Additional_Info
- Price

Cleaning and Preparing Data

The dataset which we get is usually in raw format that is it includes missing values and null attributes, categorical variables which need to be encoded in the format which machine understands, outliers which needs to be detected if any and removed etc. Through this process of cleaning and preparing the data, all unnecessary data is removed. This breakthrough is the most critical and time. Different mathematical methods and logics clean and customize knowledge in python. For example, the extracting date and time.

Analyzing the Data

Data preparation is monitored by breaking down the data, exposing the hidden patterns and applying various regression models afterwards. Similarly, from the existing features, a few features can be calculated. Flight days can be given by measuring the difference between the date of the flight and the date of collection of the details. In addition, the flight date, whether on a festive day or a weekday or a weekend, is significant. The flights scheduled during the weekends instinctively cost more than the flights on weekdays. In addition, time plays a major role
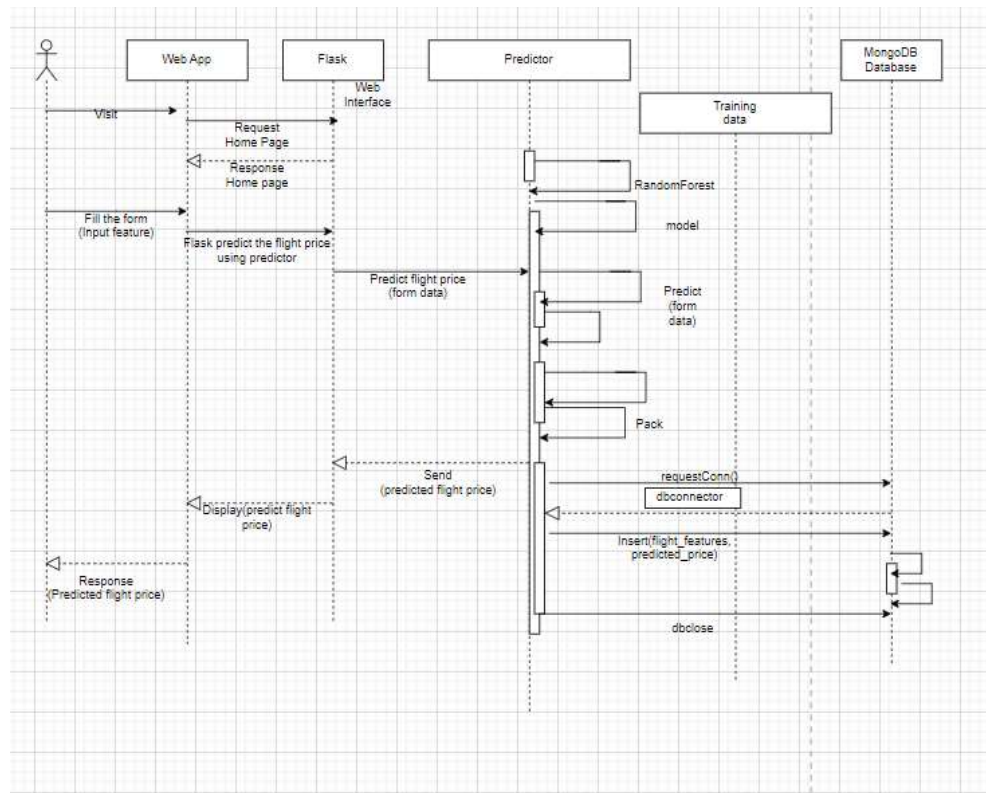
Workflow diagram or sequence diagram: -

Fig No:2 -> The above diagram depicts the sequence diagram of the entire workflow of the flight fare Prediction using the ML model end to end.

CONCLUSION

With the help of our project the travelers can find out the right time to buy their tickets at the lowest cost and also can plan accordingly. So, to get result with maximum accuracy regression analysis is used. From the studies, the feature that influences the price ticket are to be considered. In future the details about number of available seats can also improve the performance of the model.