Creating a Machine Learning Model to Improve the Accuracy of Predicting Airline Ticket Prices

Dr. Ram Bajaj[1*]

[1*]RNB Global University-Bikaner

ABSTRACT

Air travel serves as a pivotal mode of transportation for both domestic and international journeys, offering swift mobility. While it tends to be relatively costly compared to other transportation options, the price directly impacts our comfort and travel time. When demand exceeds capacity, airlines may adjust prices accordingly. Data pertaining to specific air routes, including constraints such as flight duration, arrival time, and available routes during a particular timeframe, has been collected to determine the baseline airfare.

INTRODUCTION

An important job is played via carrier cost prediction. Today, foreseeing the expense of any transportation has a huge effect.

Our undertaking's significant objective is to introduce the purchaser to an estimated cost close to the ticket's actual expe nse. The expense of activity is fundamentally scaled down by cost prediction using a PC analysis. Furthermore, it shows different evaluation levels for different organizations offering changing conveniences.

LITERATURE SURVEY

Expecting the extended cost of that day's tickets with any level of accuracy is difficult. In any case, AI models will help us foresee departures' expenses from source to destination, permitting clients to buy tickets at a discount and save some cash. The expense is lower when a client buys an airline ticket far sufficiently before the take-off date. Making AI models using different AI calculations, like Random Forest, MAE, MSE, and RSME, can accomplish these results. The AI model is made by considering and assessing the information assembled in earlier years, and the model will create the best results.

AI

AI is a subset of Computerized reasoning that helps programming models become considerably more precise without being expressly customized. Different elements for weather conditions measure foster different AI models.
These models utilize different calculations to get the ideal answer for a given issue.
For our model, we have used below techniques:

1) Random forest: Random Forest might be a classifier that contains a variety of decision trees on different subsets of the given dataset and takes the normal to support the proactive accuracy of that dataset.

2) Mean Squared error: The mean squared error (MSE) lets you know how tight a bend is to various focuses. The lower the MSE, the higher the index. Can achieve this by squaring the distances between the directions and the bend (likewise alluded to as the "blunders"). It additionally gives more weight to bigger contrasts. The figuring out is basic to destroy any bad signs. It's known as the mean squared error, as you track down the normal of various errors.
MSE equation = (1/n) * Σ(actual -
forecast)2

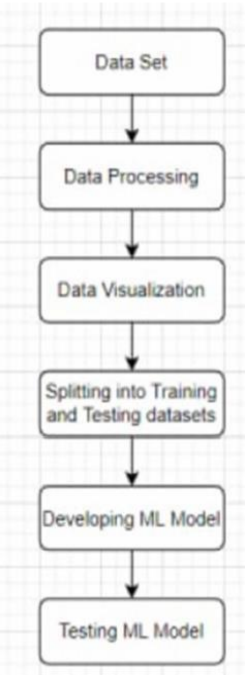
3) Mean outright Error: It might be a proportion of mistakes between matched perceptions communicating the impalpable peculiarity.
Mean Outright Mistake = (1/n) * Σ|yi - xi|

4) Root Mean Square error: It computes the change between values anticipated by a model furthermore, genuine qualities. It's one such mistake in estimating the accuracy and blunder pace of any AI calculation of a relapse issue.

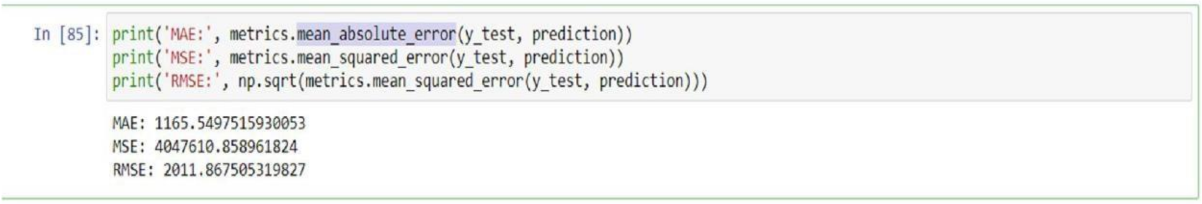
PROPOSED MODEL

We have gathered accurate information from the Kaggle source in our proposed model. Afterwards, we used some preprocessing methodology to work on the dataset's appropriateness for AI. Afterwards, we lead some investigation using a few charts from the dataset. Our informational collection has now been separated into train information for model preparation and test information for model execution assessment.
Mean Squared Error and Mean Outright Blunders measure the genuine and expected information variety.

