iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** X **Month of publication:** October 2022

DOI: https://doi.org/10.22214/ijraset.2022.47022

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

Airline Ticket Price Prediction

Dr. B. Sekhar Babu[1], Kuppala Mahi Venkat[2], Janga Manoj Reddy[3], Katragadda Rathna Sai[4]

Abstract: ***A mode of transportation is an airline. It is essential for both domestic and international travel that takes place quickly. They are relatively expensive when compared to other forms of transportation. But our comfort level and journey time will be directly impacted by this price. When more capacity is needed to meet demand, the airline may raise the prices. Data for a certain air route, including parameters like departure time, arrival time, and airways for a specific period, has been gathered to determine the minimum airfare.***

Keywords: RANDOM FOREST, Mean Squared Error, Root mean square error, Mean absolute Error

I. INTRODUCTION

A crucial role is played by airline price prediction. Today, predicting the cost of any form of transportation has a significant impact. The cost of travel is significantly reduced by price prediction utilizing a computer algorithm. Our project's major goal is to present the consumer a forecast price that is very near to the real cost of the ticket. Additionally, it displays different pricing tiers for various businesses offering varying amenities.

II. LITERATURE SURVEY

It is quite difficult to anticipate the projected price of that day's tickets with any degree of accuracy. However, machine learning models will assist us in predicting the costs of flights from source to destination, allowing users to purchase tickets at a discount and save some money. The cost is lower when a user purchases an airline ticket far enough in advance of the departure date. By creating machine learning models utilising various machine learning algorithms, such as Random Forest, MAE, MSE, and RSME, among others, these outcomes can be attained. The machine learning model is created by taking into account and evaluating the data that was gathered in prior years, and the model will produce the best outcomes.

III. MACHINE LEARNING

Machine Learning is a subset of Artificial Intelligence which helps software models to become much more accurate without being explicitly programmed to do so. Various machine learning models are developed by various entities for the forecasting of weather. These models use various algorithms to obtain the desired solution toa given problem.

For our model we have used:

1) *Random Forest Algorithm:* Random Forest may be a classifier that contains variety of decision trees on various subsets of the given dataset and takes the typical to boost the predictive accuracy of that dataset.
2) *Mean Squared Error:* The mean squared error (MSE) tells you the way close a curve is to a collection of points. this can be accomplished by squaring the distances between the points and also the curve (also referred to as the "errors").The squaring is critical to get rid of any negative signs. It also gives more weight to larger differences. It's called the mean squared error as you're finding the common of a collection of errors. The lower the MSE, the higher the forecast.
   MSE formula = (1/n) * Σ(actual – forecast)2
3) *Mean absolute Error:* The mean absolute error (MAE) may be a measure of errors between paired observations expressing the identical phenomenon.

