Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

Prediction Of Flight For Users By Using Different ML Techniques

Mr. G.Ramamohana Rao Assistant professor in the department of AI and IT at DVR & DR.HS MIC College of Technology (Autonomous), Kanchikacherla, NTR (DT).

Mr. G.Narendra MCA Student in the department of DCA at DVR & DR.HS MIC College of Technology (autonomous), Kanchikacherla, NTR(DT).

Abstract_ People who frequently travel by plane will be more knowledgeable about the best discounts and the ideal time to purchase a ticket. Many airline companies adjust their costs based on the seasons or the length of the flight.When individuals travel more, the price will rise. Estimating the highest prices of airline data for the route with features such as Duration, Source, Destination, Arrival, and Departure. The features are drawn from a selected dataset, and in this research, we employed machine learning techniques and regression algorithms to forecast the price of an airline ticket, which varies over time. We used decision tree and random forest algorithms, as well as KNN, to anticipate flight prices for customers. Random forest has the highest accuracy in predicting flight prices.

Keywords: Feature selection, Airfare price, Machine learning, Pricing Models, Prediction Model, Random Forest.

1. INTRODUCTION

Perfect time for purchasing plane ticket by the passenger's view is difficult since passengers get very less information of future business price rates. Different models figure out future business price on plane and categorise the best time to obtain flight ticket. Airlines use different strategies of pricing for their tickets, later taking the decision on price because order shows higher value for the approximation models. The causes behind the difficult system is each Planes has limited number of seats to be filed, so airlines must regulate demand. Suppose when demand is expected to increase capacity, the airline may increase prices, to decrease the rate at which seats fill. Also, seating arrangements in flight which is not occupied shows the loss of the amount invested for the business airline companies and making them purchase the ticket to fill the seats for any price this would be the best idea to get profit in loss too. Passengers should be compatible with the airline companies to get adjusted for the increase and decrease of the price. Passengers or customers should make their own planning to get the best offers available on different

airlines and travel through less price. Planes ticket prices changes as time passes, pulling out the elements which creates the difference. Reporting the correlated and models which is used to price the flight tickets. Then, using that information, building the model which helps passengers to make pull out the ticket to buy and predicting air ticket prices which progresses in the future. Duration, Arrival time, Price, Source, Destination and much more these are the attribute used for flight price prediction

2. LITERATURE SURVEY

2.1 K. Tziridis, Th. Kalampokas, G.A. Papakostas K.I. Diamantaras:

The issue of predicting ticket rates is covered in this essay. In order to achieve this, a collection of characteristics that define a typical flight are chosen, presuming that these characteristics have an impact on the cost of an airline ticket. Eight cutting-edge machine learning (ML) models are employed to forecast the pricing of airline tickets using the attributes, and the models' performance is compared to one another. This study examines the relationship between forecast accuracy and the feature set used to represent an airline, in addition to the prediction accuracy of each model.

2.2) Tao Liu, Jian Cao Yudong Tan, Quanwu Xiao:

In this study, we offer the ACER context-aware ensemble regression model, which incorporates various context-aware models and adaptively modifies context features. Context characteristics are arbitrarily chosen to efficiently cluster data, and several regression models are trained for data with various contexts. This approach is inspired by bagging and boosting. The context feature list is additionally constantly modified by removing some unnecessary elements. Our model is contrasted in the experiment on the real data set with the baseline regression model, random forest, and traditional time series models. The outcomes demonstrate that ACER outperforms the other models by a wide margin.

2.3) Viet Hoang Vu, Quang Tran Minh, Phu, H. Phung:

In this article, we provide a brand-new model that might assist the customer in anticipating price trends without relying on official airline information. Our results showed that the suggested model, despite lacking several essential components, such as the number of unsold seats on flights, can forecast trends as well as actual changes in airfare up to the departure dates using public airfare data that is

readily available online. We also determined the characteristics that have the biggest effects on changes in airfare..

3. PROPOSED SYSTEM

Estimating the highest prices of the airlines data for the route is collected with features such as Duration, Source, Destination, Arrival, Departure. Features are taken from chosen dataset

In this paper, we have used machine learning techniques such as random forest algorithm, decesion tree , KNN algorithm .

