Flight Delay Prediction

MACHINE LEARNING PROJECT

Nicholas Kuok Jin Shung
FCS12124

7 January, 2026

PROBLEM STATMENT

Flight delays often happen without warning, making it hard for passengers and airlines to plan their schedules.

Stakeholders :

- Passengers
- Airlines
- Airport Operators

Impact :

Unpredicted delays cause long waiting times for passengers and increase operational costs for airlines.

Data Overview

Source :

Kaggle flight delay dataset (2024)

Granularity :

One row represents one flight with details

Size :

1,048,575 rows and 18 columns

Target Variable:

Weather Delay & Late Aircraft Delay

Objectives & Key Questions

Objectives :

- Build a machine learning to predict flight delays
- Use flight information to classify flights as delayed or on time

Key Questions :

- Can flights data predict delays ?
- How well does the model identify delayed flights ?

Methodology

Data Preprocessing

- Dropped irrelevent columns (fl_date , wheels_off , wheels_on)
- Filtered canceled flight (cancelled == 0)
- Filled missing numeric values with column means (dep_time , taxi_out , air_time , taxi_in)

Feature Selection

- Input : month , day_of_week , dep_time , taxi_out , air_time , distance
- Target : Binary all the delay (1 = Delayed , 0 = On-Time)

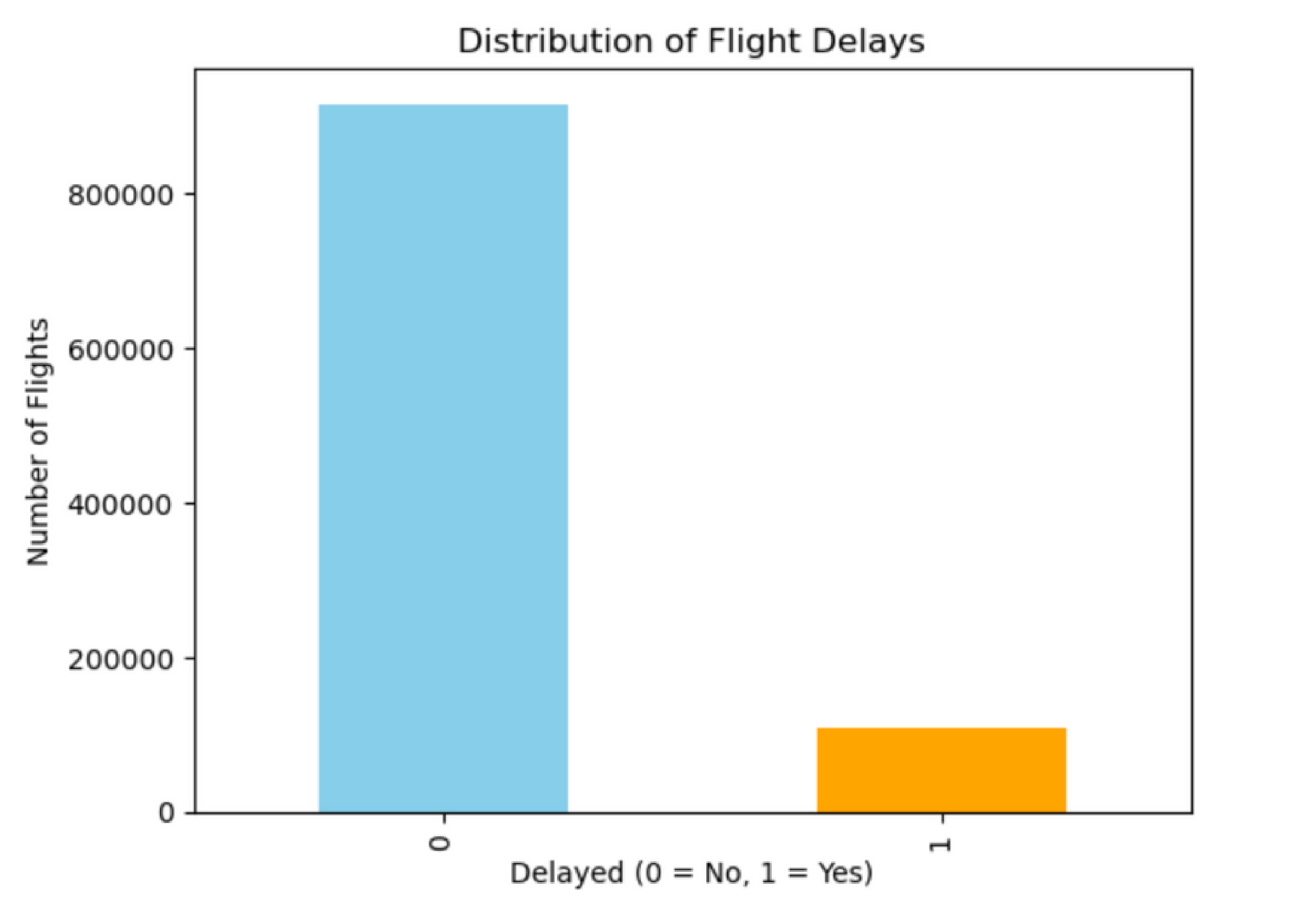
Modeling

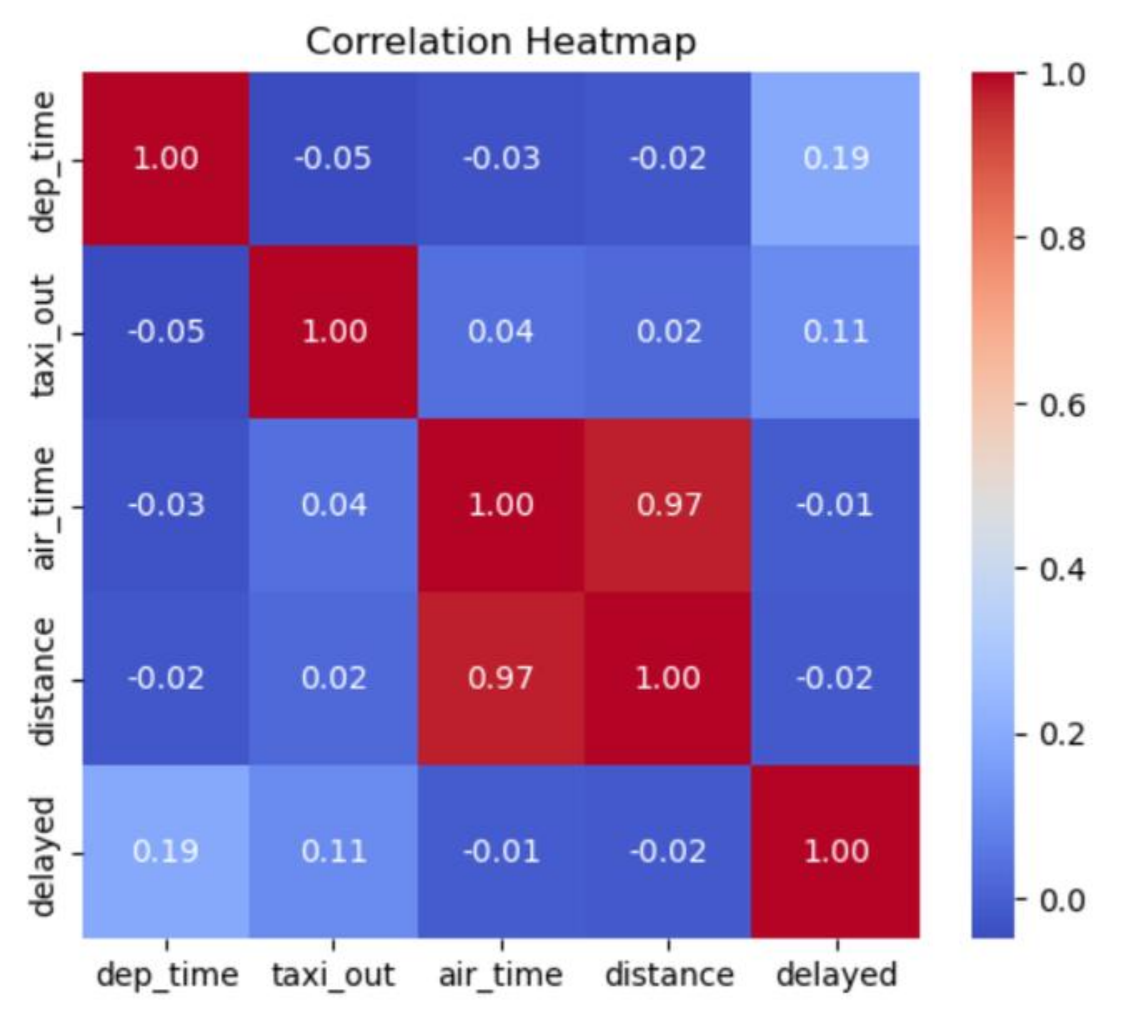
- Split dataset :
  80% train / 20% test
- Algorithm :
  Logistic Regression
  (max_iter = 1000 ,
  class_weight ='balanced')

