Flight Delay Forecasting Using Machine Learning

A Data-Driven Approach to Improve Operational Efficiency in Air Travel

Name: Minaksi Yadav

Date: June 2025

Problem Overview:
Delays in air travel remain a critical bottleneck, leading to disrupted passenger experiences and increased operational overhead for airlines. These disruptions can escalate fuel consumption, hinder crew scheduling, and negatively impact service quality and reliability.

Project Goals:
This project utilizes past flight performance data to:
- **Reveal underlying delay trends** through targeted exploratory analysis
- **Develop a dual-phase ML system** to predict delay likelihood and expected delay time
- **Deliver strategic recommendations** enabling airlines to proactively minimize delays and optimize efficiency

Dataset Overview & Preprocessing Steps

Preprocessing Steps:

- **Dropped Unnecessary Columns**: Removed year, carrier_name , airport_name.

- **Encoded Categorical Features**: Transformed carrier and airport using frequency encoding

- **Handled Missing Data**:
    - Eliminated rows with multiple missing fields
    - Filled arr_del15 using median values

- **Outlier Processing**:
    - Used IQR capping per airport for numerical delay fields

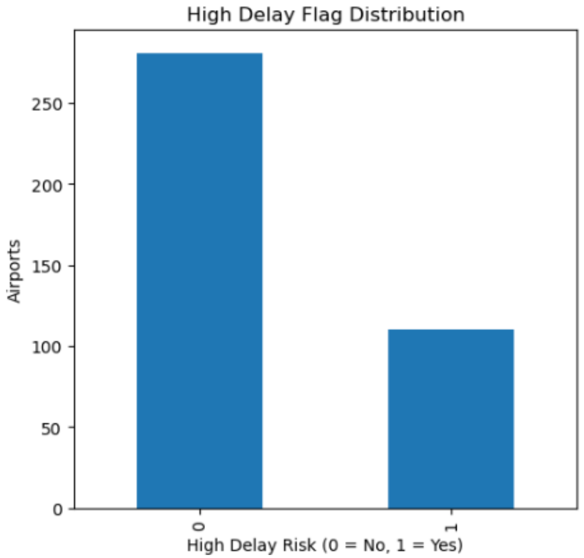
- **Filtered Sparse Airport Data**: Removed rare airports having freq=1

- **Created Custom Target Metric**:
    - Defined **OAI (Operational Adjustability Index)** using delays, cancellations, and encoded features

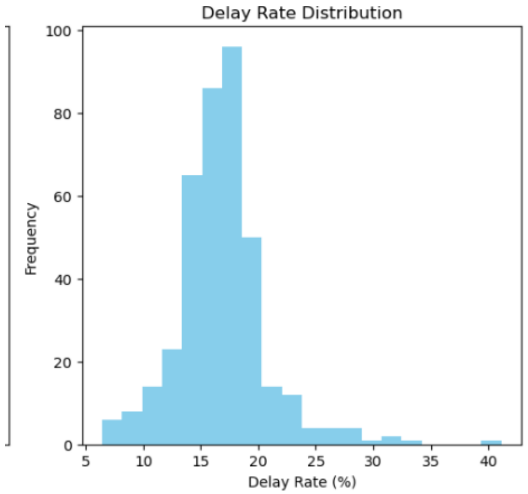
EDA from next page :

Key Variables:

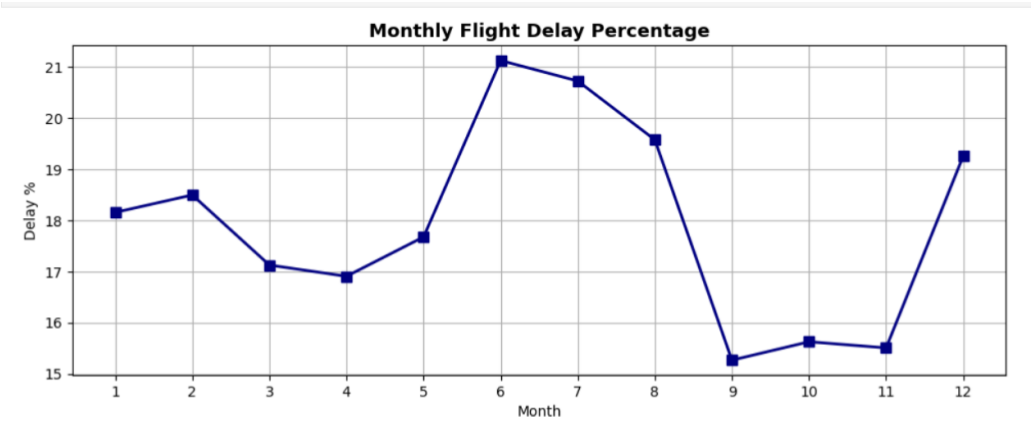
- **Delay Counts**: carrier_ct, weather_ct, nas_ct, security_ct, late_aircraft_ct
- **Delay Durations**: carrier_delay, weather_delay, nas_delay, security_delay, late_aircraft_delay
- **Flight Information**: arr_flights, arr_del15, arr_cancelled, arr_diverted, arr_delay, month, carrier, airport