*Mean Absolute Error = (1/n) * $\sum|y_i - x_i|$*

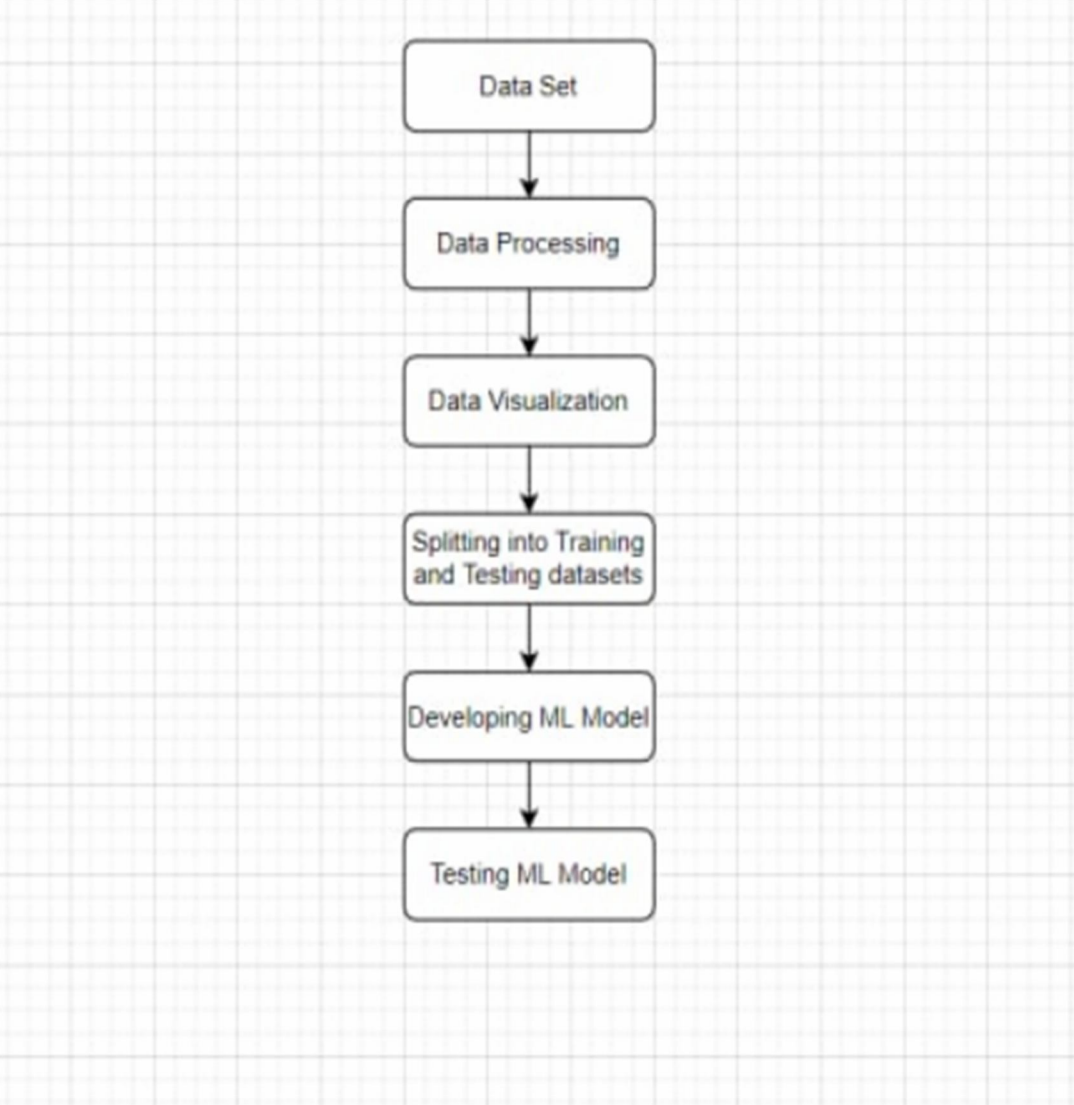
4) *Root Mean Square Error:* RSME (Root mean square error) calculates the transformation between values predicted by a model and actual values. In other words, it's one such error within the technique of measuring the precision and error rate of any machine learning algorithm of a regression problem.

IV. PROPOSED MODEL

In our proposed model ,we have collected the accurate data from Kaggle source. Later, we underwent some preprocessing procedures to improve the dataset's suitability for machine learning. Later, utilising some graphs from the dataset, we conduct some analysis. Our data set has now been divided into train data for our model training and test data for model performance evaluation. Mean Squared Error and Mean Absolute Errors are used to measure the variation between the actual and anticipated data.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

V. FLOW CHART



VI. IMPLEMENTATION

A. *Data Set*

Data Set Source: https://www.kaggle.com/datasets/nikhilmittal/flight-fare-prediction-mh

B. *Tools Used*

1) *Jupyter Notebook:* Build and exchange documents with live code, calculations, visualizations, and narrative text using this open-source web application.
2) *PyCharm:* It is an IDE to used to run flask application.

C. *Libraries Used*

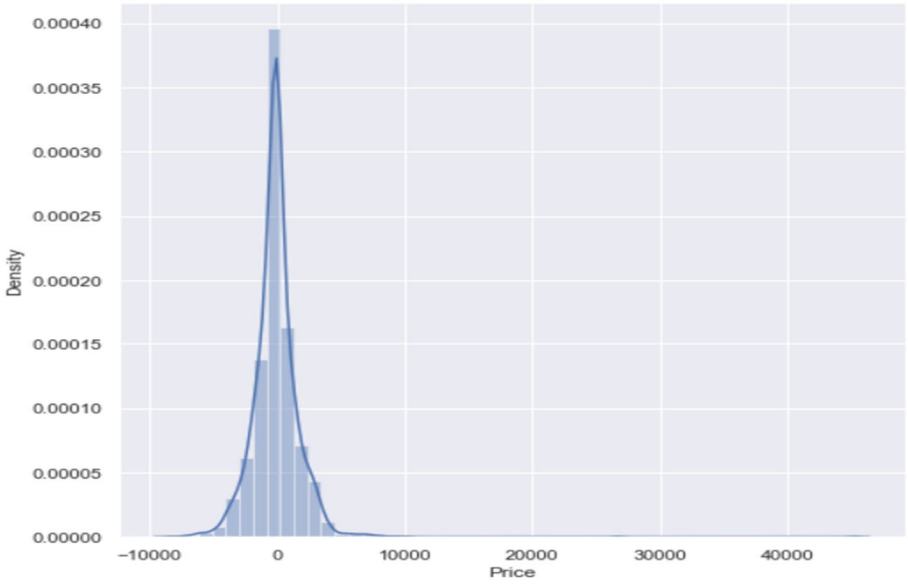
1) *Pandas:* Handle and import datasets.
2) *Numpy:* It is a library for mathematicians
3) *Seaborn:* To visualize random distributions.
4) *matplotlib.pyplo:* It is a group of operations that enable matplotlib to behave similarly to MATLAB.

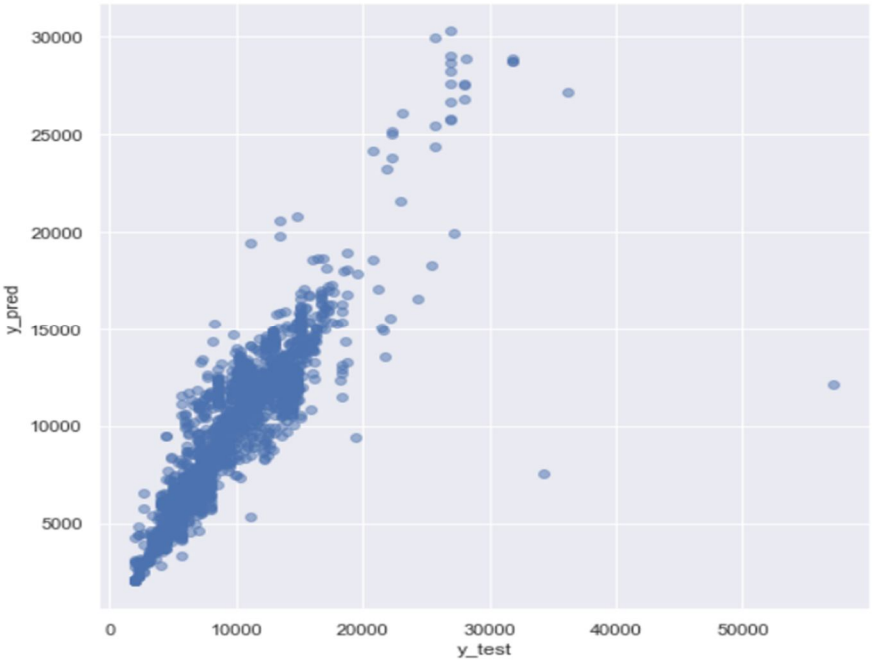
VII. EXPERIMENTAL RESULTS

We have used Random Forest as our model and Mean squared error, mean absolute error , Root Mean Square Error inorder to track how much we are far from actual value.

```
In [85]: print('MAE:', metrics.mean_absolute_error(y_test, prediction))
         print('MSE:', metrics.mean_squared_error(y_test, prediction))
         print('RMSE:', np.sqrt(metrics.mean_squared_error(y_test, prediction)))

         MAE: 1165.5497515930053
         MSE: 4047610.858961824
         RMSE: 2011.867505319827
```





ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

VIII. CONCLUSION AND FUTURE WORK

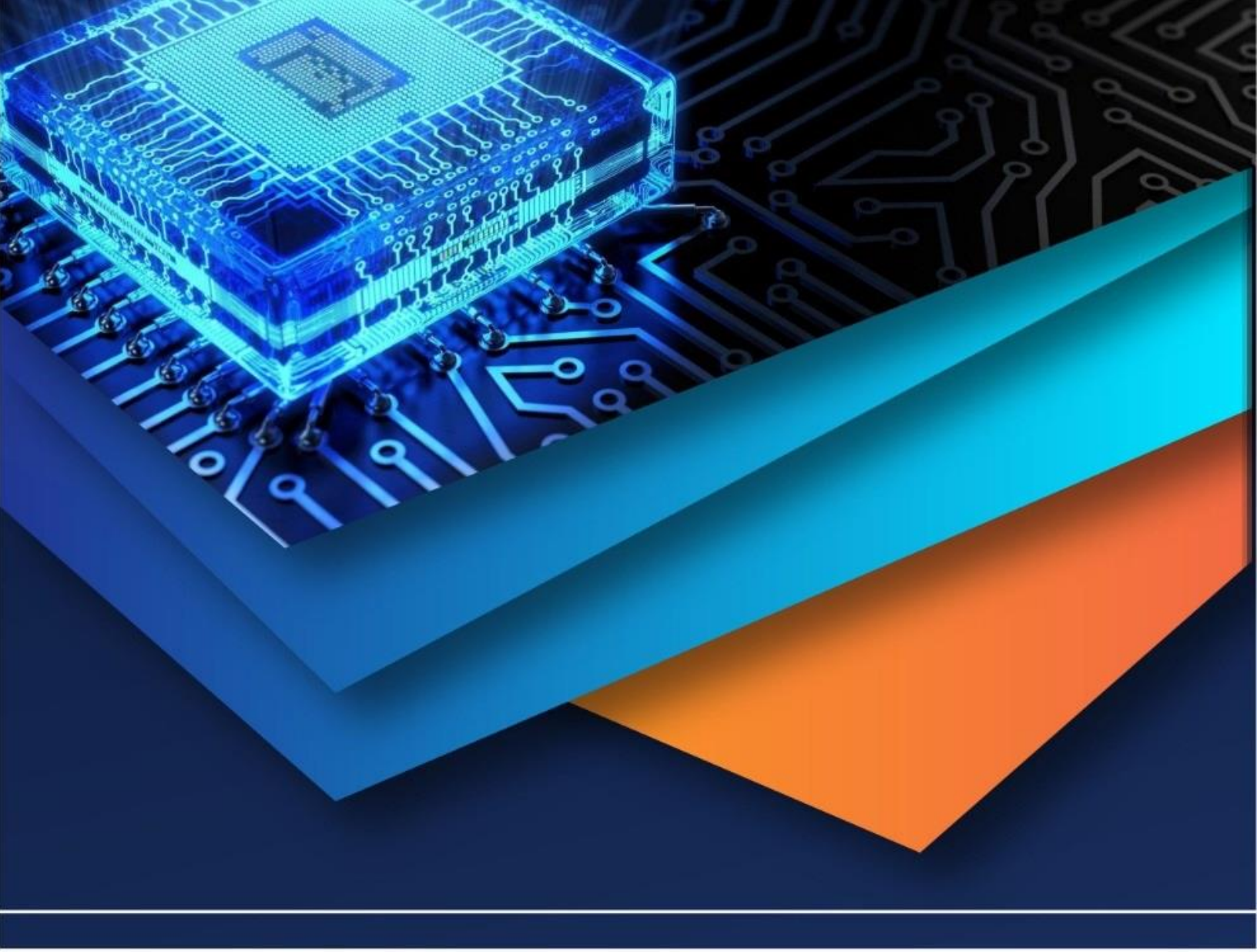
A dataset for routes for the most important Indian cities is constructed, and a trend of price fluctuation for a finite number of days is analysed. On the dataset, machine learning methods are used to forecast the dynamic flight pricing.

This provides the estimated flight fee to purchase a flight ticket at the lowest price. Information is gathered from websites. which sell the tickets for the flights, therefore only a limited amount of information accessed. The correctness of the model may be determined by looking at the mean squared error, mean absolute error, and root mean square error numbers.

The anticipated outcomes will be more accurate in the future if more data could be available, such as the present availability of seats.

REFERENCES

[1] https://www.javatpoint.com/machine-learning-random-forest-algorithm
[2] https://www.google.com/amp/s/www.geeksforgeeks.org/python-mean-squared-error/amp/
[3] https://www.statology.org/mean-absolute-error-python/
[4] https://www.google.com/amp/s/www.geeksforgeeks.org/python-mean-squared-error/amp/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY