FLIGHT FARE PREDICTION USING MACHINE LEARNING

Abstract:

Travelling through flights has become an integral part of today's lifestyle as more and more people are opting for faster travelling options. The flight ticket prices increase or decrease every now and then depending on various factors like timing of the flights, destination, duration of flights. various occasions such as vacations or festive season. Therefore, having some basic idea of the flight fares before planning the trip will surely help many people save money and time. In the proposed system a predictive model will be created by applying machine learning algorithms to the collected historical data of flights. This system will give people the idea about the trends that prices follow and also provide a predicted price value which they can refer to before booking their flight tickets to save money. This kind of system or service can be provided to the customers by flight booking companies which will help the customers to book their tickets accordingly

INTRODUCTION:

This project aims to develop an application which will predict the flight prices for various flights using machine learning model. The user will get the predicted values and with its reference the user can decide to book their tickets accordingly. In the current day scenario flight companies try to manipulate the flight ticket prices to maximize their profits. There are many people who travel regularly through flights and so they have an idea about the best time to book cheap tickets. But there are also many people who are inexperienced in booking tickets and end up falling in discount traps made by the companies where actually they end up spending more than they should have. The proposed system can help save millions of rupees of customers by proving them the information to book tickets at the right time. The proposed problem statement is “Flight Fare prediction system”.

Airline companies use complex algorithms to calculate flight prices given various conditions present at that particular time. These methods take financial, marketing, and various social factors into account to predict flight prices.

Nowadays, the number of people using flights has increased significantly. It is difficult for airlines to maintain prices since prices change dynamically due to different conditions. That’s why we will try to use machine learning to solve this problem. This can help airlines by predicting what prices they can maintain. It can also help customers to predict future flight prices and plan their journey accordingly.

REQUIREMENTS ANALYSIS:

System Requirements

I. Hardware Requirement

a. Laptop or PC
- I3 processor system or higher
- 4 GB RAM or higher
- 100 GB ROM or higher

II. Software Requirement

a. Laptop or PC
- Windows 7 or higher
- Front end: HTML, CSS, java script
- Middle ware: python, JavaScript
- PyCharm

Description of the present system of flight price prediction in python:

Flight price prediction is a popular problem in the field of machine learning and data science. In this task, we aim to predict the price of a flight ticket based on various factors such as the airline, departure date and time, destination, and number of stops.

There are various approaches to building a flight price prediction system, but one popular method is using a regression model. In Python, we can use popular machine learning libraries such as scikit-learn and TensorFlow to build regression models.

Here's a high-level overview of the steps involved in building a flight price prediction system:

Data collection: The first step is to collect data on past flight prices, along with various other factors such as the airline, departure date and time, destination, and number of stops. There are various sources for obtaining this data, including web scraping, APIs, and publicly available datasets.

Data preprocessing: Once we have collected the data, the next step is to preprocess it. This involves cleaning the data, handling missing values, and transforming the data into a format that can be used by our machine learning model.

Feature engineering: After preprocessing the data, we need to extract relevant features from the data that can help our machine learning model make accurate predictions. This may involve creating new features or transforming existing features.

Model selection: Once we have extracted relevant features from the data, we can then train and evaluate various regression models to find the best model for our task. This may involve comparing the performance of linear regression, decision trees, random forests, and other regression models.

Model training: After selecting a suitable regression model, we can then train the model on our preprocessed data.

Model evaluation: Once we have trained our model, we need to evaluate its performance on a test dataset to ensure that it is making accurate predictions.

Deployment: After evaluating our model, we can then deploy it in a production environment where it can be used to make real-time predictions on flight prices.

Overall, building a flight price prediction system in Python involves collecting and preprocessing data, extracting relevant features, selecting and training a regression model, and evaluating the model's performance

Limitation of the present system in Flight price Prediction in python

There are several limitations to the present system in flight price prediction in Python, some of which are:

Data quality: The accuracy of the predictions heavily relies on the quality of the data used to train the model. If the data used is incomplete, inaccurate or biased, it can lead to inaccurate predictions.