By using above algorithms here we are predicting predicting flight price .

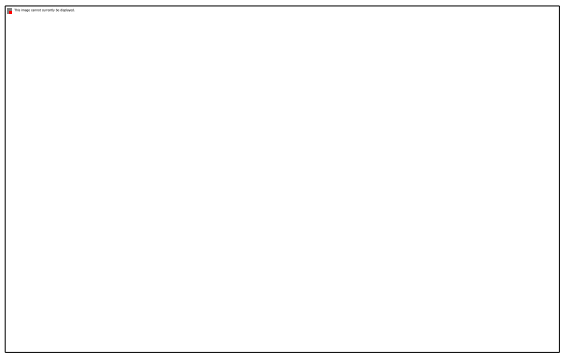


Fig 1:Architecture

3.1 IMPLEMENTATION

3.1.1 Gathering the datasets: We gather all the r data from the kaggale website and upload to the proposed model

3.1.2 Generate Train & Test Model: We have to preprocess the gathered data and then we have to split the data into two parts training data with 80% and test data with 20%

3.1.3 Run Algorithms: For prediction apply the machine learning models on the dataset by splitting the datasets in to 70 to 80 % of training with these models and 30 t0 20 % of testing for predicting

3.1.4 Predict output: in this module user will get flight price based on user input values

3.2 About Algorithms

3.2.1 KNN Algorithm

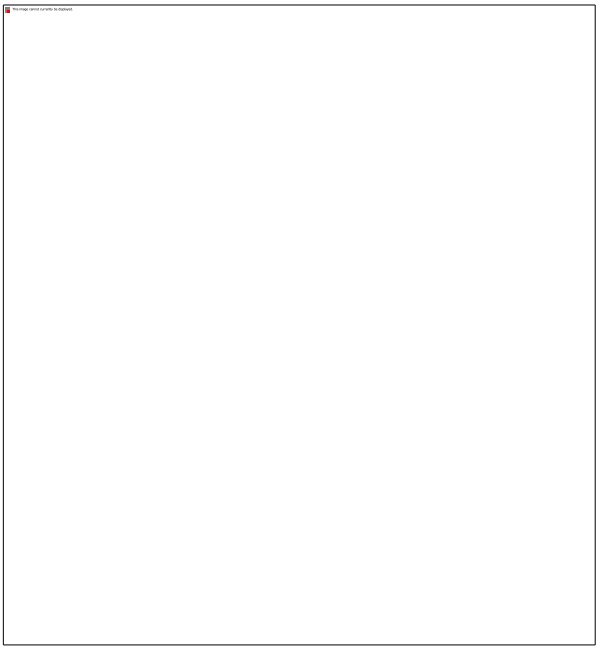
- K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.
- K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.
- K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.
- K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.
- K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 7, July : 2023

- It is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.
- KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

3.2.2 Decision Tree Classification Algorithm

- Decision Tree is a method of supervised learning that can be used to solve classification and regression problems, but it is typically used to solve classification problems. It is a classifier with a tree structure, with internal nodes representing a dataset's features, branches representing the decision rules, and each leaf node representing the result.
- The Decision Node and the Leaf Node are the two nodes in a Decision tree. Leaf nodes are the results of decisions and do not contain any additional branches, whereas Decision nodes are used to make any decision and have multiple branches.
- The features of the given dataset are used to make decisions or conduct tests.
- It is a graphical representation that is used to get all of the possible solutions to a problem or decision based on the conditions that have been given.
- It is referred to as a decision tree because, like a tree, it begins at the root node and grows into a structure similar to a tree.
- The CART algorithm, which stands for Classification and Regression Tree algorithm, is used to construct a tree.
- A decision tree simply asks a question and divides the tree into subtrees based on the answer (yes/no).
- Below graph makes sense of the general design of a choice tree:



3.2.3 Random Forest

The supervised learning method includes the well-known random forest machine learning