Evaluation & Deployment

- Metrics : Accuracy , Confusion Matrix , F1 Score
- Deployment : Gradio App

Eda Key Finding





	dep_time	taxi_out	air_time	distance	delayed
dep_time	1.00	-0.05	-0.03	-0.02	0.19
taxi_out	-0.05	1.00	0.04	0.02	0.11
air_time	-0.03	0.04	1.00	0.97	-0.01
distance	-0.02	0.02	0.97	1.00	-0.02
delayed	0.19	0.11	-0.01	-0.02	1.00

Evidence : Most flights are on time , fewer are delayed .

Interpretation : The dataset is unbalanced

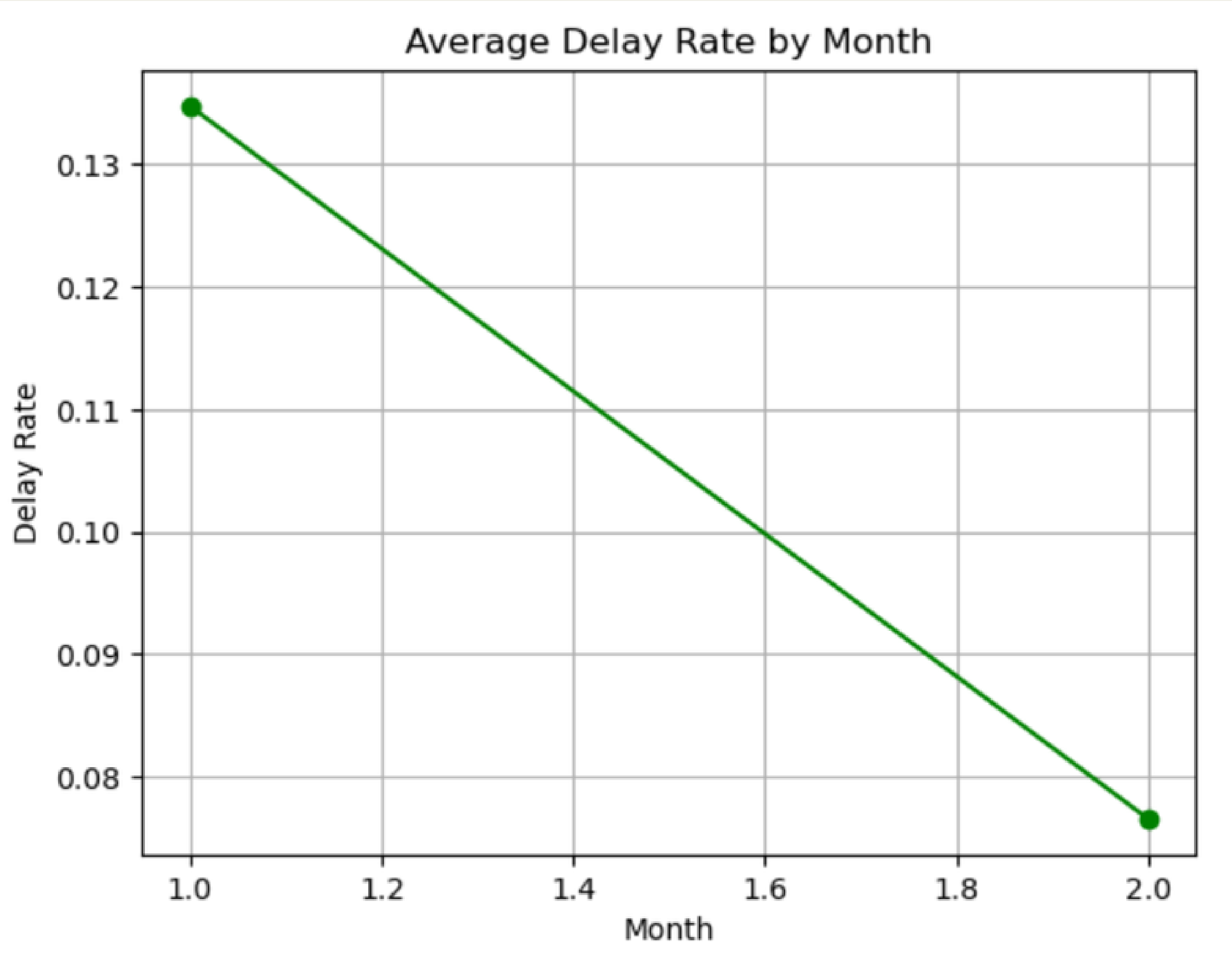
Action : Use F1-score and confusion matrix to check model , not just accuracy

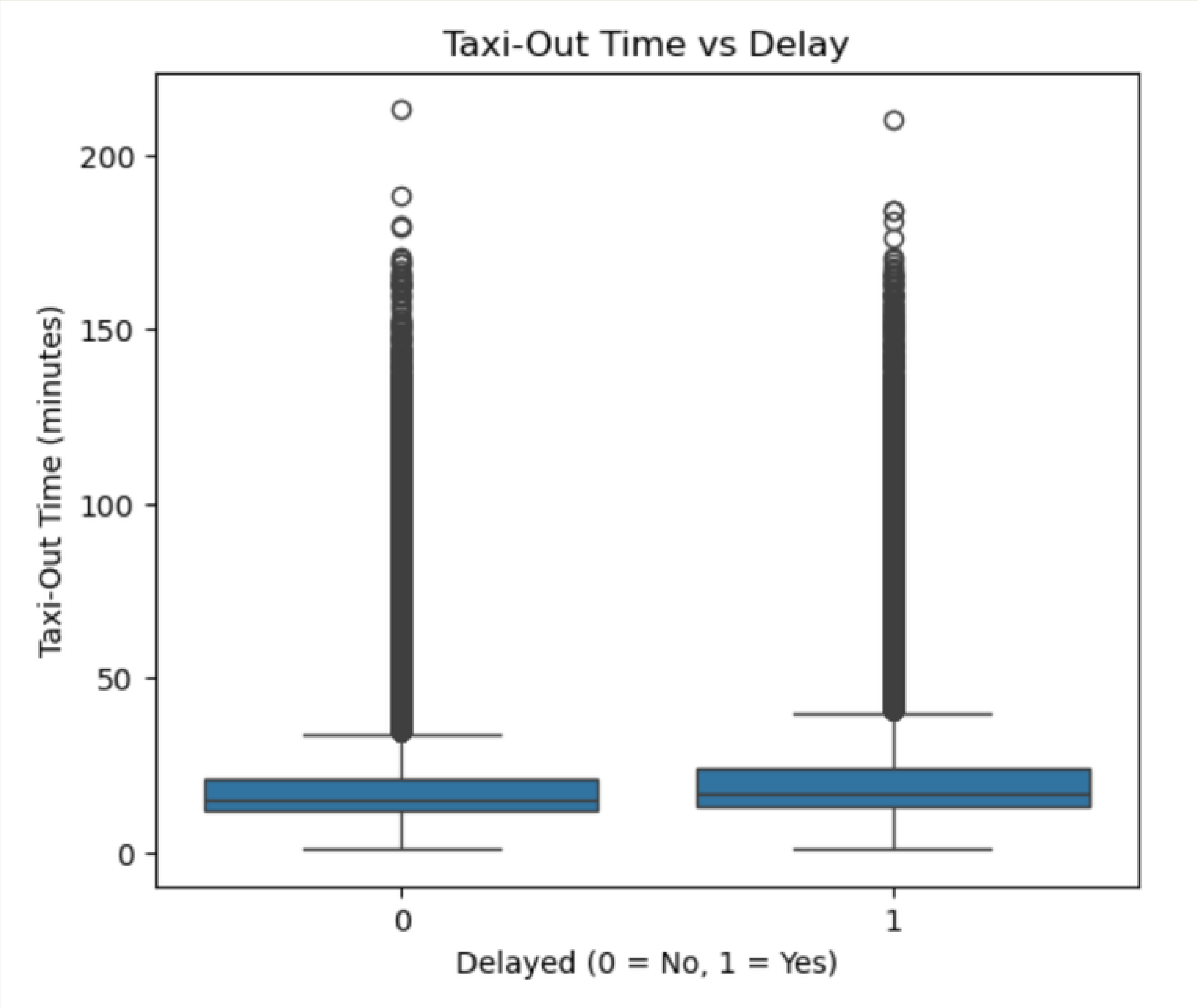
Evidence : shows **taxi_out** and **air_time** are moderately correlated with delay .

Interpretation : These features affect the chance of delay

Action : Keep these features for the model

Eda Key Finding





Evidence : Some **month** have more delays than others

Interpretation : Delays depend on the month

Action : Include **month** as a feature in the model

Evidence : Longer **taxi_out** times usually mean more delays .

Interpretation : Taxi_out time is an important factor .

Action : Include **taxi_out** as a feature in the model

Modeling Approch

Algorithms :

Used **Logistic Regression** because it works well for predicting yes/no outcomes

