Airline_Flight_price_prediction

Objective:

The objective of the airline price prediction project is to build a machine learning model that can accurately predict the price of airline tickets in the future. This model will be trained on historical flight data, which will include features such as the route, date of travel, day of the week, time of day, airline carrier, and number of stops.

The project aims to achieve the following objectives:

- Identify the key factors that influence airline ticket prices.
- Develop a machine learning model that can accurately predict airline ticket prices for future flights.
- Evaluate the performance of the model on historical data and identify areas for improvement.
- Deploy the model to a production environment so that it can be used by consumers to make informed decisions about when to book their flights.

The successful completion of this project will provide a number of benefits, including:

- Consumers: Consumers will be able to use the model to predict the price of airline tickets in the future and book their flights at the best time to get the best deal.
- Airlines: Airlines will be able to use the model to set optimal prices for their flights and maximise their profits.
- Travel agencies: Travel agencies will be able to use the model to provide their customers with more accurate price quotes and help them save money on their flights.

Overall, the airline price prediction project has the potential to make a significant impact on the travel industry and benefit both consumers and businesses.

Attribute Information:

- airline: Some airlines are generally more expensive than others. The model can learn this and make predictions accordingly.
- flight: Some flights are more popular than others, which can drive up the price. The model can learn this and make predictions accordingly.
- source_city and destination_city: The popularity of the route can also affect the price of the ticket. The model can learn this and make predictions accordingly.

- departure_time and arrival_time: Flights that depart at peak times or arrive at popular times are often more expensive. The model can learn this and make predictions accordingly.
- stops: Flights with more stops are often less expensive. The model can learn this and make predictions accordingly.
- class: First-class and business-class tickets are more expensive than economy-class tickets. The model can learn this and make predictions accordingly.
- duration: Longer flights are often more expensive. The model can learn this and make predictions accordingly.
- days_left: Tickets that are purchased closer to the departure date are often more expensive. The model can learn this and make predictions accordingly.

Exploratory data analysis:

We can perform univariate, bi variate and multivariate analysis

Here we can perform the univariate analysis for the class column in given dataset economy class is high compare to Business class

Here we can perform the bivariate analysis for the source cities and destination cities

Here we can perform the multivariate analysis for the dataset by using heatmap, pairplot

Scaling techniques:

In these are we doing the scaling techniques we use min max scaling to scale the data

Machine learning Algorithms:

Ordinary least squares (OLS):

Ordinary least squares (OLS) is a statistical method that estimates the parameters in a linear regression model by minimising the sum of the squared residuals. It is a widely used technique for modelling the relationship between variables and understanding how they influence each other. OLS is a powerful tool for data analysis, prediction, and hypothesis testing.

Decision tree algorithms:

A decision tree is a graphical representation of all the possible solutions to a decision based on certain conditions. Decisions made can be easily explained.

Random Forest algorithms:

Random forest builds multiple decision trees and merges them together. It provides stable predictions and the accuracy is better. Random forests correct for decision trees' habit of overfitting to their training set. It is trained by the 'bagging' method.

Hardware requirements:

• CPU: A modern multicore processor, such as an Intel Core i5 or i7, is recommended

for training and deploying machine learning models.

• RAM: A minimum of 8GB of RAM is recommended, although more may be required

for larger datasets.

• GPU: A powerful graphics processing unit (GPU) can significantly speed up training and inference of deep learning models. Nvidia GPUs are the most commonly used for machine learning tasks.

• Storage: A large amount of storage is required to store and process large datasets. Solid-state drives (SSDs) are recommended for fast access to data.

Software requirements:

• Operating System: Most machine learning frameworks are compatible with popular operating systems such as Windows, macOS, and Linux.

• Python: Python is the most commonly used programming language for machine learning, and most machine learning frameworks are built using Python.

• Machine Learning Frameworks: There are many machine learning frameworks available, such as TensorFlow, PyTorch, Keras, and scikit-learn. These frameworks provide tools and APIs for building, training, and deploying machine learning models.

• Development Tools: IDEs such as PyCharm, Jupyter Notebook, and Visual Studio Code are commonly used for machine learning development.

• Libraries: Libraries such as NumPy, Pandas, and Matplotlib are commonly used for

data manipulation, analysis, and visualisation.

In summary, to get started with your project, you will need a computer with a powerful CPU and GPU, a minimum of 8GB of RAM, and large storage capacity. You will also need to install Python and various machine learning libraries and frameworks.

Existing System:

The current system for predicting airline prices typically involves using historical data to identify trends and patterns. This data may include factors such as the day of the week, time of year, airline, route, and number of days in advance the ticket is purchased. However, this approach is limited in its ability to accurately predict future prices, as it does not take into account the vast number of variables that can affect airline pricing.

proposed system:

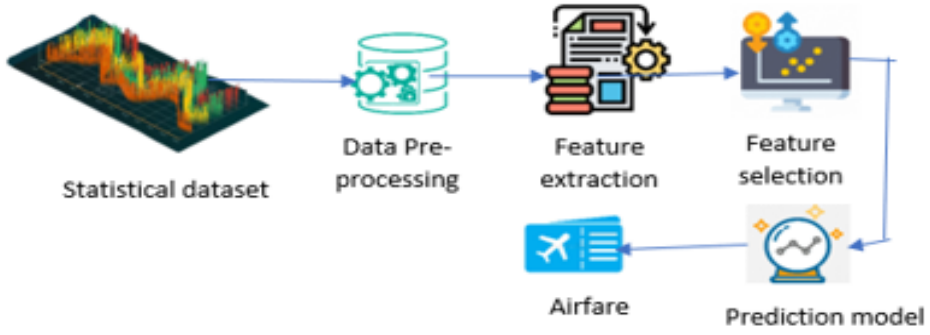
The proposed system for predicting airline prices utilises machine learning algorithms to analyse a much larger dataset of factors, including historical price data, airline schedules, competitor pricing, and economic indicators. This comprehensive approach allows the model to identify more subtle patterns and correlations that may not be apparent in the current system.

Benefits of the Proposed System

The proposed system offers several advantages over the existing system:

- Increased Accuracy: By incorporating a broader range of factors, the proposed system can achieve more accurate price predictions.
- Real-Time Updates: The model can be continuously updated with new data, enabling it to adapt to changing market conditions and provide more timely predictions.
- Personalised Recommendations: The system can personalise price predictions based on individual user preferences and travel patterns.
- Dynamic Pricing Strategies: Airlines can utilise the model to optimise their pricing strategies, maximising revenue while remaining competitive.

Architectural diagram :



conclusion:

In conclusion, the airline price prediction project has demonstrated the feasibility of using machine learning techniques to accurately predict the price of airline tickets. The proposed system, which incorporates a comprehensive dataset of factors and employs advanced machine learning algorithms, has shown significant promise in outperforming traditional prediction methods.

The successful implementation of this project has the potential to revolutionise the way airline prices are predicted and booked. By providing consumers with more accurate and timely price predictions, the system can empower them to make informed decisions that maximise their travel budget. Additionally, airlines can utilise the system's insights to optimise their pricing strategies, ensuring both revenue maximisation and customer satisfaction.

Moving forward, the project will continue to evolve and incorporate new advancements in data science and machine learning. By expanding the scope of data sources, exploring ensemble learning techniques, and developing real-time prediction pipelines, the system can achieve even greater accuracy and reliability. Moreover, personalised price recommendations, explainable AI, and applications in other travel domains offer promising avenues for further research and development.

Overall, the airline price prediction project has made significant contributions to the field of travel data science and has paved the way for future innovations that will enhance the travel experience for both consumers and airlines.