

# Flight Ticket Price Predictor Using Machine Learning

## Overview

This project introduces a machine learning-based system to predict flight ticket prices by analyzing key factors that influence airfare trends. The aim is to create a user-friendly platform that helps travelers plan their trips cost-effectively.

### 1. Understanding Key Factors:

- Examines elements affecting flight ticket prices such as locations, travel dates, airlines, seasonal trends, and route popularity.
- Integrates market situations and customer behavior data for accurate predictions.

### 2. Utilization of Machine Learning Models:

- Algorithms used:
- Linear Regression, Random Forest, XGBoost, SVM, and Neural Networks.

### 3. Programming Implementation:

- Implemented in Python using Pandas, Scikit-learn, TensorFlow, and XGBoost.
- Sample Code: 

```
```python
import pandas as pd
from sklearn.model_selection import
train_test_split
from sklearn.ensemble import RandomForestRegressor
```

```
df = pd.read_csv("flight_prices.csv")
```

```
X = df[["airline", "departure", "arrival", "duration", "stops", "date"]]
```

```
y = df["price"]
```

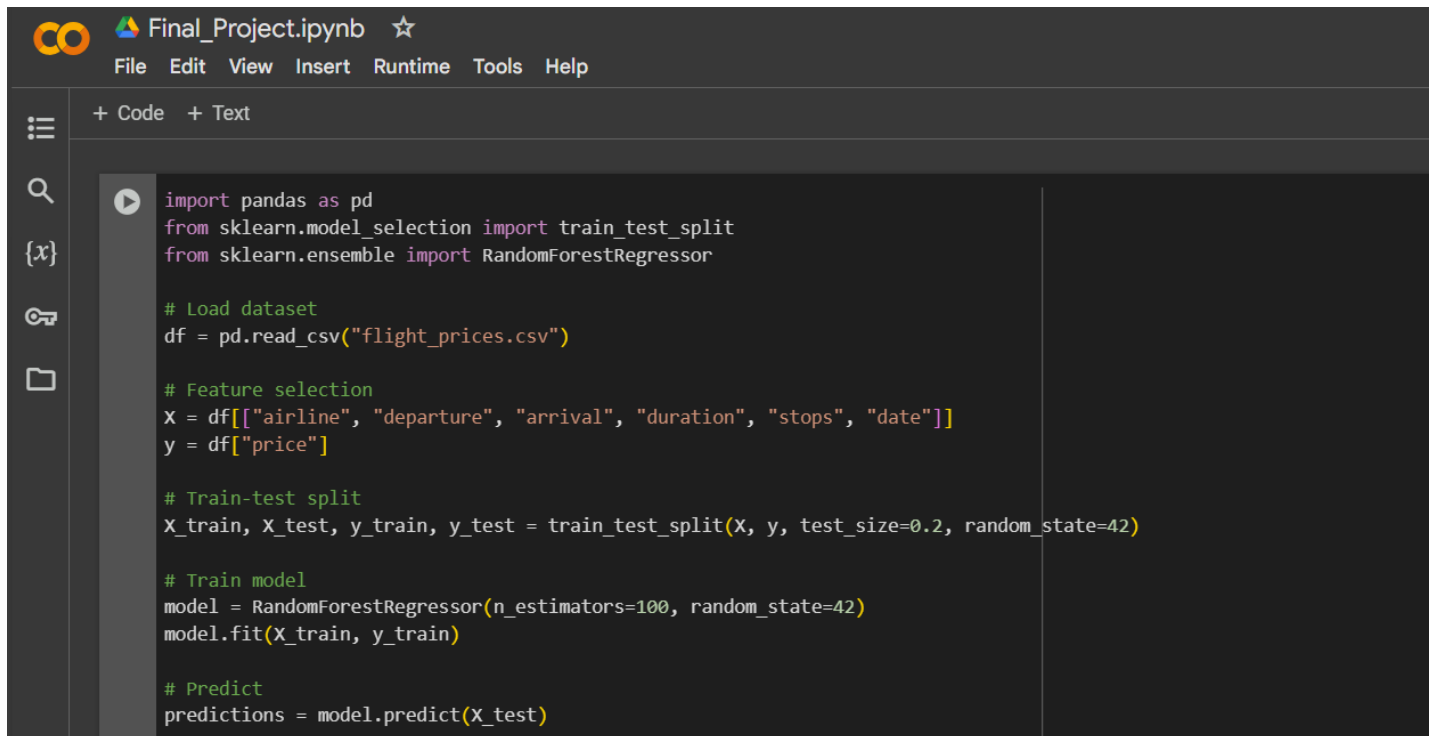
## Flight Ticket Price Predictor Using Machine Learning

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
model = RandomForestRegressor(n_estimators=100, random_state=42)
```

```
model.fit(X_train, y_train)
```

```
predictions = model.predict(X_test)
```

A screenshot of a Jupyter Notebook interface. The top bar shows the file name 'Final\_Project.ipynb' and a star icon. Below it is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. The left sidebar contains icons for a menu, search, a variable '{x}', a key, and a folder. The main area is a code editor with a dark background and light-colored text. It contains Python code for loading a dataset, selecting features, splitting data, training a model, and making predictions. The code is as follows:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor

# Load dataset
df = pd.read_csv("flight_prices.csv")

# Feature selection
x = df[["airline", "departure", "arrival", "duration", "stops", "date"]]
y = df["price"]

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Train model
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Predict
predictions = model.predict(X_test)
```

### 4. Dataset Description:

- Includes flight price records with details like airline name, departure/arrival locations, flightduration, number of stops, travel dates, and seasonal indicators.
- Datasets used:
  - Kaggle Flight Price Dataset
  - OpenSky Network Data
  - Airline APIs (Skyscanner, Amadeus)

## **Flight Ticket Price Predictor Using Machine Learning**

### **5. Prediction Approach:**

- Data preprocessing includes handling missing values, encoding variables, and feature scaling.
- Model training using cross-validation for performance optimization.
- Predicts ticket prices based on input features, with ensemble techniques enhancing accuracy.

### **6. Practical Benefits for Users:**

- Provides insights on the best times to book flights and save money.
- Adapts to market changes for real-time accuracy and reliability.

### **7. Expected Outcomes:**

- Fast and reliable price predictions using advanced data analysis.
- Improved user experience through a user-friendly interface.