**Project Report: Flight Delays Prediction Using Machine Learning**

**1. Introduction**

This project aims to predict whether a flight will be delayed using machine learning models trained on synthetic or real-world flight data. It is deployed as a Flask web application, enabling users to input flight details and receive a delay prediction. The system supports improved travel planning, operational efficiency for airlines, and decision-making for airport authorities.

**2. Objectives**

* Predict flight delays greater than 15 minutes.
* Develop a web-based interface for delay prediction.
* Train a machine learning model using historical flight data.
* Demonstrate model performance and usability.

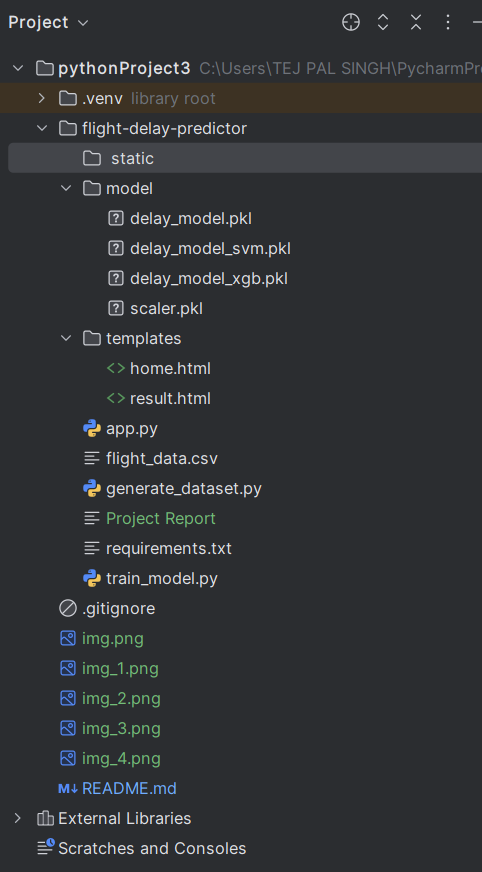
**3. Features**

* Predicts delays based on input flight parameters (date, airline, origin/destination, time).
* User-friendly web interface using Flask.
* Random Forest Classifier used for classification.
* Synthetic dataset generation capability.
* Modular code structure with model training, data generation, and app deployment separated.

**4. Tools and Technologies Used**

* **Programming Language:** Python 3
* **Web Framework:** Flask
* **Machine Learning:** Scikit-learn, Random Forest
* **Data Handling:** Pandas, NumPy
* **Serialization:** Joblib

**5. Project Structure**

****

**6. How to Run the Project**

1. **Clone the Repository:**

git clone https://github.com/yourusername/flight-delay-predictor.git

cd flight-delay-predictor

1. **Install Dependencies:**

pip install -r requirements.txt

Or manually:

pip install flask pandas scikit-learn joblib numpy

1. **Generate Dataset (Optional):**

python generate\_dataset.py

1. **Train the Model:**

python train\_model.py

1. **Run the Web App:**

python app.py

Visit: [http://localhost:5000](http://localhost:5000/)

**7. Sample Prediction Inputs**

* Month: 7
* Day: 15
* Day of Week: 5 (Friday)
* Airline: AA
* Origin: ATL
* Destination: JFK
* Scheduled Departure: 0930

**8. Results and Use Cases**

* **Travelers**: Can assess flight reliability before booking.
* **Airlines**: Can optimize resource allocation and scheduling.
* **Airport Authorities**: Can forecast and mitigate delays and congestion.

**9. Challenges Faced**

* Generating realistic synthetic data for initial training.
* Model accuracy fluctuating with small or unbalanced datasets.
* Building an intuitive yet informative web interface.

**10. Future Enhancements**

* Integrate real-time APIs (weather, traffic).
* Display prediction probabilities and feature importance.
* Add visual analytics dashboard.
* Deploy on cloud platforms like Heroku or Render.

**11. License**

This project is intended for educational and academic use only.

**12. GitHub Repository**

<https://github.com/kimrahd3782/flight-delay-predictor>

*Submitted by: Goonj Chaudhary*  
*Registration Number: 23bce11504*