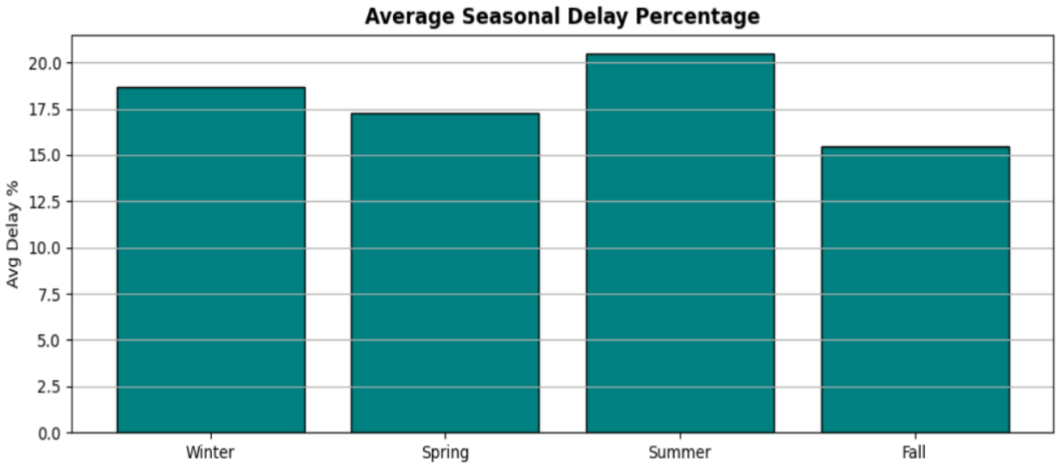
- **Majority of airports (≈70%)** show **low delay risk**.
- Only a **smaller portion (~30%)** are flagged as **high delay risk**.
- Highlights an **imbalance**, indicating that **delays are concentrated** at specific airports.
- Suggests potential for **targeted interventions** rather than broad policy changes.

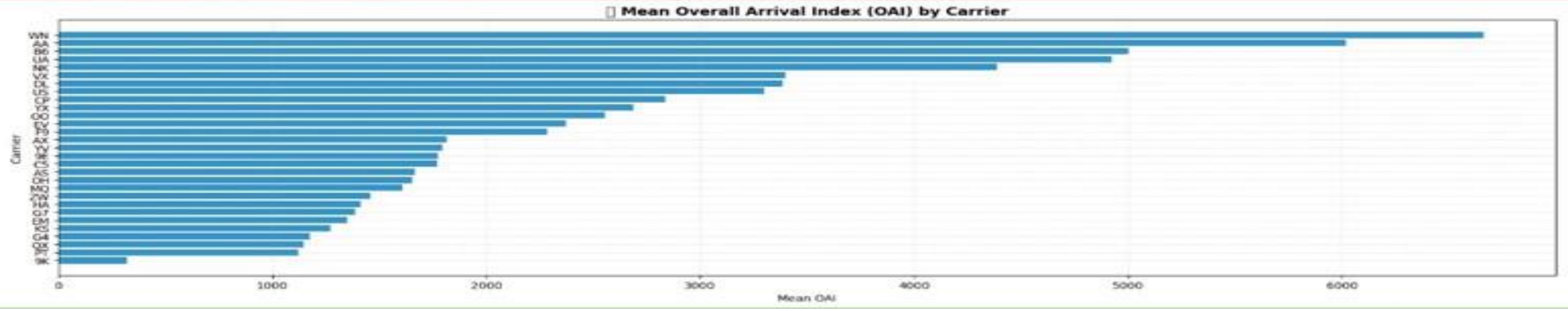


- Most airports have a **delay rate between 13% and 20%**, indicating a typical moderate delay trend.
- **Few airports exceed 25% delay rate**, suggesting high delays are rare but significant.
- The distribution is **right-skewed**, highlighting a concentration of airports with lower delay rates.



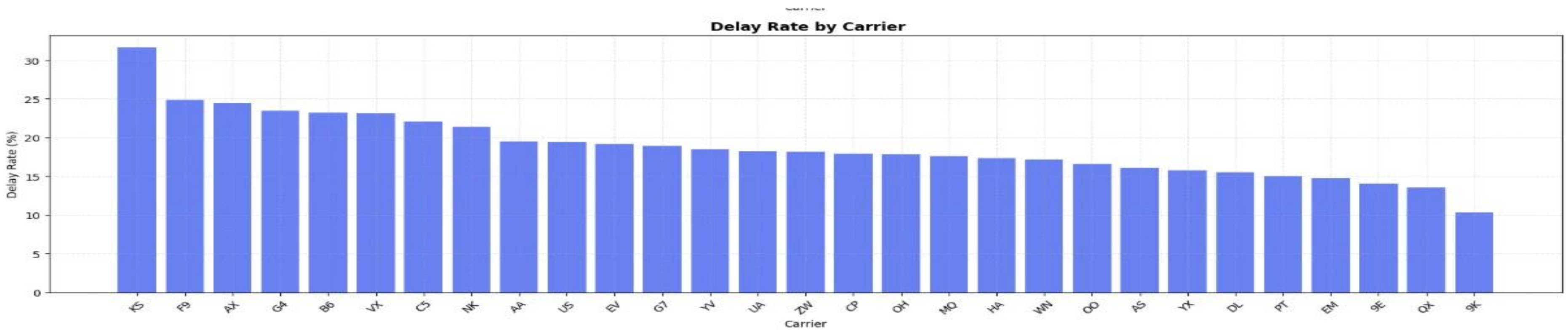
- **June and July** experience the **highest delay percentages** (over 21% and 20%), likely due to peak summer travel.
- **September and November** show the **lowest delays**, indicating more efficient operations or lower traffic.





WN (Southwest Airlines) has the highest mean Overall Arrival Index (OAI), indicating strong performance in on-time arrivals.

- **9K** shows the lowest OAI, suggesting relatively poor arrival punctuality among carriers



- **KS** and **F9** carriers have the highest delay rates, exceeding **30%** and **25%** respectively.
- **9K**, **QX**, and **9E** exhibit the lowest delay rates, staying under **15%**, indicating more reliable on-time performance.

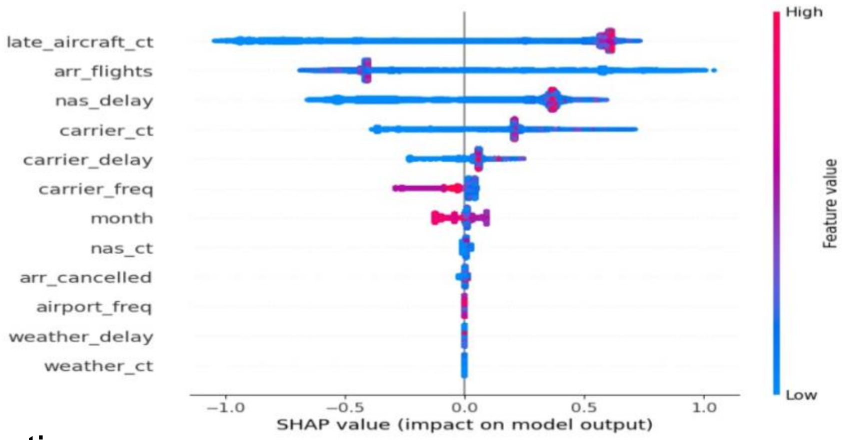
Model Architecture Overview

1. Classification Module (XGBoost Classifier)

Goal: Determine if an upcoming flight is likely to face a delay.

