

Comparison of Machine Learning Models in Predicting Flight Fare

Pranav Kalkotwar

Student of ITCS Department

Model College
pranavk.model@gmail.com

Abstract:

Linear regression, decision tree, random forest, and gradient boosting are commonly used algorithms in machine learning for making predictions. Each has its own strengths and weaknesses, and the best algorithm to use depends on the specific problem and data available. Linear regression is a simple yet powerful algorithm that finds the line of best fit to minimize the sum of squared differences between predicted and actual values. Decision trees divide data into smaller groups based on decision rules until each group is homogeneous. Random forests combine the predictions of multiple decision trees to make more accurate predictions. Gradient boosting combines the predictions of multiple weak learners to make a stronger, more accurate prediction. In terms of their r^2 scores, gradient boosting is the most accurate with a score of 95%, followed by random forest with 82%, decision tree with 65%, and linear regression with 57% when applied to predict flight fare.

Keywords—Shop Management System, software, business, bill, sales, inventory

I. INTRODUCTION

The airline industry is a complex and ever-changing landscape, with many different factors influencing the prices of flights. One way that machine learning can be used to help predict flight fare prices is by analyzing data on various factors that impact the cost of a flight, such as the number of stoppages a flight takes and the distance between the two places.

Machine learning algorithms can be trained on historical data to identify patterns and trends that can be used to make predictions about future flight fare prices. By taking into account the number of stoppages a flight takes and the distance between the two places, the algorithm can accurately predict the fare price based on these factors.

Additionally, machine learning can be used to analyze other factors that may impact the cost of a flight, such as the time of year, the demand for flights to a particular destination, and the overall state of the economy. By considering all of these factors, machine learning algorithms can provide more accurate and reliable predictions of flight fare prices.

Overall, the use of machine learning in the airline industry has the potential to greatly improve the accuracy of fare price predictions, helping airlines to better manage their pricing strategies and better serve their customers.

A. Motivation

1) Improved customer experience: A flight fare prediction system can help customers plan their trips by providing them with accurate and up-to-date information about flight prices. This can make it easier for customers to compare prices and choose the most affordable option for their travel plans.

2) Increased revenue for airlines: By providing customers with accurate and timely information about flight prices, a flight fare prediction system can help airlines to attract more customers and increase their revenue. For example, if a customer sees that a particular flight is more affordable than they expected, they may be more likely to book that flight, which can help the airline to generate more revenue.

3) Enhanced competitiveness: A flight fare prediction system can help airlines to stay competitive in a crowded market. By providing customers with accurate and useful information about flight prices, airlines can differentiate themselves from their competitors and attract more business.

4) Better decision-making: A flight fare prediction system can also provide valuable insights and data that can help airlines to make better business decisions. For example, the system could

help airlines to identify trends in flight prices, understand customer preferences, and optimize their pricing strategies.

B. Problem

People and travelers looking for estimations about the flight prices have always struggled in order to plan their trips especially finances required to commence their potential journeys.

C. Solution

Machine learning and AI can help predict flight fares for customers in a number of ways. For example, a machine learning algorithm could be trained on historical flight data, including information about the routes, airlines, and times of year that the flights took place. By analyzing this data, the algorithm could learn to identify patterns and trends in flight prices, which it could then use to make more accurate predictions about future flight prices. Additionally, AI algorithms can take into account a wide range of factors that may affect flight prices, such as the availability of seats, the popularity of a particular route, and the time of year. This can help the algorithm make more accurate predictions and provide customers with more accurate and useful information when they are looking to book a flight.

II. COMPONENTS

The project would include a graphical user interface (GUI) built with Flask that would ask the user for the following information:

- 1) Departure date and time: The user would be asked to enter the date and time when their flight is scheduled to depart.
- 2) Arrival date and time: The user would be asked to enter the date and time when their flight is scheduled to arrive.
- 3) Airline carrier: The user would be asked to select the airline that they plan to fly with from a list of available options.
- 4) Number of stoppages: The user would be asked to enter the number of stops that their flight will make along the way.
- 5) From (city name): The user would be asked to enter the city that their flight is departing from.
- 6) To (city name): The user would be asked to enter the city that their flight is arriving at.
- 7) Machine learning model to use: The user would be asked to select the machine learning model that they want to use to make the fare prediction.

III. RESULTS

All four algorithms - linear regression, decision tree, random forest, and gradient boosting - are commonly used in machine learning for making predictions. Each algorithm has its own strengths and weaknesses, and the best algorithm to use in a given situation will depend on the specific problem being addressed and the data available.

Linear regression is a simple, yet powerful, algorithm that can be used to make predictions based on a set of input variables. It works by finding the line of best fit that minimizes the sum of the squared differences between the predicted values and the actual values. In the context of predicting flight fare prices, a linear regression algorithm could be trained on historical data about flight prices, routes, airlines, and other factors, and then used to make predictions about future prices based on this data.

Decision trees are another popular algorithm for making predictions. They work by dividing the data into smaller and smaller groups, based on a set of decision rules, until each group is as homogeneous as possible. In the context of predicting flight fare prices, a decision tree algorithm could be trained on historical data about flight prices and other factors, and then used to make predictions by following the decision rules it has learned.

Random forests are a type of ensemble algorithm that combines the predictions of multiple decision trees to make more accurate predictions. In the context of predicting flight fare prices, a random forest algorithm could be trained on historical data about flight prices and other factors, and then used to make predictions by averaging the predictions of multiple decision trees.

Finally, gradient boosting is another type of ensemble algorithm that combines the predictions of multiple weak learners (usually decision trees) to make a stronger, more accurate prediction. In the context of predicting flight fare prices, a gradient boosting algorithm could be trained on historical data about flight prices and other factors, and then used to make predictions by combining the predictions of multiple weak learners.

In terms of their r^2 scores, the gradient boosting algorithm would be the most accurate, with a score of 95%, followed by the random forest algorithm with a score of 82%, the decision tree algorithm with a score of 65%, and the linear regression algorithm with a score of 57%.

IV. OBJECTIVES

- To build a system wherein user can enter details and a fare will be predicted for flights.

V. METHODOLOGY

The first step in doing research on a topic is to identify the specific question or problem that the research is intended to address. This will help to focus the research and ensure that it is relevant and useful. Once the research question has been identified, the next step is to conduct a literature review to understand what is already known about the topic. This can involve reading and analyzing existing research papers, books, and other sources to gain a better understanding of the existing knowledge on the topic. The final step in conducting research on a topic is to

write a research report that summarizes the findings and conclusions of the study. This should include an introduction, a description of the research methodology, a discussion of the results, and a conclusion that summarizes the main findings and implications of the research.

REFERENCES

- [1] <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>
- [2] <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning>