A FLIGHT PRICE PREDICTOR SYSTEM USING PYTHON

[1]Tessy Abraham,[2] Chippy T,[3] Sreelekshmi B

[1,2,3] Assistant Professor

[1] Department of Computer Science and Engineering

[1] Mahaguru Institute of Technology, Kerala,India

Abstract: This study describes a machine learning-powered flight price prediction system. As the price of an airline ticket is determined by numerous factors, the advent of flight price prediction using machine learning approach ,aided clients in purchasing the most convenient and reasonable air ticket. This study provided a model that correlates with the elements of an ideal system, which include ticket pricing, seat availability, seasonal change and price rise, accuracy, market trends, and route time. Overall, this study structured the benefits and negatives of the current system, imprinting the future idea for improved air travel and accommodation.

Index Terms –**Flight Prediction, Air Ticket, Machine learning, Python**

I. INTRODUCTION

The Second World War resulted in significant expansion in the aviation sector, resulting in a dramatic transformation in both air travel and management. The twentieth century saw the introduction of digitalization in aviation, which advanced flight instrumentation; the twenty-first century saw the rise of drones for military, civilian, and recreational usage.

Domestic and international airlines structure the rating and pricing distribution for the sophisticated revenue management techniques. Currently, airlines are attempting to manipulate airfares in order to maximize profits through discounts and phony testimonies. Airlines are currently seeking to manipulate ticket pricing in order to maximize revenues. [3]Many individuals fly regularly, so they are aware of the best times to purchase cheap tickets. However, many consumers have little expertise ordering tickets and consequently end up falling into the trap of discounts from firms, resulting in them spending more than they should have. The approach that is suggested can help consumers save millions of rupees by giving them with the information they require in order to reserve tickets at the right time.

II. PROBLEM STATEMENT

To develop a flight pricing forecasting system based on machine learning principles in order to produce an optimal system for optimizing cost-effective plane tickets and facilities, resulting in thousands of satisfied clients.

III. LITERATURE REVIEW

Customer satisfaction and quality assurance are crucial in the present areas of civil aviation. Beginning in 2017, researchers in the aviation sector have been working hard to reach these targets. Fare, convenience, customer friendliness, and availability of plane tickets are all important considerations. Regression models such as Linear Regression (LR), Support Vector Machines (SVMs), and Random Forests (RF) are widely utilized in forecasting correct flight prices, according to the most recent update. [2] Over the last decade, academics have used machine learning algorithms and data mining techniques to better estimate observed prices.[3]According to the statistics of the International Civil Aviation Organization, 58% of individuals flew by plane [4]. Researchers used eight machine learning models, including ANNs, RF, SVM, and LR, to forecast ticket prices and compared their performance. The best regression model has an 88% accuracy. In their analysis, Bagging Regression Tree is recognized as the best model since it is resilient and does not change when multiple input feature sets are used [5]. Deep Regressor Stacking was proposed to improve prediction accuracy. The suggested method is a unique multi-target strategy that uses RF and SVM as regressors and is easily applicable to other comparable problem domains [6].

IV. EXISTING SYSTEM

Prediction of Airfare Prices Using machine learning techniques, a data set of 1814 Aegean Airlines data flights was collected and utilized to train a machine learning model for a research article. A variety of characteristics were utilized to train several models, demonstrating how feature selection may affect accuracy model [7]. In a case study [8] by William Groves, an agent is presented who can optimize the time of a customer's purchase. Partial least squares regression is used to build a model. The desired model is applied in the [9] research using the San Francisco-New York course's linear quantile mixed regression approach, where tickets are offered on a daily basis via an internet website.

V. PROPOSED SYSTEM

We need to combine legitimate data with our stage for predicting flight ticket prices. Organisations at the international level World Air Transport Statistics is a dependable but has a comprehensive aviation database platform. [1]They refresh their database once a year by combining it with a wide range of industry-related variables such as flight duration, quality, availability, rating, and feedback. They all work together to make the ambitions of each aviation industry a reality. So, by combining the aircraft origin and survey data with economic data covering current trends, seasonal changes, thriving tourism, and so on, they produce data for processing. They finally filter out their features for further duplication and reciprocation after analysing each target particular data. By repeating similar tests and statistical analyses, you may extract a system with the fewest mistakes and the highest profit.

VI. TOOLS AND METHODOLOGY

Hardware Requirements:1Intel core i5,8GB Ram,Fast Ethernet,LCD Monitor,Windows
Software Requirements:PyScript
Language Used:Python
Editor:Visual Studio

VII. PROPOSED METHODOLOGY

Our suggested system is a unidirectional structure made up of predetermined parts. The command begins by entering the necessary parameters in the working setup, such as destination, date, flight name, and route. After accessing the website, those details will extract features, and from the listed set of features, they will select the most relevant and frequently requested data's, which will then be automatically routed to the back-end system, which will use machine learning techniques in Python, filtering them to produce reasonable airfare details for the customer. [8]Machine learning techniques are used in our system to generalise correct data. Since machine learning is one of the most commonly acknowledged domains of AI, it focuses on data and algorithms to improve its efficiency. They were specifically trained to give classifications and predictions that disclose the crucial data mining observation.
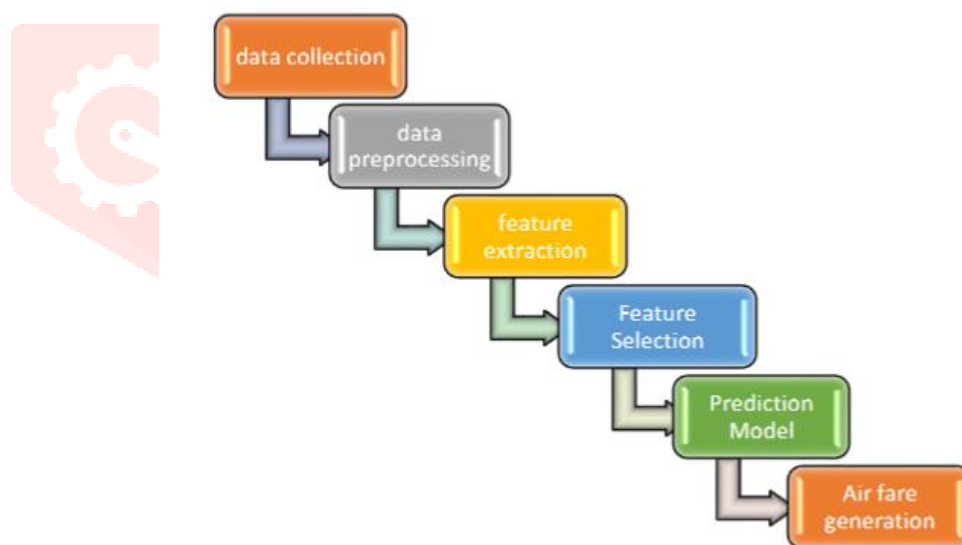

Fig.1 Architecture of the proposed method

Fig.2 Implementation method

The first phase is data selection, which involves collecting past flight data for a price prediction model. Our dataset includes over 10,000 flight-related data entries and pricing. The dataset's functionalities include source, destination, departure date and time, number of stops, arrival time, expenses, and more. During the exploratory data analysis phase, we sanitized the dataset by removing duplicate and null values. If these values are not removed, the model's accuracy suffers.

The following stage is data pre-processing, where we discovered that the majority of the data was stored in string format. The data for each characteristic is received, such as the day and month from the trip's date in integer format, as well as the hours and minutes from the departure time. Because the source and destination characteristics were categorical, they had to be converted to values. Categorical variables are turned into model-identifiable values for this One utilizing hot-coding and label coding techniques. The feature selection process entails picking the important qualities that are more closely related to pricing.