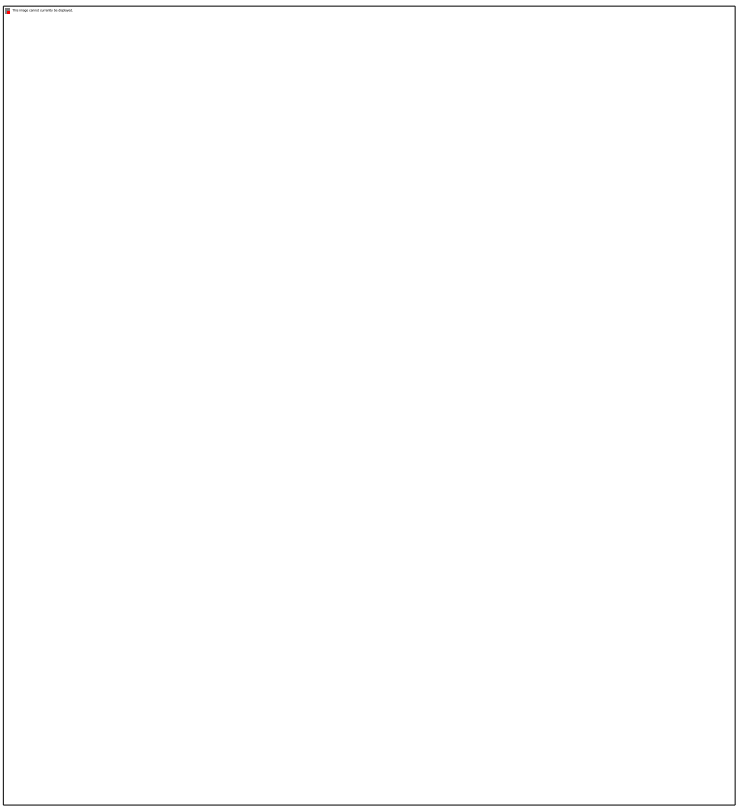
algorithm. In ML, it can be utilized for both regression and classification issues. It is based on the idea of ensemble learning, in which multiple classifiers are combined to solve a complex problem and boost the model's performance.

"Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset," as the name suggests, "Random Forest is a classifier." Rather than depending on one choice tree, the irregular timberland takes the forecast from each tree and in light of the greater part votes of expectations, and it predicts the last result.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

4.RESULTS AND DISCUSSIONS

The below diagram explains the working of the Random Forest algorithm:



