Flight price prediction using IBM watson

Submitted by:

KODATI PAVAN SRIRAM

(20BCI7054)

VEEDHULA YASHASWI SAI

(20BCD7029)

BUDHARAJU SARATH CHANDRA VARMA

(20BCR7033)

KANMARLAPUDI SATVIKA

(20BCD7034)

Table of contents:

S.NO	TOPIC	PAGE NO.
1	Introduction	3
2	Literature Survey	3
3	Theoretical analysis	4
4	Experimental Investigations	4
5	Flowcharts	5
6	Results	6
7	Advantages and Disadvantages	11
8	Applications	12
9	Future Scope	12

Introduction:

Air travel has become an integral part of our lives, Right from the business trips, vacations to visiting loved ones we use Air travel a lot. However, the constantly changing prices of flight tickets can make it challenging to plan the budget effectively.

The aim of this project is to harness the power of Data analysis and Machine learning models to predict the flight prices to an extent accurately. By using previous history of flight ticket price, type of airline, Duration of the travel, flights available and some other relevant factors, We can develop a model that can predict the future flight prices.

We used different analysis to find out the factors that influence the flight ticket prices, providing us to make smarter choices when it comes to booking your flights.

Literature Survey:

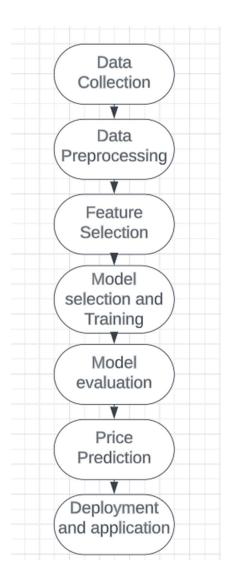
We read a couple of research papers regarding this project and browsed on the internet. We analyzed problems that contribute to the dynamic flight ticket price changing, such as the fuel prices, competition and market demand. Also the factors such as seat availability, booking time, travel seasons we often notice that prices during the peak travel times goes up because the demand for the tickets is more and competition is also high, Understanding this price volatility is much important for prediction of ticket price.

We have to also check for the data limitations we have, we can only predict the price of the ticket based on the previous history of prices, In reality the situation can be different. Addressing the issues related to price volatility, dynamic pricing, seasonality, profit or financial management and data quality, Considering all these we tried to develop a robust prediction regression model that can provide more reliable results.

Theoretical analysis:

We used multiple platforms Visual studio, Google colab and used flask library in python to deploy the project for use.

BLOCK DIAGRAM:



Experimental Investigations:

The first step is Defining the problem:

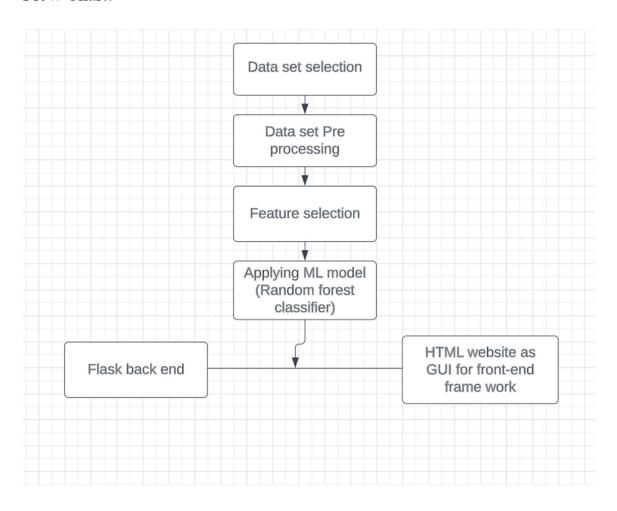
The price prediction depends upon many factors and some of them like fuel cost vary daily and some of them vary based on the duration of travel and some factors like season makes it difficult to predict the price accurately. We evaluate all these factors and developed a Machine learning model using Random forest classifier algorithm to predict the price.

