FLIGHT FARE PREDICTION

ATHIRA KP
08/07/2024

TABLE OF CONTENTS

OBJECTIVE

The aim of our study is to answer the below research questions:

a) Does price vary with Airlines?

b) How is the price affected when tickets are bought in just 1 or 2 days before departure?

c) Does ticket price change based on the departure time and arrival time?

d) How the price changes with change in Source and Destination?

e) How does the ticket price vary between Economy and Business class?

OVERVIEW OF THE PROBLEM STATEMENT

An airline is a company that provides air transport services for traveling passengers and freight. Airlines use aircraft to supply these services and may form partnerships or alliances with other airlines for codeshare agreements, in which they both offer and operate the same flight. Generally, airline companies are recognized with an air operating certificate or license issued by a governmental aviation body. Airlines may be scheduled or charter operators.

Airlines assign prices to their services in an attempt to maximize profitability. The pricing of airline tickets has become increasingly complicated over the years and is now largely determined by computerized yield management systems.

The price of an Airline Ticket is affected by a number of factors, such as flight duration, days left for departure, arrival time and departure time etc. Airline organizations may diminish the cost at the time they need to build the market and at the time when the tickets are less accessible. They may maximize the costs. The price may rely upon different factors. Each factor has its own proprietary rules and algorithms to set the price accordingly. Recent advances in Artificial Intelligence (AI) and Machine Learning (ML) makes it possible to infer such rules and model the price variation.

DATA COLLECTION AND METHODOLOGY

Dataset contains information about flight booking options from the website kaggle.com for flight travel between India's top 6 metro cities. There are 300261 datapoints and 11 features in the cleaned dataset.

- kaggle/input/flight-price-prediction/Clean_Dataset.csv

This is the processed dataset obtained after merging the economy and business dataset and performing basic feature transformation.

FEATURES

The various features of the cleaned dataset are explained below:

1) Airline:
The name of the airline company is stored in the airline column. It is a categorical feature having 6 different airlines.
2) Flight:
Flight stores information regarding the plane's flight code. It is a categorical feature.
3) Source City:
City from which the flight takes off. It is a categorical feature having 6 unique cities.
4) Departure Time:
This is a derived categorical feature obtained created by grouping time periods into bins. It stores information about the departure time and have 6 unique time labels.
5) Stops:
A categorical feature with 3 distinct values that stores the number of stops between the source and destination cities.
6) Arrival Time:
This is a derived categorical feature created by grouping time intervals into bins. It has six distinct time labels and keeps information about the arrival time.
7) Destination City:
City where the flight will land. It is a categorical feature having 6 unique cities.
8) Class:
A categorical feature that contains information on seat class; it has two distinct values: Business and Economy.
9) Duration:
A continuous feature that displays the overall amount of time it takes to travel between cities in hours.
10)Days Left:
This is a derived characteristic that is calculated by subtracting the trip date by the booking date.
11) Price:
Target variable stores information of the ticket price.

—---