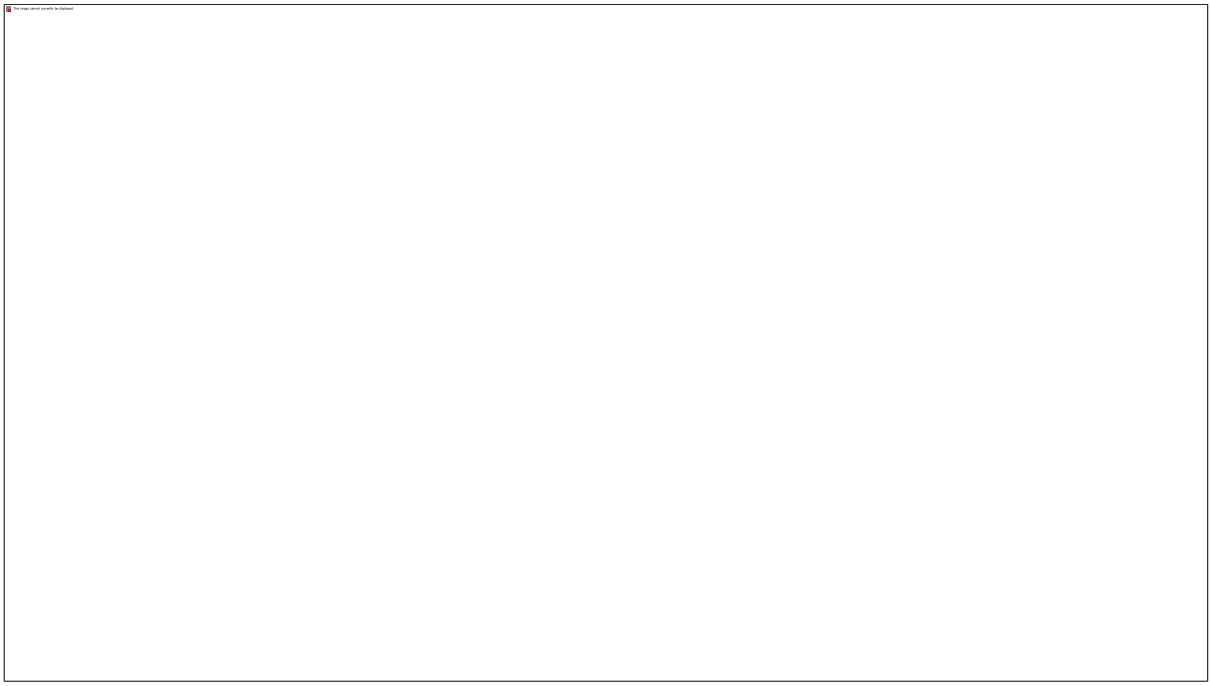


Fig 2: Predicted flight price based on input parameters

Fig 3: Predicted flight price based on input parameters

5.CONCLUSION

A dataset is collected, pre-processed, data modelling is performed, and a value difference

for the number of restricted days by the passengers for travel is explored while evaluating the algorithmic rule. Machine Learning methods with square measure for projecting airline fare and giving accurate value of plane price ticket at limited and highest value. Information is gathered from Kaggle websites that offer aircraft tickets, limiting data that is often viewed. The outcomes of the random forest, decision tree, and KNN algorithms. Here, random forest has superior accuracy, but the best accuracy is anticipated by the Random forest algorithm, as indicated in the above study. When a large amount of information, such as comprehensive information in a dataset, is accessible in the future, the projected outcomes are highly accurate. If you want to expand on it, you should request different sources of historical data or be very organised in collecting knowledge manually over a long period of time. A number of possible combinations of planes will be crossed. There is a potential that planes differ in their implementation ideas based on the plane's attributes. Finally, it is intriguing to compare the accuracy of our model to that of the business models available today.

REFERENCES

[1] Bachis, E., & Piga, C. A. (2011). "Low-cost airlines and online price dispersion. International Journal of Industrial Organization", 29(6), 655– 667. doi:10.1016/j.ijindorg.2011.02.006

[2] Groves, W. and Gini, M., 2021. "A Regression Model For Predicting Optimal Purchase Timing For Airline Tickets.". Available at: <https://conservancy.umn.edu/handle/11299/215872>

[3] T. Wohlfarth, S. Clemencon, F. Roueff and X. Casellato, "A Data-Mining Approach to Travel Price Forecasting," doi: 10.1109/ICMLA.2011.11.

[4] T. Wang et al., "A Framework for Airfare Price Prediction: A Machine Learning Approach," doi: 10.1109/IRI.2019.00041.

[5] V. H. Vu, Q. T. Minh and P. H. Phung, "An airfare prediction model for developing markets," doi: 10.1109/ICOIN.2018.8343221.

[6] K. Tziridis, T. Kalampokas, G. A. Papakostas and K. I. Diamantaras, "Airfare prices prediction using machine learning techniques," doi: 10.23919/EUSIPCO.2017.8081365.

[7] Y. Chen, J. Cao, S. Feng and Y. Tan, "An ensemble learning based approach for building airfare forecast service," doi: 10.1109/BigData.2015.7363846

[8] T. Liu, J. Cao, Y. Tan and Q. Xiao, "ACER: An adaptive context-aware ensemble regression model for airfare price prediction," doi: 10.1109/PIC.2017.8359563.[9] Boruah A., Baruah K., Das B., Das M.J., Gohain N.B. (2019) “A Bayesian Approach for Flight Fare Prediction Based on Kalman Filter,” https://doi.org/10.1007/978-981-13-0224-4_18

[10] T. Wang et al., "A Framework for Airfare Price Prediction: A Machine Learning Approach," doi: 10.1109/IRI.2019.00041.

[11] G.A. Papakostas, K.I. Diamantaras and T. Papadimitriou, “Parallel pattern classification utilizing GPU-Based kernelized slackmin algorithm,” doi:10.1016/j.jpdc.2016.09.001

[12] G. Francis, A. Fidato, and I. Humphreys, “Airport–airline interaction: the impact of low-cost carriers on two european airports,” doi:10.1016/s0969-6997(03)00004-8

.