Data Collection and preprocessing is the second step:

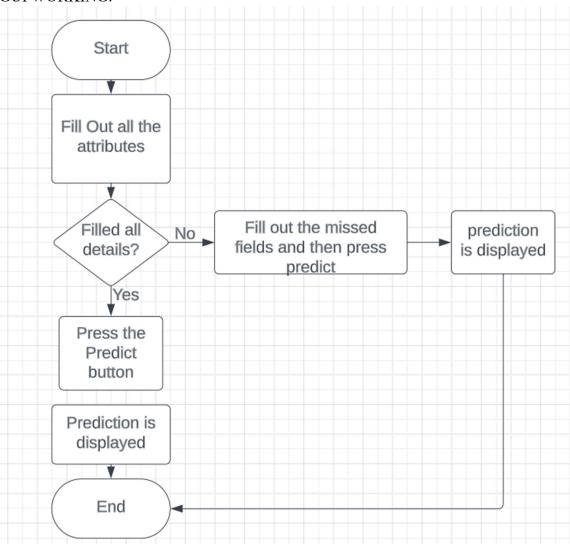
We took flight price dataset and applied pre processing techniques to make sure the dataset is free from null values and other inappropriate formats. We used Heatmap, K-best and feature map selection to find out which feature in the provided dataset will provide best relation with the target variable.

We applied Random forest classifier algorithm to find out the best feature to predict the price correctly. We used performance metrics like MSE, RMSE, MAE to assess the accuracy the model is producing.

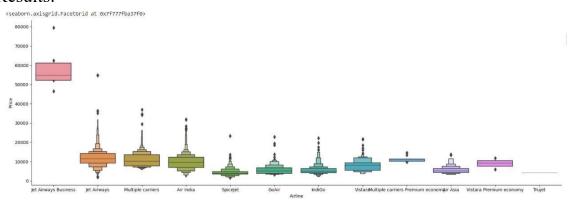
Flow Chart:



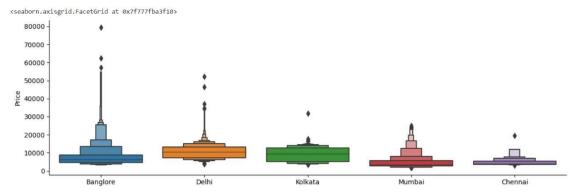
GUI WORKING:



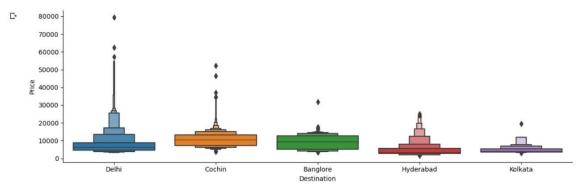
Results:



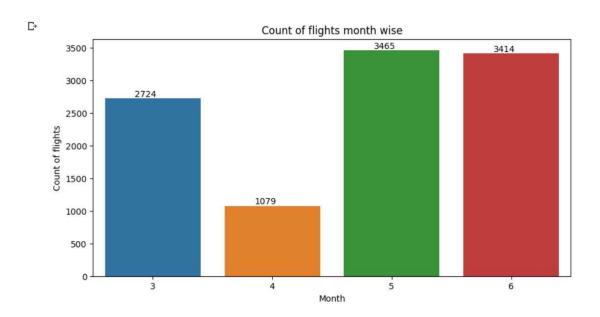
we can observe that jet airways bussiness is he costliest airways



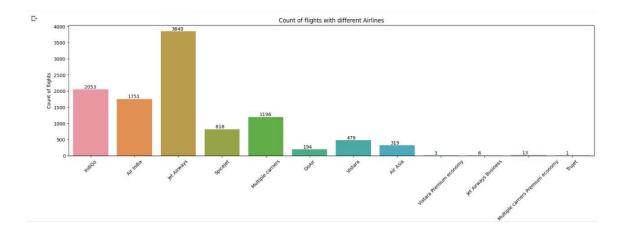
If we are going from Banglore the prices are slightly higher as compared to other cities