Lack of real-time data: Flight prices can change frequently, and the availability of real-time data is crucial for accurate predictions. However, the present system may not have access to real-time data, which can lead to outdated predictions.

Limited feature set: The present system may not include all the relevant features that affect flight prices…

Airline companies use complex algorithms to calculate flight prices given various conditions present at that particular time. These methods take financial, marketing, and various social factors into account to predict flight prices.

Nowadays, the number of people using flights has increased significantly. It is difficult for airlines to maintain prices since prices change dynamically due to different conditions. That's why we will try to use machine learning to solve this problem. This can help airlines by predicting what prices they can maintain. It can also help customers to predict future

Overflow of the model:

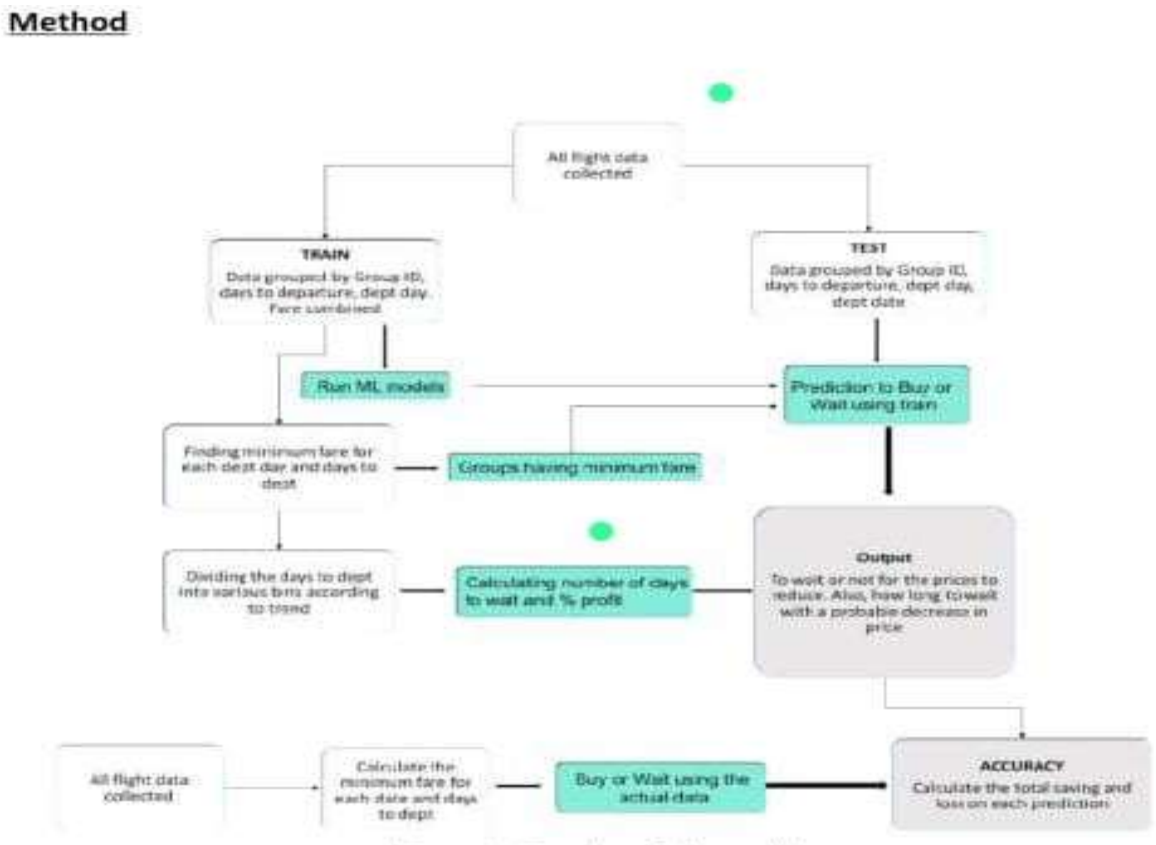
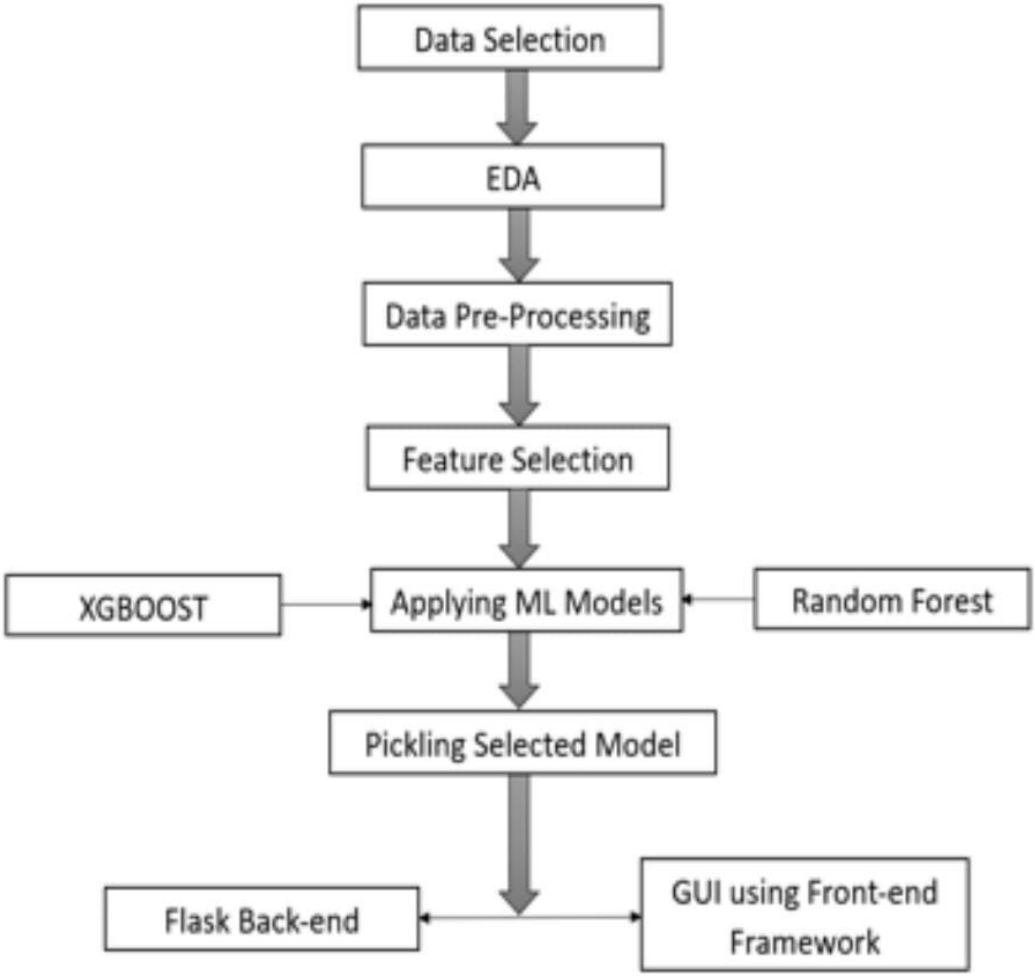
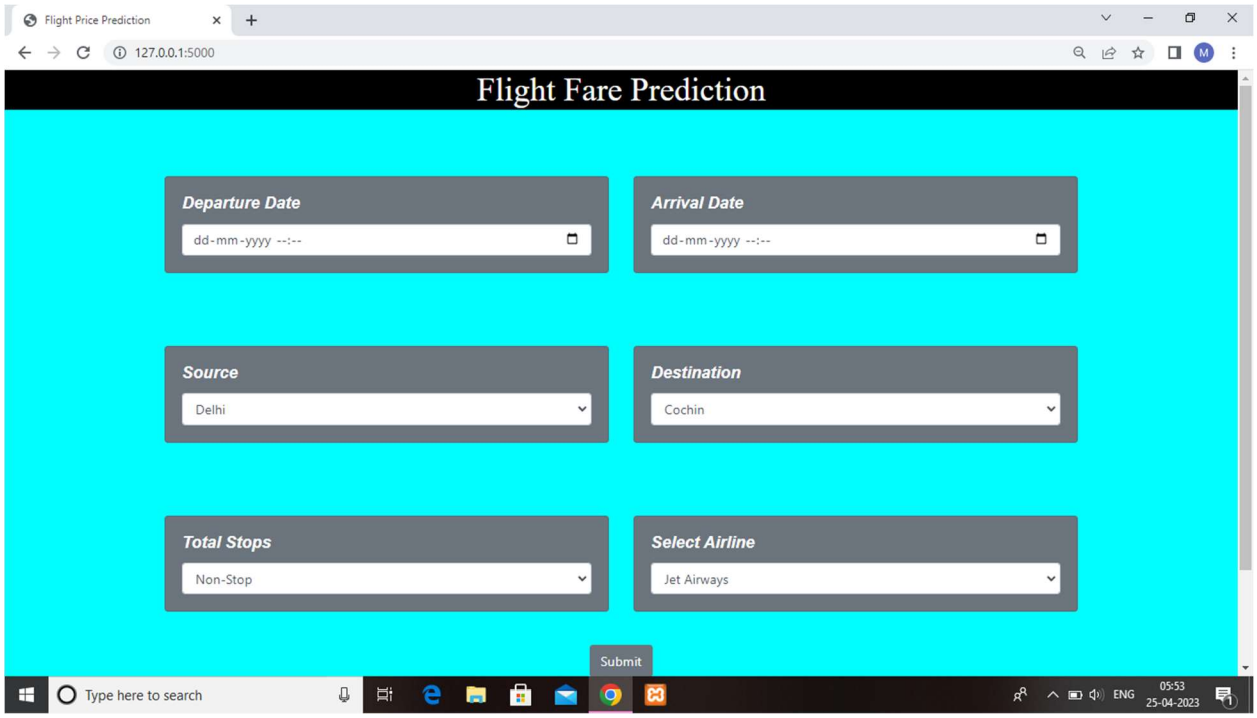


