FLIGHT FARE PREDICTION SYSTEM

A Synopsys Submitted
In Partial Fulfillment of the
Requirements for the
Degree of

CSE(AIML)

Fourth Semester Project Synopsis

in

Machine Learning

by

NISHANT KUMAR (2201331530132)
YUVRAJ KUMAR SINGH (2201331530217)
OMKAR SINGH (2201331530137)
ABHISHEK SINGH (2201331530008)

CONCERNED FACULTY:

OSHIN MISHRA(Assistant professor)
ARUSHI THUSU(Assistant professor)

To the

FACULTY OF SOCET(CSE[AIML])

Noida Institute of Engineering and Technology, Greater Noida

An Autonomous Institute

May-2024

ABSTRACT

The Flight Fare Prediction System offers a comprehensive solution to the challenge of accurately forecasting flight ticket prices amidst the dynamic landscape of the airline industry. With the industry's continuous expansion and evolving fare structures, predicting flight fares has become increasingly complex. This system harnesses the power of machine learning algorithms and extensive historical flight data to deliver precise fare predictions. Drawing from a vast dataset encompassing various factors such as travel dates, destinations, airlines, departure times, and other pertinent variables, the system employs advanced machine learning techniques to discern patterns and relationships, thereby enabling reliable predictions of future flight fares. Utilizing a blend of regression algorithms and ensemble methods, the Flight Fare Prediction System ensures high accuracy in its predictions. It takes into account a multitude of factors influencing ticket prices, including seasonality, market demand, fuel costs, competition, and other dynamic variables. By integrating real-time data updates, the system ensures that predictions remain current and reflective of the latest market trends.

INTRODUCTION

The Flight Fare Prediction System represents a machine learning initiative dedicated to estimating aircraft ticket costs using relevant features and historical data. This approach serves travellers, travel firms, and airlines, aiding them in projecting trip expenses for planning, budgeting, and making informed choices.

The project's objective lies in constructing a dependable machine-learning model for forecasting flight expenses by encompassing various factors such as travel class, airline, departure and arrival destinations, travel dates, and other relevant details. To train the algorithm effectively, an extensive dataset comprising historical flight data, including ticket pricing and related attributes, will be utilized.

Users of the Flight Fare Prediction System will gain access to a user-friendly interface where they can input their travel information and receive an estimated flight fare. The system will meticulously analyse input data and generate precise predictions through feature engineering, data preprocessing, and machine learning methodologies. Rigorous evaluation criteria will be applied to ensure the accuracy and reliability of the model and its associated attributes.

While the accuracy of predictions heavily relies on the quality of the training and prediction data, the project places a premium on data quality and integrity. Data preprocessing techniques such as data cleaning, handling missing values, and feature scaling will be employed to validate the legitimacy and trustworthiness of the data utilized for training and prediction purposes.

The Flight Fare Prediction System holds the potential to assist consumers in planning their travel budgets, aid travel agencies in offering competitive pricing to their clientele, and support airlines in devising effective pricing strategies and revenue management tactics. By leveraging machine learning, the system aims to offer valuable insights and advantages to the travel industry, precisely estimating airline fares and enhancing decision-making processes.

The project's ultimate aim is to establish a dependable and accurate flight fare prediction system that delivers flight rates based on key parameters. Evaluation metrics such as forecast accuracy, model performance, and usability will be leveraged to assess systems thoroughly. Ethical considerations, including the handling of personal data and ensuring fairness in assumptions, are carefully incorporated into the project framework.

In conclusion, the Flight Fare Prediction System signifies a machine learning endeavor aimed at developing a system capable of accurately predicting trip expenses based on historical data and relevant attributes. This technology harbors the potential to enhance decision-making within the travel industry and offer valuable insights and benefits to travelers, travel agencies, and airlines alike.

TECHNOLOGY USED

1.MACHINE LEARNING

Machine learning is a field of artificial intelligence that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without being explicitly programmed. It encompasses various techniques such as supervised learning, unsupervised learning, and reinforcement learning, and finds applications in diverse areas such as finance, healthcare, marketing, and robotics.

2.SERVER DEPLOYMENT

Server deployment involves the process of setting up and configuring server hardware and software to make applications or websites accessible over a network. This includes tasks such as installing the operating system, configuring network settings, setting up security measures, installing necessary software dependencies, and deploying the application code or website files.

3.PYTHON

Python is a high-level, interpreted programming language known for its simplicity, readability, and versatility. It was created by Guido van Rossum and first released in 1991. Python emphasizes code readability and expressiveness, making it ideal for beginners and experienced programmers alike.

4.CSS

CSS is a style sheet language used to control the presentation and layout of HTML documents. It allows web developers to define styles such as colours, fonts, spacing, and positioning for HTML elements. CSS works by selecting HTML elements and applying styles to them through rules defined in style sheets.

5.HTML

HTML is the standard markup language used to create and design web pages. It provides the structure and layout for content on the World Wide Web. HTML consists of a series of elements, each enclosed in angle brackets, which define the structure of a webpage. These elements include headings, paragraphs, lists, links, images, and more.

PROBLEM STATEMENT

Everyone knows that holidays always call for a much-needed vacation and planning the travel itinerary becomes a time-consuming task. The commercial aviation business has grown tremendously and has become a regulated marketplace as a result of the worldwide growth of the Internet and E-commerce. Hence, for Airline revenue management, different strategies like customer profiling, financial marketing, and social factors are used for setting ticket fairs. When tickets are booked months in advance, airfares are often reasonable, but when tickets are booked in a hurry, they are often higher. But, the number of days/hours until departure isn't the only factor that decides flight fare, there are numerous other factors as well. Customers find it quite difficult to obtain a perfect and lowest ticket deal due to the aviation industry's complex pricing methodology. This project aims to address the need for a dependable and accurate flight fare prediction system that provides travellers, travel agencies, and airlines with reliable estimates of flight fares.

LITERATURE SURVEY

A. "Airline Ticket Price Prediction: A Machine Learning Approach" by M. L. Ahirrao, et al. (2018): This research paper proposes a flight fare prediction model using machine learning techniques such as regression algorithms and time-series analysis. The study explores various factors influencing ticket prices and compares the performance of different algorithms in predicting fare trends.

B. Flight Fare Prediction using Historical Data and Machine Learning Techniques" by A. Kumar, et al. (2019): The paper presents a flight fare prediction system that combines historical flight data and machine learning algorithms to forecast ticket prices. It analyzes factors such as departure time, travel duration, and airline popularity to generate accurate fare predictions. The study compares the performance of different algorithms and discusses the potential for improving prediction accuracy.

C. "Airline Fare Prediction Using Machine Learning" by A. L. Rodrigues, et al. (2020): This work focuses on predicting airline fares using machine learning techniques. The study considers various parameters, including airline popularity, route distance, and historical fare data, to train a predictive model. The authors explore the performance of different algorithms and discuss the implications of their findings for fare prediction accuracy.

D. "Predicting Airfare Using Machine Learning Techniques" by S. Aruna, et al. (2020): The paper presents a comparative analysis of different machine learning algorithms for predicting airfare. The study considers factors such as seasonality, time of booking, and flight class to develop a prediction model. The authors evaluate the performance of regression algorithms, including linear regression, support vector regression, and random forest regression.

E. "Flight Fare Prediction Using Ensemble Learning Techniques" by M.Sharma, et al. (2021): This research focuses on the application of ensemble learning techniques for flight fare prediction. The study combines multiple machine learning models, including decision trees, random forests, and gradient boosting, to improve prediction accuracy. The authors compare the performance of individual models and ensemble methods to identify the most effective approach.

F. “Flight Fare Prediction using Machine Learning Techniques” by K. Kumar and team (2017). This study compares the performance of various machine learning techniques, including decision trees, support vector machines, k-nearest neighbours, and random forests for flight fare prediction. This study also employs features engineering techniques to extract relevant features from flight data and evaluates the models using metrics such as mean squared error (MSE) and R-squared.

METHODOLOGY

1. Data Collection:

Gather historical flight fare data from reliable sources such as airline databases, online travel agencies, and public datasets.Collect relevant features including departure and arrival destinations, travel dates, airline, booking class, and fare prices.

2. Data Preprocessing:

Cleanse the collected data by handling missing values, outliers, and inconsistencies. Normalize or scale numerical features to ensure uniformity in data distribution. Encode categorical variables using techniques such as one-hot encoding or label encoding. Perform feature engineering to create additional features that may enhance the predictive power of the model.

3. Model Selection:

Experiment with various machine learning algorithms suitable for regression tasks, such as linear regression, decision trees, random forests, support vector machines (SVM), and gradient boosting algorithms. Assess the performance of each algorithm using appropriate evaluation metrics such as mean absolute error (MAE), mean squared error (MSE), or root mean squared error (RMSE). Select the best-performing algorithm based on evaluation results and computational efficiency.

4. Model Training and Evaluation:

Split the preprocessed data into training and testing sets for model training and evaluation, respectively. Train the selected machine learning model on the training data, optimizing model parameters using techniques such as cross-validation or grid search. Evaluate the trained model's performance on the testing data using evaluation metrics and assess its ability to accurately predict flight fares.

5. Model Deployment:

Deploy the trained machine learning model in a production environment, making it accessible via a user-friendly interface. Develop a user interface that allows travellers to input their travel preferences and receive estimated flight fares. Implement backend logic to process user inputs, make fare predictions using the deployed model, and display results to the user. Ensure scalability, reliability, and security of the deployed system to handle real-time user requests and protect sensitive data.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Airline	e_of_Jour	Source	Destinatior	Route	Dep_Time	rrival_Tim	Duration	Total_Stop	ditional_Ir	Price	
2	IndiGo	24/03/201	Banglore	New Delhi	BLR → DEL	22:20	01:10 22 M	2h 50m	non-stop	No info	3897	
3	Air India	1/05/2019	Kolkata	Banglore	CCU → IXR	05:50	13:15	7h 25m	2 stops	No info	7662	
4	Jet Airway	9/06/2019	Delhi	Cochin	DEL → LKC	09:25	04:25 10 Ju	19h	2 stops	No info	13882	
5	IndiGo	12/05/201	Kolkata	Banglore	CCU → NA	18:05	23:30	5h 25m	1 stop	No info	6218	
6	IndiGo	01/03/201	Banglore	New Delhi	BLR → NA	16:50	21:35	4h 45m	1 stop	No info	13302	
7	SpiceJet	24/06/201	Kolkata	Banglore	CCU → BLI	09:00	11:25	2h 25m	non-stop	No info	3873	
8	Jet Airway	12/03/201	Banglore	New Delhi	BLR → BOI	18:55	10:25 13 M	15h 30m	1 stop	In-flight m	11087	
9	Jet Airway	01/03/201	Banglore	New Delhi	BLR → BOI	08:00	05:05 02 M	21h 5m	1 stop	No info	22270	
10	Jet Airway	12/03/201	Banglore	New Delhi	BLR → BOI	08:55	10:25 13 M	25h 30m	1 stop	In-flight m	11087	
11	Multiple ca	27/05/201	Delhi	Cochin	DEL → BOI	11:25	19:15	7h 50m	1 stop	No info	8625	
12	Air India	1/06/2019	Delhi	Cochin	DEL → BLR	09:45	23:00	13h 15m	1 stop	No info	8907	
13	IndiGo	18/04/201	Kolkata	Banglore	CCU → BLI	20:20	22:55	2h 35m	non-stop	No info	4174	
14	Air India	24/06/201	Chennai	Kolkata	MAA → CC	11:40	13:55	2h 15m	non-stop	No info	4667	
15	Jet Airway	9/05/2019	Kolkata	Banglore	CCU → BO	21:10	09:20 10 M	12h 10m	1 stop	In-flight m	9663	
16	IndiGo	24/04/201	Kolkata	Banglore	CCU → BLI	17:15	19:50	2h 35m	non-stop	No info	4804	
17	Air India	3/03/2019	Delhi	Cochin	DEL → AM	16:40	19:15 04 M	26h 35m	2 stops	No info	14011	
18	SpiceJet	15/04/201	Delhi	Cochin	DEL → PN	08:45	13:15	4h 30m	1 stop	No info	5830	
19	Jet Airway	12/06/201	Delhi	Cochin	DEL → BOI	14:00	12:35 13 Ju	22h 35m	1 stop	In-flight m	10262	
20	Air India	12/06/201	Delhi	Cochin	DEL → CCU	20:15	19:15 13 Ju	23h	2 stops	No info	13381	
21	Jet Airway	27/05/201	Delhi	Cochin	DEL → BOI	16:00	12:35 28 M	20h 35m	1 stop	In-flight m	12898	
22	GoAir	6/03/2019	Delhi	Cochin	DEL → BOI	14:10	19:20	5h 10m	1 stop	No info	19495	

9540	Jet Airway	9/05/2019	Kolkata	Banglore	CCU → BO	14:05	18:15 10 M	28h 10m	1 stop	No info	13584
9541	Air India	21/05/201	Delhi	Cochin	DEL → AM	05:00	07:40 22 M	26h 40m	2 stops	No info	11806
9542	GoAir	1/04/2019	Delhi	Cochin	DEL → AM	07:25	13:35	6h 10m	1 stop	No info	4563
9543	Jet Airway	27/03/201	Delhi	Cochin	DEL → MA	05:10	04:25 28 M	23h 15m	2 stops	In-flight m	9412
9544	Jet Airway	27/06/201	Delhi	Cochin	DEL → BOI	21:50	04:25 28 Ju	6h 35m	1 stop	In-flight m	12898
9545	Vistara	27/04/201	Delhi	Cochin	DEL → COI	14:40	17:50	3h 10m	non-stop	No info	6216
9546	Jet Airway	21/05/201	Delhi	Cochin	DEL → BOI	13:00	04:25 22 M	15h 25m	1 stop	In-flight m	12898
9547	Jet Airway	9/06/2019	Delhi	Cochin	DEL → BOI	22:50	19:00 10 Ju	20h 10m	1 stop	In-flight m	10262
9548	Air India	01/03/201	Banglore	New Delhi	BLR → DEL	13:20	16:10	2h 50m	non-stop	No info	12526
9549	Air India	6/03/2019	Delhi	Cochin	DEL → TRV	05:15	06:50 07 M	25h 35m	1 stop	No info	11522
9550	Jet Airway	12/06/201	Mumbai	Hyderabad	BOM → HY	02:55	04:20	1h 25m	non-stop	No info	5678
9551	Vistara	09/03/201	Banglore	New Delhi	BLR → DEL	11:30	14:10	2h 40m	non-stop	No info	9918
9552	Jet Airway	1/05/2019	Kolkata	Banglore	CCU → BO	21:10	08:15 02 M	11h 5m	1 stop	No info	14781
9553	Air India	27/03/201	Delhi	Cochin	DEL → AM	05:00	19:15 28 M	38h 15m	2 stops	No info	8026
9554	Air India	24/03/201	Banglore	New Delhi	BLR → DEL	13:20	16:10	2h 50m	non-stop	No info	7276
9555	Jet Airway	1/04/2019	Kolkata	Banglore	CCU → BO	08:25	18:15	9h 50m	1 stop	In-flight m	7064
9556	Multiple ca	21/03/201	Delhi	Cochin	DEL → BOI	13:00	21:00	8h	1 stop	No info	13062
9557	Multiple ca	27/03/201	Delhi	Cochin	DEL → BOI	11:30	21:00	9h 30m	1 stop	No info	12537
9558	Jet Airway	15/03/201	Banglore	New Delhi	BLR → BOI	18:55	00:45 13 M	5h 50m	1 stop	No info	21934
9559	Jet Airway	24/06/201	Delhi	Cochin	DEL → BOI	13:00	19:00	6h	1 stop	In-flight m	10262
9560	Vistara	3/06/2019	Banglore	Delhi	BLR → DEL	17:00	19:35	2h 35m	non-stop	No info	4878
9561	Jet Airway	15/06/201	Delhi	Cochin	DEL → BOI	14:00	19:00	5h	1 stop	No info	14714

EXPECTED OUTCOME OF THE PROJECT

The developed machine learning model demonstrates high accuracy in predicting flight fares, with low mean absolute error (MAE) and root mean squared error (RMSE) values compared to baseline models. The user-friendly interface allows travellers to input their travel details and receive estimated flight fares quickly and efficiently. The flight fare prediction system provides tangible benefits such as cost savings, improved decision-making, and enhanced customer satisfaction for travellers, travel agencies, and airlines. Future enhancements may include integrating real-time data feeds, expanding coverage to include additional features, and ongoing maintenance to ensure continued effectiveness and relevance in the travel industry.

REFERENCES

1. "Predicting Airline Ticket Prices Using Historical Data" Authors: R. B. Patterson and R. M. Patel Published in: International Journal of Computer Applications, Vol. 52, No. 5, 2012.

2. Flight Fare Prediction using Historical Data and Machine Learning Techniques" Authors: A. Kumar, et al. Published in: Proceedings of the 3rd International Conference on Computer, Communication, and Signal Processing, 2019.

3. "Flight Fare Prediction Using Machine Learning Techniques" Authors: S. G. Sonawane and A. N. Kadam Published in: International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 7, Issue 7, 2017.

4. "Airline Fare Prediction Using Machine Learning" Authors: A. L. Rodrigues, et al. Published in: Proceedings of the International Conference on Data Engineering and Communication Technology, 2020.

5. "Flight Fare Prediction Using Machine Learning Techniques" Authors: R. N. Sahoo, B. Mishra, and S. P. Dash Published in: Proceedings of the 3rd International Conference on Computational Intelligence in Data Science (ICCIDS), 2019. DOI: 10.1109/ICCIDS.2019.9010314

6. "Predicting Airfare using Machine Learning Algorithms" Authors: K. Gupta and S. Agarwal Published in: International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 9, Issue 3, 2019.

Student Signature

Guide Signature
Signature

For Office Use Only

Approved/ Not Approved

Suggestions (if Any) ____

Project Convener Head of Department