If we are going to New Delhi the prices are slightly higher as compared to other cities



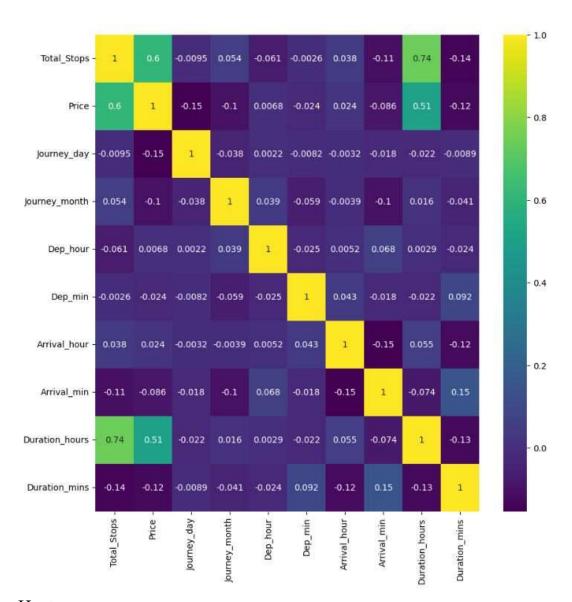
we can observe that month 5 has the highest count of flights as compared to other months



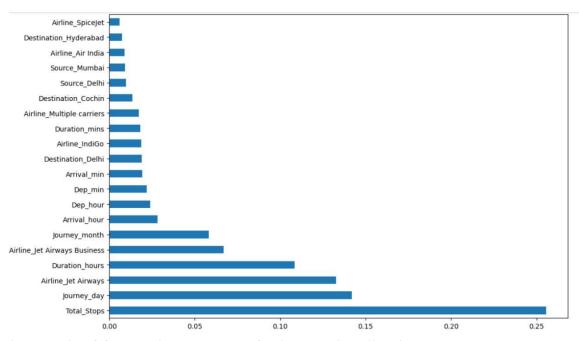
we can observe that the jet airways run the hightest number of flights



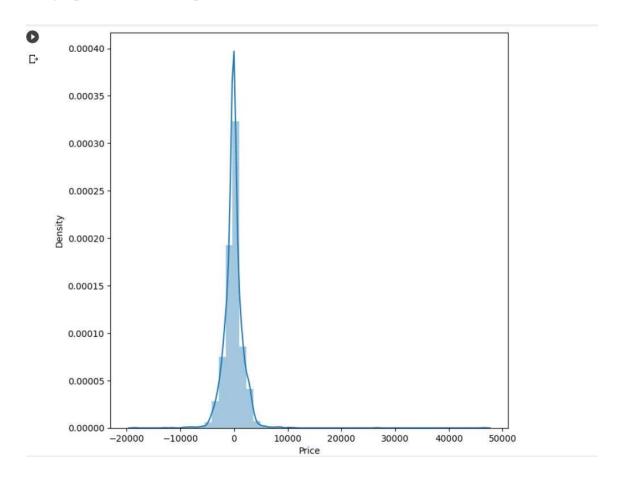
acc to the data, price is directly prop to the no. of stops