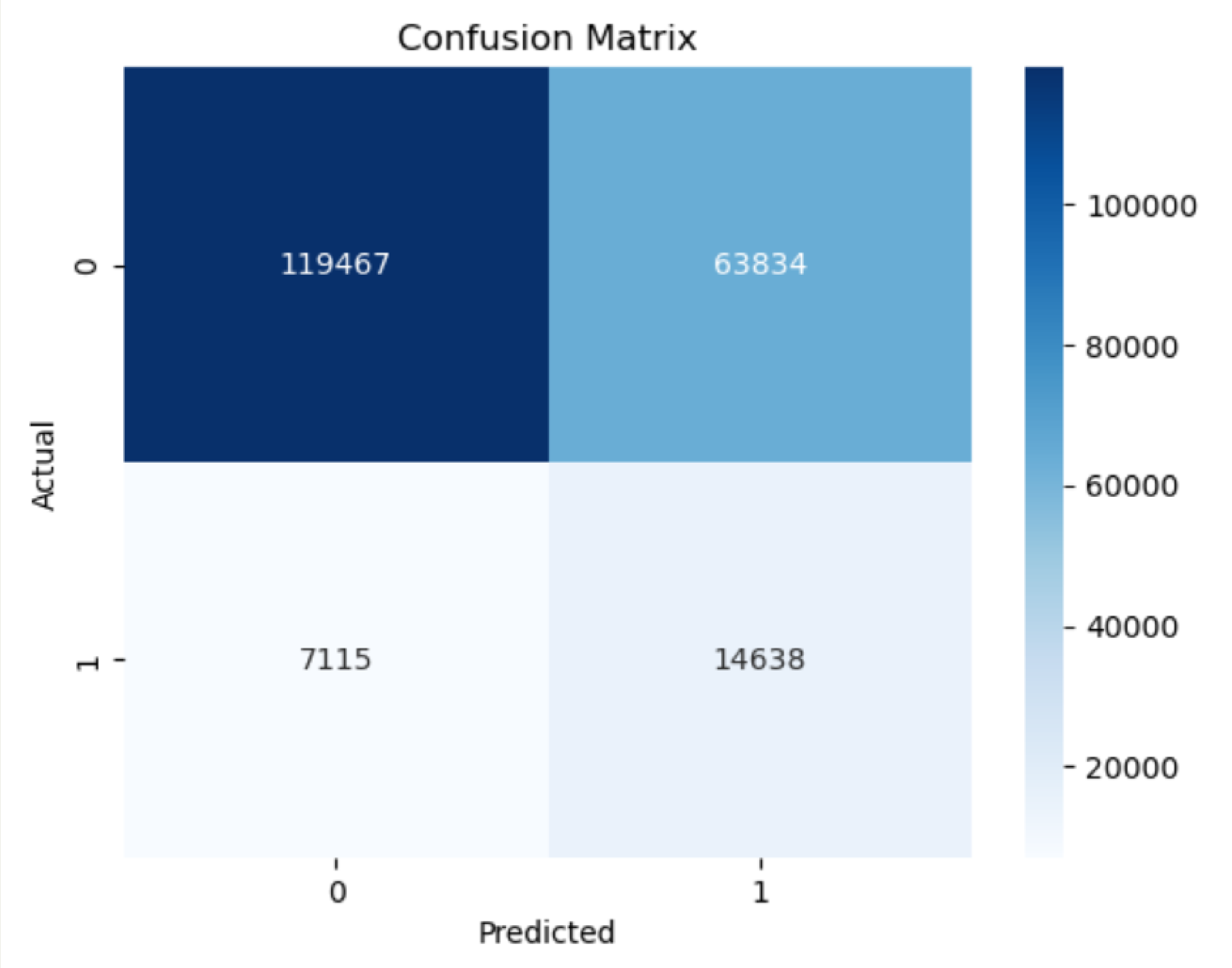
Validation :

Split data into 80% traning and 20% testing

Feature Engineering :

- Converted departure time to numbers
- Filled missing values with the mean
- All features are numeric , so no encoding needed

Result & Evaluation



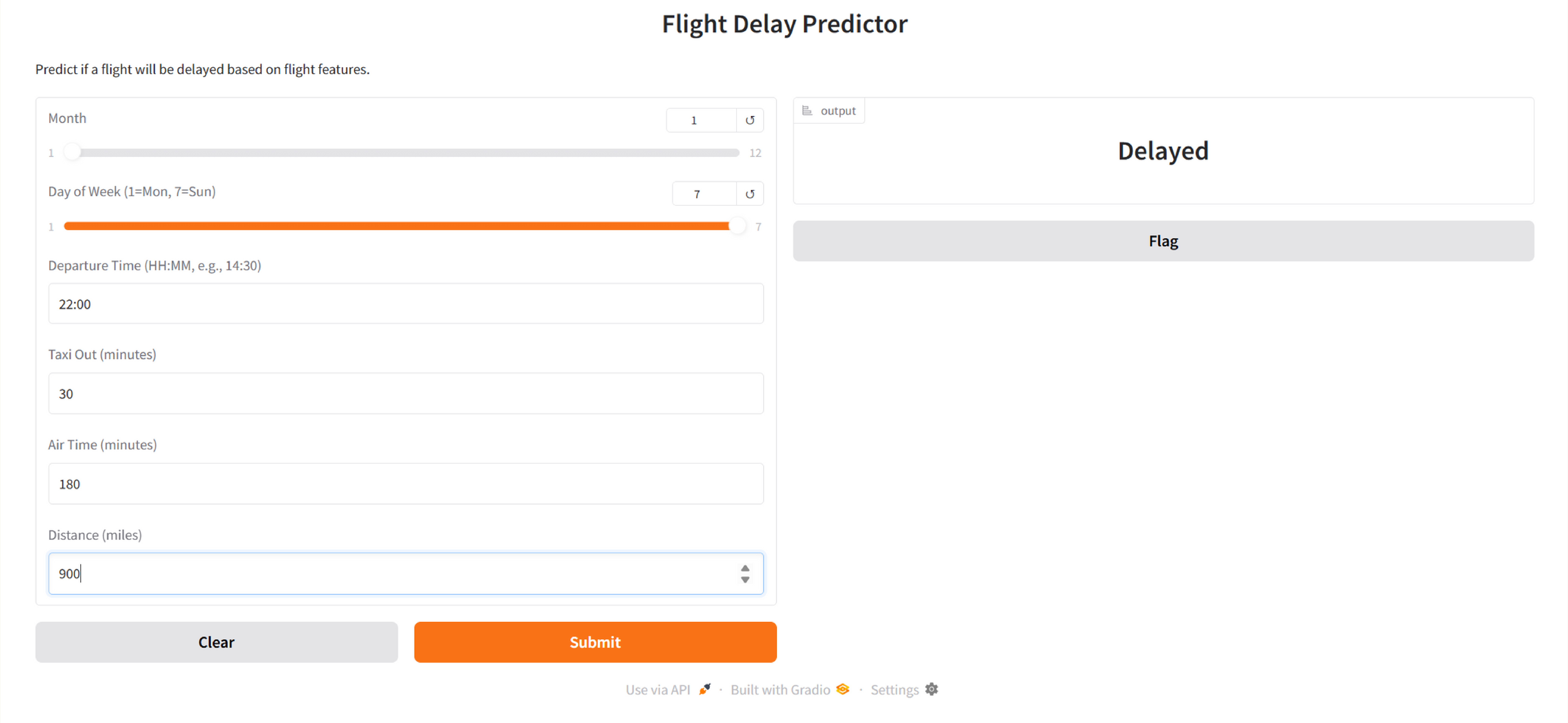
Primary Metrics :
Accuracy : 65.40% → Most flights were predicted correctly as on-time or delayed
F1 Score : 0.29 → The model is able to detect delayed flights , but performance is still moderate due to class imbalance

“So What” ? :

- The model can correctly predict most on-time flights , which is useful for planning and scheduling .

Project Demo

http://127.0.0.1:7861/



Flight Delay Predictor

Predict if a flight will be delayed based on flight features.

Month: 1

Day of Week (1=Mon, 7=Sun): 7

Departure Time (HH:MM, e.g., 14:30): 22:00

Taxi Out (minutes): 30

Air Time (minutes): 180

Distance (miles): 900

output: Delayed

Measure of Success

The model achieved an F1 Score of 0.29 , which shows it can identify delayed flights , but there is still room for improvement .

- The model can identify flights likely to be delayed
- This helps airlines plan better and reduce passenger complaints

Challanges & Limitations

- **The dataset has fewer delayes flights compared to on-Time flights**
  - focused on evaluation metrics like accuracy and confusion matrix instead of accuracy alone .

- Some data values were missing and had to be filled using the mean of each feature

- Only a limited number of features were used in the model

Future Work & Recommendations

- **Use more data such as detailed weather information and airport traffic data to improve prediction accuracy**

- **Try more advanced models (Random Forest or XGBoost)**

Tech Stack

Language : Python

Libraries : Pandas , Scikit-Learn , Joblib , Gradio , Matplotlib/Seaborn

Infastructure : Github , Gradio

Thank you!