- **F1 Score:** 0.8569
- **F2 Score:** 0.8868 *(puts more weight on recall for delay prediction)*
- **ROC-AUC Score:** 0.8959
- **Key Drivers:** late_aircraft_ct, arr_flights, nas_delay, carrier_ct, carrier_delay

Insight: Effectively detects high-risk flights, enabling early alerting and focused mitigation.



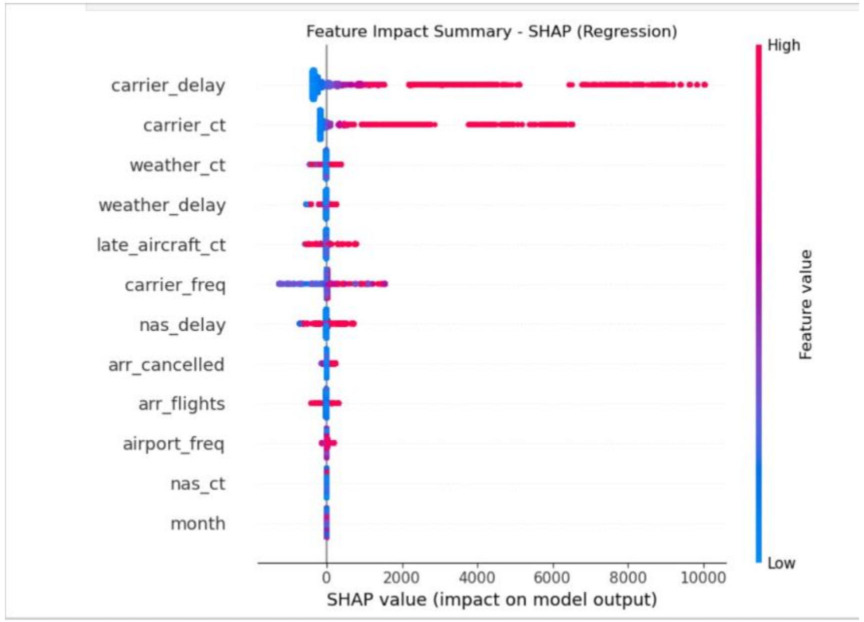
2. Regression Module (XGBoost Regressor)

Goal: Estimate delay magnitude (in minutes) for predicted delayed flights.

- **Mean Absolute Error (MAE):** 247.30 minutes
- **Root Mean Squared Error (RMSE):** 1041.00 minutes
- **R² Score:** 0.7882 *(solid explanatory strength)*
- **Top Predictors:**

carrier_delay, carrier_ct, weather_ct, weather_delay, late_aircraft_ct

Insight: Provides a quantitative view of delay severity to better prioritize airline operations.



3. Unified Prediction Pipeline

Function: predict_flight_delay(input_df, xgb_model_clf, xgb_model_reg)

- Merges classification (will it delay?) and regression (how much?)
- Runs severity estimation only when delay is predicted
- Outputs clear, actionable values for each flight

Insight: An end-to-end intelligent pipeline for delay risk forecasting and impact assessment—deployment-ready.

Strategic Takeaways & Operational Suggestions

Key Observations:

- **Carriers with high OAI potential** (e.g., WN, AA, DL) are mostly affected by manageable delays — making them ideal candidates for operational fine-tuning.
- **Leading contributors to delays** are typically associated with **Late Aircraft** and **Carrier-based delays**.
- **Carriers with low OAI** (like 9K, EM, PT) often experience delays driven by external variables, leaving minimal room for direct intervention.

Suggested Actions for Delay Mitigation:

Flight Schedule Rebalancing:
Reorganize timings on routes with heavy congestion or frequent delays to ease pressure on operational systems.

Ground Handling Efficiency:
Streamline refueling, baggage movement, and boarding through tighter coordination to speed up aircraft turnaround times.

Passenger-Focused Messaging:
Use real-time delay updates and alternative travel options to manage expectations and reduce inconvenience.

Targeted Resource Deployment:
Strategically assign more manpower and support at airports that historically experience frequent delays, especially during seasonal peaks or bad weather.

Cross-Entity Collaboration:
Encourage stronger alignment between airlines and airport authorities to resolve persistent infrastructure and logistics challenges.