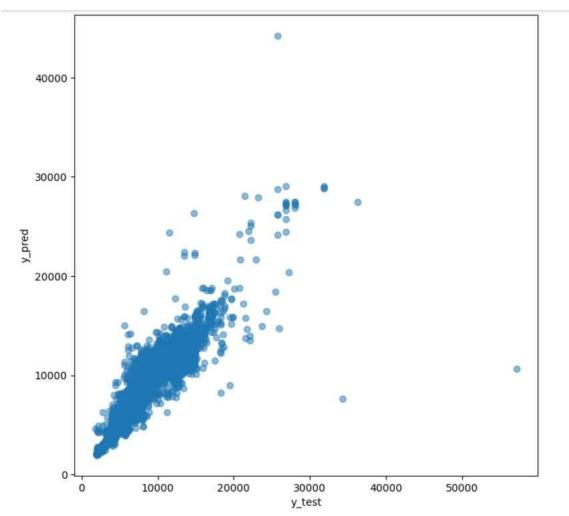


Heatmap



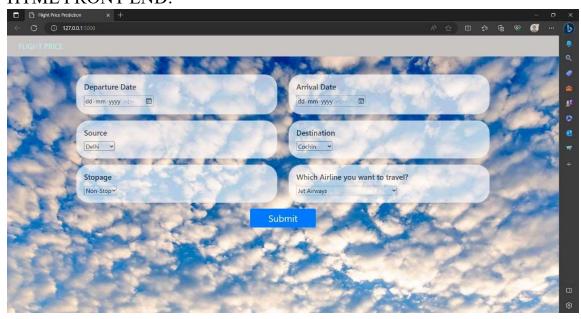
lot graph of feature importances for better visualization

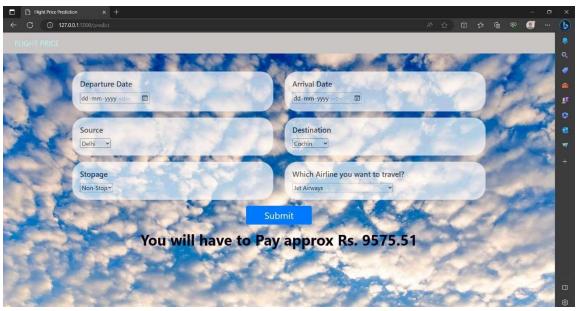




last two are for random forest ml model

HTML FRONT END:





PREDICTION:

r2 score: 0.8113485163969612

Advantages:

- 1. Cost Savings: Accurate flight predictions can help the passengers to get the tickets at a lower price.
- 2. Planning: With the help of price trends, travellers can plan their trips better, selecting the best cost effective date of travel and airlines etc..,
- 3. Insights to Dynamic pricing: Analyzing historical flight data, valuable insights can be gained into the factors influencing flight prices, such as seasonality, demand patterns and other market dynamics.

Disadvantages:

- 1. We may be able to predict the traavelling charges but not the overall cost which may include baggage costs etc.,
- 2. Due to dynamic functioning of the airline industry it becomes difficult to predict the price of the ticket accurately
- 3. Flight ticket prices also depends upon the nature and natural disasters which are entirely unpredictable
- 4. Obtaining Flaw free, reliable flight price history is not an easy task and error in the data can make the prediction go wrong

Applications:

Predicting flight ticket prices has applications in travel industry and beyond:

- 1. Travel planning websites such as MakeMyTrip(MMT), Trivago, etc.., allows users to search for flights and receive real-time predictions n future price trends. This helps travellers to select the most cost effective option.
- 2. Price alert notifications can be generated using this data and enables the users to take advantage of favourable prices and plan their trips accordingly
- 3. Travel insurance companies can leverage flight price prediction to determine the coverage and pricing for trip cancellation or trip interruption policies
- 4. Predictions can be used to determine the future prices enabling us to compare different airline prices over different times and plan the trip accordingly
- 5. Flight price prediction models generate valuable insights into market trends and pricing dynamics. Researchers and industry analysts can utilize this information to conduct market research, identify pricing patterns, and understand the factors driving price fluctuations in the airline industry.

Conclusion:

In conclusion, flight price prediction projects offer valuable insights and benefits both travellers and the travel industry. By using historical flight data and employing Machine Learning algorithms, these projects aim to provide accurate predictions of future flight prices.

For travellers, it enables them to plan their trips better and choose cost effective prices.

