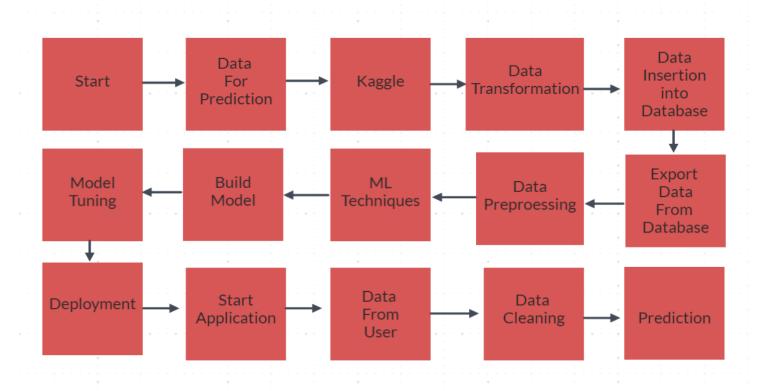
Flight Fare Prediction

Objective:

- ▶ The project aims to create a flight fare price prediction system that harnesses machine learning techniques to analyze historical flight data, traveler preferences, and dynamic market conditions.
- ▶ By doing so, we intend to empower travelers with reliable estimates of airfare, facilitating budget-conscious travel planning. This system's objective is to enhance travel affordability, enabling users to make well-informed decisions while promoting transparency in the airline pricing landscape.

Architecture:



Data Insertion into Database:

- After Transformation of all data I insert data into database and for that I used MySql Workbench
- ► Table Creation: I create Table Called flight_price_data into database called same as table name.
- Insertion Of Data: After Creating Table all Transformed Data are inserted into Table.

Model Training:

- Export Data from DB: I Export Data from mysql workbench in CSV form for training the model.
- Impute: I replace Null value with median and mode of numerical and categorical column respectively.
- Encoder: I encode some column using One hot encoder. Like Airline, source and Destination.
- ▶ Remove Duplicate: I remove All Duplicate rows from the Dataset.
- Normalize : To normalize Data I used Standard Scaler which converts any data using z-score

Model Selection :

- ▶ I Took all Models MI models and train on data and and I got top three model. Catboost, Xgboost, gradient Boost Among these Three cateboost Gave highest R2 Score on Test Data
- ► Therefore I select catboost for my Project

► Model Tuning:

▶ After Selecting catboost model I done Hyperparameter Tuning on catboost in which I used Randomized search cv with 30 iteration.

Prediction:

- ▶ For Prediction I make Prediction Pipeline which takes data from user.
- first it Preprocessed that by giving the raw data to preprocesser object which returns preprocessed data
- After getting processed data it gives the data to MI model object which predict the result on given data.
- After predicting, the result is returns to user.

Q&A

- ▶ Q1) What is the source of Data?
 - ▶ The data for training is Given in Project Details. which also Available on Kaggle.
- Q2) What is the type of Data?
 - ▶ The data is mix of Categorical and numerical values.
- Q3) What is the Complete Flow you followed in this Project ?
 - Refer Slide 3 for better understanding

- ▶ Q4) How Logs are Managed ?
 - ▶ I make one Logger file which is responsible for all logs fires in this project.
- Q5) What techniques Were you used for data Pre-processing?
 - ▶ Please refer slide 5 for better understanding.
- Q6) How training was done or What models are used ?
 - After processing a data, data are given to all models for training and predicting.
 - In which catboost, xgboost performed very well on test data with highest R2 Score.

- ▶ Q7) How Prediction Was done?
 - ▶ Please refer slide 7 for better understanding.
- ▶ Q8) Where you Deployed your project?
 - ► I used <u>Render</u> for Deploy a Project which is free and Easy for deploying a machine learning Models.
- Q9) How can I reach project you deployed?
 - ▶ Please follow the link: https://flight-fare-price-pred.onrender.com

Thank you!!!