FLOW CHART

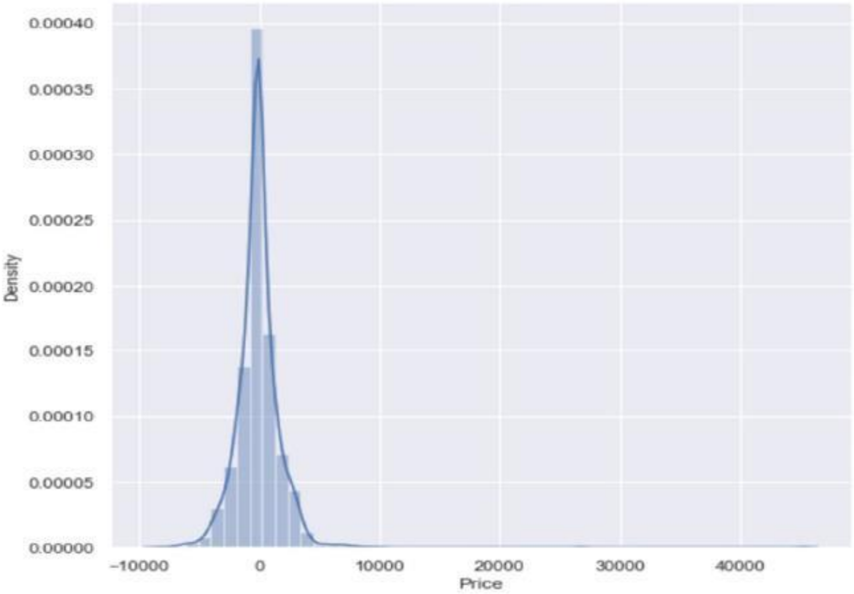


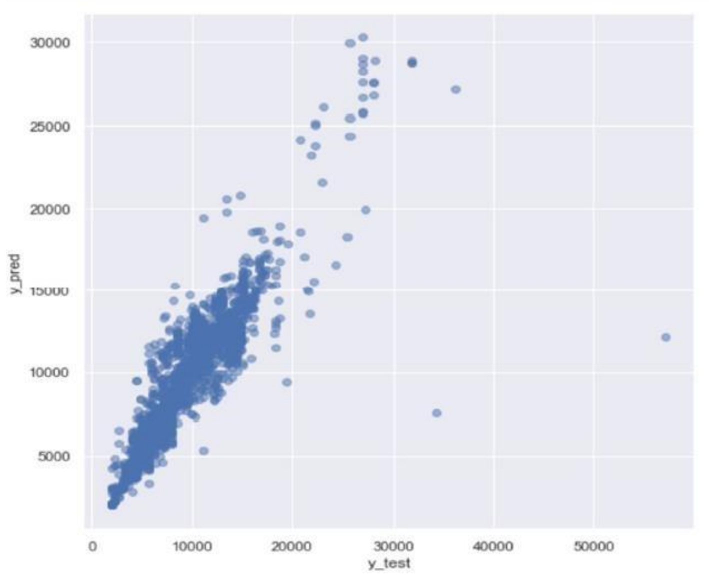
EXPERIMENTAL RESULTS



```
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
```





CONCLUSION AND FUTURE WORK

days is examined. AI techniques are utilized on the dataset to figure out
A dataset for courses for the main Indian dynamic flight evaluation.
urban communities is built, and a pattern of cost variance for a limited number of This gives the assessed flight charge to buy a flight ticket at the least cost. Data is accumulated from sites, which sell the tickets for the flights. Consequently, just a restricted measure of data is gotten to. May decide the rightness of the model by

REFERENCES

1. taking a gander at the mean squared mistake, mean outright blunder, and root mean square mistake numbers. The expected results will be more accurate if more information is accessible, for example, the current accessibility of seats.
2. https://www.javatpoint.com/machine-learning-random-forest-algorithm
3. https://www.google.com/amp/s/www.geeksforgeeks.org/python-mean-squared-error/amp/
4. https://www.statology.org/mean-absolute-error-python/
5. https://www.google.com/amp/s/www.geeksforgeeks.org/python-mean-squared-error/amp/