However they can also sometimes face challenges as the weather, Dynamic nature of prices makes it difficult to predict the prices.

Overcoming these challenges requires careful collection of data and preprocessing, robust model development and training, and continuous refinement of the prediction algorithms.

Despite these challenges, flight price prediction has practical applications in travel planning websites and apps, booking platforms, Travel insurance and etc..,

Overall, the flight price prediction has the power to revolutionize the way travellers plan and book flights, offering them more cost-saving options.

With the ongoing advancements in data Analytics and machine learning techniques, the accuracy and applicability of flight price prediction is expected to improve further.

Future Scope:

The flight price prediction continues to evolve, and there is so much scope for further advancement in this area. Some of them are:

- 1. Integration of Real Time Data:

 Incorporating real time data such as fuel prices, weather conditions, market trends and etc.., can further result in accurate results in predicting flight prices.
- 2. Ensemble Methods:

Exploring ensemble techniques such as combining multiple algorithms can result in more accurate predictions and robustness of flight price predictions.

- 3. Incorporating Customer preferences:
 - By Considering the merging of Customer preferences and historical booking data we can personalize the flight price predictions, by considering individual travel booking history, layover preferences, class of travel, etc.., we can offer more tailored recomendations.
- 4. Advancement of machine learning techniques: With the enhancement of algorithms we would enable us to analyze complex relationships among data resulting in higher accuracy.
- 5. Sentiment Analysis: Analyzing social media data and incorporating sentiment analysis and finding out the customers opinion on a particular airline and integrating such data can help us in identifying the factors that influence pricing decisions.
- 6. Dynamic pricing Strategies: Integrating dynamic pricing strategies, such as demand based or personalized pricing, into flight prediction models can enable airlines to optimize their revenue generation.
- 7. Collaboration with Industry Stakeholders: Collaboration with airlines, travel agencies and other industry stakeholders can provide access to proprietary data and domain expertise, leading to more accurate and relevant prediction models.

Bibliography:

- 1. https://medium.com/@skillcate/flight-fare-prediction-machine-learning-project-bc7363e6d9eb
- 2. https://www.ijser.in/archives/v10i11/SE221105023044.pdf
- 3. https://www.analyticsvidhya.com/blog/2022/01/flight-fare-prediction-using-machine-learning/

CODE:

```
<html lang="en">
  <style>
   body {
    background-image: url('/static/img/bg5.jpg');
   background-repeat: no-repeat;
   background-attachment: fixed;
   background-size: cover;
}

img {
   display: block;
   margin-left: auto;
```

```
margin-right: auto;
     align-self: center;
  }
  .navbar {
     background-color: #c9c5c5;
  a {
     color: #d2f8f8;
  a:hover {
     color: #604f4f;
  .card {
     border-radius: 1rem;
     background-color: rgb(241, 12, 127);
<html lang="en">
<style>
  body {
     background-image: url('/static/img/bg5.jpg');
     background-repeat: no-repeat;
     background-attachment: fixed;
     background-size: cover;
  }
  img {
     display: block;
     margin-left: auto;
     margin-right: auto;
     align-self: center;
  }
  .navbar {
     background-color: #c9c5c5;
  a {
     color: #d2f8f8;
  }
  a:hover {
     color: #604f4f;
```

```
.card {
    border-radius: 1rem;
    background-color: rgb(241, 12, 127);
  }
</style>
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Flight Price Prediction</title>
  <!-- BootStrap -->
  link
                                                                     rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.0/css/bootstrap.min.css"
    integrity="sha384-
9aIt2nRpC12Uk9gS9baDl411NQApFmC26EwAOH8WgZl5MYYxFfc+NcPb1dKGj7Sk
" crossorigin="anonymous">
</head>
<body>
  <nav class="navbar navbar-inverse navbar-fixed-top">
    <div class="container-fluid">
       <div class="navbar-header">
         <a class="navbar-brand" href="/" style='color: #b7f8f5;'>FLIGHT PRICE</a>
       </div>
    </div>
  </nav><div class="container my-5">
    <form action="\predict" method="post">
       <div class="row my-3">
         <div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
              <div class="card-body">
                <h5 class="card-title">Departure Date</h5>
                <input type="datetime-local" name="Dep Time" id="Dep Time"
required="required">
              </div>
           </div>
         </div>
```

```
<div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
             <div class="card-body">
                <h5 class="card-title">Arrival Date</h5>
                <input type="datetime-local" name="Arrival Time" id="Arrival Time"
required="required">
             </div>
           </div>
         </div>
      </div>
<div class="row my-3">
         <div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
             <div class="card-body">
                <h5 class="card-title">Source</h5>
                <select name="Source" id="Source" required="required">
                  <option value="Delhi">Delhi</option>
                  <option value="Kolkata">Kolkata
                  <option value="Banglore">Banglore
                  <option value="Mumbai">Mumbai
                  <option value="Chennai">Chennai
                </select>
             </div>
           </div>
         </div>
         <div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
             <div class="card-body">
                <h5 class="card-title">Destination</h5>
                <select name="Destination" id="Destination" required="required">
                  <option value="Cochin">Cochin</option>
                  <option value="Delhi">Delhi</option>
                  <option value="Banglore">Banglore
                  <option value="Hyderabad">Hyderabad</option>
                  <option value="Kolkata">Kolkata
                </select>
             </div>
           </div>
         </div>
      </div>
<div class="row my-3">
         <div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
```

```
<div class="card-body">
               <h5 class="card-title">Stopage</h5>
               <select name="stops" required="required">
                  <option value="0">Non-Stop</option>
                 <option value="1">1</option>
                 <option value="2">2</option>
                 <option value="3">3</option>
                  <option value="4">4</option>
               </select>
             </div>
           </div>
         </div>
         <div class="col-sm-6">
           <div class="card" style='border-radius: 2.15rem;opacity: 0.7;'>
             <div class="card-body">
               <h5 class="card-title">Which Airline you want to travel?</h5>
               <select name="airline" id="airline" required="required">
                 <option value="Jet Airways">Jet Airways
                 <option value="IndiGo">IndiGo</option>
                 <option value="Air India">Air India
                  <option value="Multiple carriers">Multiple carriers
                 <option value="SpiceJet">SpiceJet</option>
                 <option value="Vistara">Vistara
                 <option value="Air Asia">Air Asia
                 <option value="GoAir">GoAir</option>
                 <option value="Multiple carriers Premium economy">Multiple
carriers Premium economy
</option>
                             value="Jet
                                                        Business">Jet
                 <option
                                           Airways
                                                                         Airways
Business</option>
                 <option value="Vistara Premium economy">Vistara Premium
economy</option>
                  <option value="Trujet">Trujet</option>
               </select>
             </div>
           </div>
         </div>
      </div>
      <div style='text-align:center'>
         <button type="submit" value='Submit' class="btn btn-primary px-5"</pre>
style="font-size: 25px;">Submit</button>
      </div>
    </form>
```

```
<h1 style='text-align: center;color: rgb(122, 43, 234);'><b>{{ predictions}
}}</b></h1>
    <footer class='text-light bg-dark position-fixed fixed-bottom '>
       </footer>
  </div>
<!-- JavaScript -->
  <script src="https://code.jquery.com/jquery-3.5.1.slim.min.js"</pre>
    integrity="sha384-
DfXdz2htPH0lsSSs5nCTpuj/zy4C+OGpamoFVy38MVBnE+IbbVYUew+OrCXaRkfj"
    crossorigin="anonymous"></script>
  <script src="https://cdn.jsdelivr.net/npm/popper.js@1.16.0/dist/umd/popper.min.js"</pre>
    integrity="sha384-
Q6E9RHvbIyZFJoft+2mJbHaEWldlvI9IOYy5n3zV9zzTtmI3UksdQRVvoxMfooAo"
    crossorigin="anonymous"></script>
  <script src="https://stackpath.bootstrapcdn.com/bootstrap/4.5.0/js/bootstrap.min.js"</pre>
    integrity="sha384-
OgVRvuATP1z7JjHLkuOU7Xw704+h835Lr+6QL9UvYjZE3Ipu6Tp75j7Bh/kR0JKI"
    crossorigin="anonymous"></script>
</body>
</html>
FLASK FILE:
from flask import Flask, request, render template
from flask cors import cross origin
import pickle
import pandas as pd
model = pickle.load(open('flight rf.pkl','rb'))
app = Flask( name )
@app.route('/')
@cross origin()
def home():
return render template('home.html')
@app.route('/predict',methods=['GET','POST'])
@cross origin()
def predict():
```

```
if request.method=='POST':
    dep time = request.form['Dep Time']
    Journey day = pd.to datetime(dep time,format="%Y-%m-%dT%H:%M").day
    Journey month = pd.to datetime(dep time,format="%Y-%m-%dT%H:%M").month
    Departure hour = pd.to datetime(dep time,format="%Y-%m-%dT%H:%M").hour
    Departure min = pd.to datetime(dep time,format="%Y-%m-%dT%H:%M").minute
    arrival time = request.form['Arrival Time']
    Arrival hour = pd.to datetime(arrival time,format="%Y-\%m-\%dT\%H:\%M").hour
    Arrival min
                                       pd.to datetime(arrival time,format="%Y-%m-
%dT%H:%M").minute
    Total stops = int(request.form['stops'])
    dur hour = abs(Arrival hour-Departure hour)
    dur min = abs(Arrival min-Departure min)
 airline=request.form['airline']
    if(airline=='Jet Airways'):
       Jet Airways = 1
      IndiGo = 0
      Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
       Jet Airways Business = 0
       Vistara Premium economy = 0
       Trujet = 0
    elif (airline=='IndiGo'):
       Jet Airways = 0
      IndiGo = 1
      Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
      Jet Airways Business = 0
       Vistara Premium economy = 0
      Trujet = 0
elif (airline=='Air India'):
      Jet Airways = 0
      IndiGo = 0
```