Figure 1 : Overview of the model

Proposed System:



Screen Design:



Conclusion:

Machine Learning algorithms are applied on the dataset to predict the dynamic fare of flights. This gives the predicted values of flight fare to get a flight ticket at minimum cost. Data is collected from the websites which sell the flight tickets so only limited information can be accessed. The values of R-squared obtained from the algorithm give the accuracy of the model. In the future, if more data could be accessed such as the current availability of seats, the predicted results will be more accurate. Finally, we have created the entire process of predicting an airline ticket and given a proof of our predictions based on the previous trends with our prediction.

Reference:

[1] K. Tziridis T. Kalampokas G.Papakostas and K. Diamantaras \"Airfare price prediction using machine learning techniques\" in European Signal Processing Conference (EUSIPCO), DOI:

10.23919/EUSIPCO .2017.8081365L. Li Y. Chen and Z. Li" Yawning detection for monitoring driver fatigue based on two cameras" Proc. 12th Int. IEEE Conf. Intel. Transp. Syst. pp. 1-6 Oct. 2009. [2] William Groves and Maria Gini \"An agent for optimizing airline ticket purchasing\" in proceedings of the 2013 international conference on autonomous agents and multi-agent systems. [3] J. Santos Dominguez-Menchero, Javier Rivera and Emilio TorresManzanera \"Optimal purchase timing in the airline market\". [4] Supriya Rajankar, Neha sakhrakar and Omprakash rajankar "Flight fare prediction using machine learning algorithms" International journal of Engineering Research and Technology (IJERT) June 2019. [5] Tianyi wang, samira Pouyanfar, haiman Tian and Yudong Tao \"A Framework for airline price prediction: A machine learning approach\" [6] T. Janssen \"A linear quantile mixed regression model for prediction of airline ticket prices\" [7] Wohlfarth, T.clemencon, S.Roueff "A Dat mining approach to travel price forecasting" 10th international conference on machine learning Honolulu 2011. [8] Vinod Kimbhaune, Harshil Donga, Ashutosh Trivedi, Sonam Mahajan and Viraj Mahajan research paper on flight fare prediction system. [9] W. Groves and M. Gini, ?An agent for optimizing airline ticket purchasing, ? 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), St. Paul, MN, May 06 - 10, 2013, pp. 1341-1342. [10] Viet Hoang Vu, Quang Tran Minh and Phu H. Phung,?An Airfare Prediction Model for Developing Markets?, IEEE paper 2018. [11] Wohlfarth, T. Clemencon, S.Roueff, ?A Dat mining approach to travel price forecasting?, 10 th international conference on machine learning Honolulu 2011. [12] Dominguez-Menchero, J.Santo, Reviera, ?optimal purchase timing in airline markets? ,2014 [13] medium.com/analytics-vidhya/mae-mse-rmse-coefficient-ofdetermination-adjusted-r-squared-which-metric-is bettercd0326a5697e article on performance metrics [14] www.keboola.com/blog/random-forest-regression article on random forest [15] https://towardsdatascience.com/machine-learning-basics-decisiontree-regression-1d73ea003fda article on decision tree regression. [16] O. Etzioni, R. Tuchinda, C. A. Knoblock, and A. Yates. To buy or not to buy: mining airfare data to minimize ticket purchase price. [17] Manolis Papadakis. Predicting Airfare Prices. [18] Groves and Gini, 2011. A Regression Model for Predicting Optimal Purchase TimingFor Airline Tickets. [19] Modeling of United States Airline Fares – Using the Official Airline Guide (OAG) and Airline Origin and Destination Survey (DB1B), Krishna Rama-Murthy, 2006. [20] B. S. Everitt: The Cambridge Dictionary of Statistics, Cambridge University Press, Cambridge (3rd edition, 2006). ISBN 0-521-69027-7. [21] Bishop: Pattern Recognition and Machine Learning, Springer, ISBN 0-387-31073-8. [22] E. Bachis and C. A. Piga. Low-cost airlines and online price dispersion. International Journal of Industrial Organization, In Press, Corrected Proof, 2011. [23] P. P. Belobaba. Airline yield management. an overview of seat inventory control. Transportation Science, 21(2):63, 1987. [24] Y. Levin, J. McGill, and M. Nediak. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. Management Science, 55(1):32–46, 2009 [25] B. Smith, J. Leimkuhler, R. Darrow, and Samuels, ?Yield managementat american airlines,?Interfaces, vol.22, pp. 8–31, 1992. [26] T. Janssen, ?A linear quantile mixed regression model for prediction of airline ticket prices,? Bachelor Thesis, Radboud University, 2014. [27] S.B. Kotsiantis, ?Decision trees: a recent overview,? Artificial Intelligence Review, vol. 39, no. 4,

pp. 261-283, 2013. [28] L. Breiman, ?Random forests, ? Machine Learning, vol. 45, pp. 5-32, 2001. [29] S. Haykin, Neural Networks – A Comprehensive Foundation. Prentice Hall, 2nd Edition, 1999. [30] H. Drucker, C.J.C. Burges, L. Kaufman, A. Smola and V. Vapnik, ?Support vector regression machines, ? Advances in neural information processing systems, vol. 9, pp. 155-161, 19.