Following the selection of price-related characteristics, the following phase involves the use of a machine algorithm and model creation. Because our dataset contains labelled data, we will employ supervised machine learning methods, as well as regression algorithms similar to ours. [10]The functions in the dataset have continuous values. Regression models are used to describe the relationship between dependent and independent variables. In this project, we will use the following machine learning algorithms: Multiple linear regressions will be used to evaluate the connection between two or more independent variables and one dependent variable. There is only one independent and dependent function in basic linear regression. However, our dataset comprises a large number of independent functions on which the pricing can vary.

The multiple linear regression model is depicted below: $Y = \beta 0 x1 + \ldots + \beta n x n + \varepsilon$
Y = the expected value of the dependent variable Xn independent variables

When all other parameters are zero, the y-intercept is equal to n = coefficients of independent variables. As decision nodes, the tree selects an independent variable from the dataset. When test data is entered into the model, the model determines the outcome by determining which subsection the data point belongs to.[4] Depending on which subsection the data point belongs to, the decision tree's output will be the average value of all the data points in the subsection.

VIII. RESULTS

Through this study, we discovered the shortcomings of existing systems and, by marginalizing them, we were able to see new ideas and solutions for issue solving. We learn how to utilize computer language and technology while constructing the flight price forecast system. This will also introduce us to several platforms that take alternative approaches to the same problem, as well as refresh our Python modules and technology. We generate the result by comparing multiple regression models with feature extraction utilizing the available data and techniques. Through this effort, we achieve our aim of identifying the elements that improve model performance.
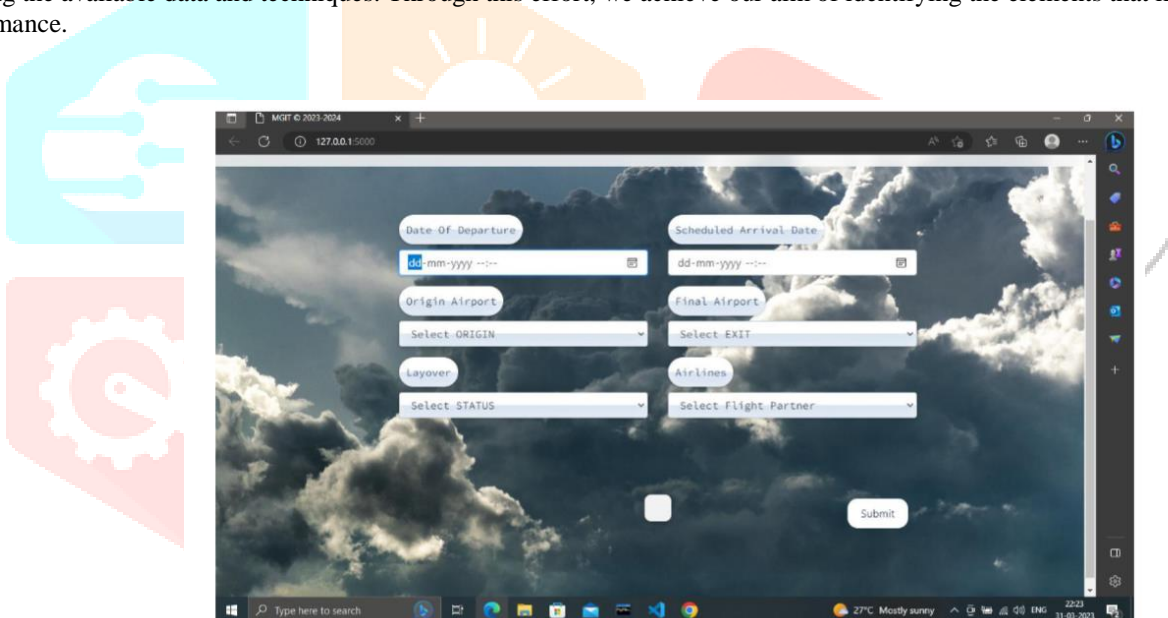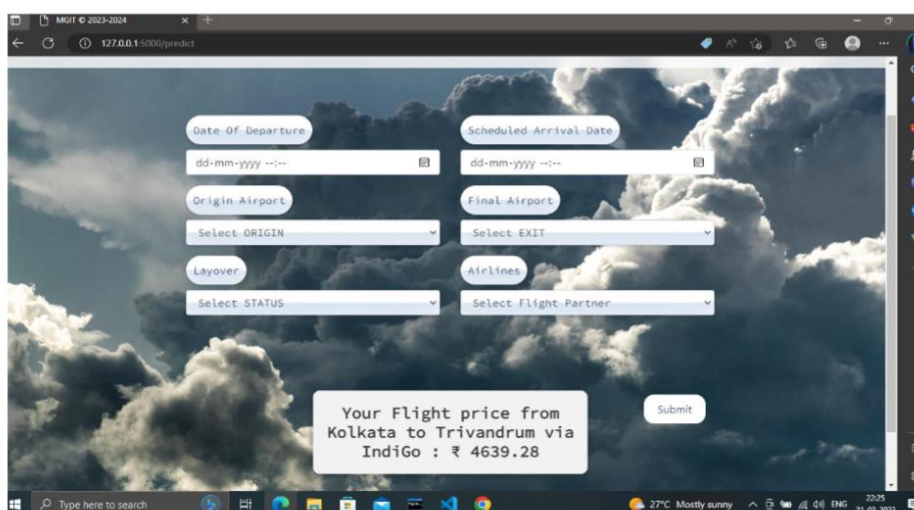

Fig. 2 Page to enter details

Fig. 3 Flight Price Predicted

IX. CONCLUSION

This working platform improves the efficiency and quality of air transport and lodging by combining real-time data and AI technologies using Python. The goal here is to create a system with few mistakes and maximum efficiency that supports appropriate airfare based on several parameters inside a given dataset. Using this configuration, a consumer may obtain correct information about dynamic flight fares and so obtain the cheapest airline ticket.

X. FUTURE SCOPE

If time allows, we would improve the working environment by include real-time feedback and comments from pre-occupied clients. So that we can assist our customers in researching each airline individually and rating them based on previous feedback, facilities offered, travel environment, complaint portal, and so on. that will ultimately give them confidence in choosing airlines for travel. In the long run, we may provide unique discounts and benefits for frequent customers, bolstering capacity growth. Finally, recognizing the critical necessity and popularity of air transportation will aid in the development of distinctive characteristics such as seat placement, covered auxiliary items, and so on. By merging such data with current market and worldwide trends, an important and comprehensive air fare forecast system is created.

XII. ACKNOWLEDGMENT

XIII. REFERENCES

[1] https://www.unwto.org

[2] Wang, Tianyi & Pouyanfar, Samira & Tian, Haiman & Tao, Yudong & Alonso, Miguel & Luis, Steven & Chen, Shu-Ching. (2019). A Framework for Airfare Price Prediction: A Machine Learning Approach.200-207.10.1109/IRI.2019.00041.

[3] T. Janssen, T. Dijkstra, S. Abbas, and A. C. van Riel, "A linear quantile mixed regression model for prediction of airline ticket prices," Radboud University, 2014.

[4] https://www.icao.int

[5] K. Tziridis, T. Kalampokas, G. A. Papakostas, and K. I. Diamantaras,"Airfare prices prediction using machine learning techniques," in the 25th IEEE European signal processing conference, 2017, pp. 1036–1039.

[6] E. J. Santana, S. M. Mastelini, and S. Barbon Jr, "Deep regressor stacking for air ticket prices prediction," in the XIII Brazilian symposium on information systems: information systems for participatory digital governance. Brazilian Computer Society (SBC), 2017, pp. 25–31.

[7] Joshi, Achyut & Sikaria, Himanshu & Devireddy, Tarun. (2017). Predicting Flight Prices in India.

[8] William Groves and Maria Gini, Department of Computer Science and Engineering, University of Minnesota, USA groves_gini@cs.umn.edu

[9] Supriya Rajankar, Neha Sakharkar, Omprakash Rajankar, "Predicting The Price Of A Flight Ticket With The Use Of Machine Learning Algorithms", INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 8, ISSUE 12, ISSN 2277-8616, DECEMBER 2019.

[10] https://www.ibm.com/topics/machine-learning