Air India = 1

```
Multiple carriers = 0
  SpiceJet = 0
  Vistara = 0
  GoAir = 0
  Multiple carriers Premium economy = 0
  Jet_Airways_Business = 0
  Vistara Premium economy = 0
  Trujet = 0
elif (airline=='Multiple carriers'):
  Jet Airways = 0
  IndiGo = 0
  Air India = 0
  Multiple carriers = 1
  SpiceJet = 0
  Vistara = 0
  GoAir = 0
  Multiple carriers Premium economy = 0
  Jet Airways Business = 0
  Vistara Premium economy = 0
  Trujet = 0
   elif (airline=='SpiceJet'):
  Jet Airways = 0
  IndiGo = 0
  Air India = 0
  Multiple carriers = 0
  SpiceJet = 1
  Vistara = 0
  GoAir = 0
  Multiple carriers Premium economy = 0
  Jet Airways Business = 0
  Vistara Premium economy = 0
  Trujet = 0
elif (airline=='Vistara'):
  Jet Airways = 0
  IndiGo = 0
  Air India = 0
  Multiple carriers = 0
  SpiceJet = 0
  Vistara = 1
  GoAir = 0
  Multiple carriers Premium economy = 0
  Jet Airways Business = 0
  Vistara Premium economy = 0
  Trujet = 0
elif (airline=='GoAir'):
```

```
Jet Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 1
       Multiple carriers Premium economy = 0Jet Airways Business = 0
       Vistara Premium economy = 0
       Trujet = 0
    elif (airline=='Multiple carriers Premium economy'):
       Jet Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 1
       Jet Airways Business = 0
       Vistara Premium economy = 0
       Trujet = 0
    elif (airline=='Jet Airways Business'):
       Jet Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
       Jet Airways Business = 1
       Vistara Premium economy = 0
       Trujet = 0
elif (airline=='Vistara Premium economy'):
       Jet Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
       Jet Airways Business = 0
       Vistara Premium economy = 1
       Trujet = 0
```

```
elif (airline=='Trujet'):
       Jet Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
       Jet Airways Business = 0
       Vistara_Premium_economy = 0
       Trujet = 1
else:
       Jet_Airways = 0
       IndiGo = 0
       Air India = 0
       Multiple carriers = 0
       SpiceJet = 0
       Vistara = 0
       GoAir = 0
       Multiple carriers Premium economy = 0
       Jet Airways Business = 0
       Vistara Premium economy = 0
       Trujet = 0
    Source = request.form["Source"]
    if (Source == 'Delhi'):
       s Delhi = 1
       s Kolkata = 0
       s Mumbai = 0
       s Chennai = 0
    elif (Source == 'Kolkata'):
       s Delhi = 0
       s Kolkata = 1
       s Mumbai = 0
       s Chennai = 0
    elif (Source == 'Mumbai'):
       s Delhi = 0
       s Kolkata = 0
       s Mumbai = 1
       s Chennai = 0 elif (Source == 'Chennai'):
       s Delhi = 0
       s Kolkata = 0
       s Mumbai = 0
       s Chennai = 1
```

```
else:
  s Delhi = 0
  s Kolkata = 0
  s Mumbai = 0
  s_Chennai = 0
Destination = request.form["Destination"]
if (Destination == 'Cochin'):
  d Cochin = 1
  d Delhi = 0
  d Hyderabad = 0
  d Kolkata = 0
elif (Destination == 'Delhi'):
  d Cochin = 0
  d Delhi = 1
  d Hyderabad = 0
  d Kolkata = 0
elif (Destination == 'Hyderabad'):
  d Cochin = 0
  d Delhi = 0
  d Hyderabad = 1
  d Kolkata = 0 elif (Source == 'Chennai'):
  s Delhi = 0
  s Kolkata = 0
  s Mumbai = 0
  s Chennai = 1
else:
  s Delhi = 0
  s Kolkata = 0
  s_Mumbai = 0
  s Chennai = 0
Destination = request.form["Destination"]
if (Destination == 'Cochin'):
  d Cochin = 1
  d Delhi = 0
  d Hyderabad = 0
  d Kolkata = 0
elif (Destination == 'Delhi'):
  d Cochin = 0
  d Delhi = 1
```

```
d Hyderabad = 0
       d Kolkata = 0
    elif (Destination == 'Hyderabad'):
       d Cochin = 0
       d_Delhi = 0
       d Hyderabad = 1
       d Kolkata = 0
elif (Destination == 'Kolkata'):
       d Cochin = 0
       d Delhi = 0
       d_Hyderabad = 0
       d Kolkata = 1
    else:#Banglore
       d Cochin = 0
       d Delhi = 0
       d Hyderabad = 0
       d Kolkata = 0
    output = model.predict([[Total stops,
       Journey_day,
       Journey month,
       Departure hour,
       Departure min,
       Arrival hour,
       Arrival min,
       dur hour,
       dur min,
       Air India,
       GoAir,
       IndiGo,
       Jet Airways,
       Jet Airways Business,
       Multiple_carriers,
       Multiple carriers Premium economy,
       SpiceJet,
       Trujet,
Vistara,
       Vistara Premium economy,
       s Chennai,
       s Delhi,
       s Kolkata,
       s Mumbai,
       d Cochin,
       d Delhi,
       d Hyderabad,
       d Kolkata]])
```

```
output = round(output[0],2)
    return render_template('home.html',predictions='You will have to Pay approx Rs.
{}'.format(output))

if _name_ == '_main_':
    app.run(debug=True)
```