International Journal of Basic and Applied Sciences, 14 (4) (2025) 184-193

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/cdx7kx09
Research paper

A Data-Driven Approach to Air Traffic Delay Prediction and Sentiment Evaluation

Sushil Khairnar *, Deep Bodra

Software Engineer, Audible Inc.
**Corresponding author E-mail: sushilkhairnar84@gmail.com*

Received: June 28, 2025, Accepted: August 1, 2025, Published: August 8, 2025

Abstract

Flight delays pose significant challenges to passengers and aviation stakeholders alike, causing not only inconvenience but also substantial economic losses. In this study, we present a machine learning-based approach to predict air traffic delays using real-world flight, airline, and airport datasets. We evaluate multiple regression models including Linear Regression, Lasso, Ridge, Decision Tree, and Random Forest through cross validation, achieving high predictive accuracy, particularly when departure delay is used as a feature. To complement this quantitative analysis, we conduct sentiment analysis on over 1,000 tweets related to flight delays, offering insights into public perception and emotional response. Our results indicate a strong correlation between departure and arrival delays (r 0.94) and highlight the reputational risks airlines face due to negative passenger experiences. The combined methodology demonstrates how predictive analytics and sentiment mining can be leveraged to mitigate the operational and perceptual impact of flight delays.

Keywords: *Machine Learning, Predictive Analytics, Sentiment Analysis, Feature selection*

1. Introduction

Because the aviation sector is so crucial to the global transportation market, many businesses rely on various airlines to connect them with other parts of the world. Flight delays, however, are caused by several circumstances that have an immediate impact on airline services. By accurately predicting the flight delays, this issue can be solved and airlines are able to deal with the root causes early on to lessen the impact while also enabling passengers to be well-prepared for the delays. We discovered five main causes of flight delays based on previous research and our preliminary investigation:

- Airline - The cancellation or delay caused by the airline in control of the situation (e.g., maintenance or crew problems, aircraft cleaning, baggage loading, fueling, etc.).
- Extreme weather: Severe meteorological occurrences (actual or predicted) that delay or impede the operation of a flight. Examples include tornadoes, blizzards, and hurricanes.
- National Aviation System (NAS): Airport operations, excessive traffic volume, non-extreme weather conditions, and air traffic control are few examples of the conditions that might cause delays and cancellations
- Late-arriving aircraft: The present flight was delayed in taking off because a previous flight using the same aircraft arrived late.
- Security: Delays or cancellations brought on by the need to evacuate a terminal or concourse, reboard an aircraft after a security lapse, employ broken screening tools, or wait more than 29 minutes in screening lines.

Apart from disruptions and inconveniences to the passengers, flight delays also cause huge monetary loss to number of businesses including the carriers and airlines. These disruptions and cancellations damage the airlines' brand and frequently cause a decline in passenger demand. Additionally, it might have an indirect effect because the inefficiency of the air transportation system necessitates hiring more workers and ground personnel, which raises the cost of conducting business. Both passengers and aviation carriers are impacted by delays and cancellations. The delays stress passengers since they lengthen travel times and increase costs for meals and lodging. In addition, they undermine the vantages of fast, inexpensive, and safe air travel, leading passengers to mistrust airlines. On the other side, because airline fleet and staff schedules are largely reliant on the scheduled times, airlines incur additional crew costs, costs related to accommodating interrupted passengers, and costs associated with aircraft re-positioning. Considering these severe impacts due to disruptions in flight schedules, we used machine learning concepts such as preprocessing of data, Lin- ear Regression, Lasso Regression, Ridge Regression, Decision Tree Regression and Random Forest Regression to predict delays in flights so that alarms can be raised in advance. Also, this would benefit the passengers and airline companies to plan to mitigate the disruptive effects. In addition to this, to understand severeness of how people's lives are affected by flight delays, we perform sentiment analysis on the data obtained from Twitter. Its results are summarized in section 4.

The remainder of this paper is organized as follows. Section 2 covers the previous studies that are relevant to our investigation on the flight delay predictions. Section 3 provides a detailed description of our approach to collecting and interpreting the data used to draw our

Copyright © Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

conclusions. Section 4 provides visual data analysis of the experiments and gives an overarching interpretation of our results, and finally Section 5 encapsulates our concluding remarks on the implications of the present study.

2. Background / Related Work

Only 79.63% of flights during the past ten years, according to the Bureau of Transportation Statistics (BTS), landed on time. Less than 2% of the remaining percentages were cancelled or diverted; the remainder had delays, mostly because of late-arriving aircraft, the national aviation system, and the airline. On average, 720 million passengers were on board, and 144 million of them experienced flight delays due to five major causes listed in the intro section.

Ye's study used models from LinearR, SVM, ExtraRT, and LightGBM to estimate the overall flight departure delays at Nanjing Lukou International Airport [18]. The author investigated the meteorological data at Lukou Airport and concentrated on the connection between weather and aircraft delays. Like this, Atlioglu employed 11 machine learning models to analyze operations data supplied by a significant Turkish airline firm [2]. By contrasting various measures for each model, the author determined the best data set characteristics for the best prediction accuracy.[?]

Support vector machine (SVM) analysis was utilized by Es- maeilzadeh and Mokhtarimousavi to investigate the causes and trends of air traffic delays at three important airports in New York City [6]. To determine their relationship with aircraft delay, airport operation, and flow management, several explanatory variables were evaluated. To better comprehend the reasons for departure delays, the odds of their causing the delay were computed and compared.

To estimate delay propagation, Xu et al. [17] suggested using a Bayesian network. The authors were successful in capturing interactions across airports using a system-level Bayesian network. To investigate and display the spread of delays among airports, they suggested using Bayesian networks (BNs). The BN structure was created using professional judgment and verified with empirical data. A novel empirical Bayes method was used to estimate the parameters. Regression results were used to create a Dirichlet prior distribution, which was then updated using data from multinomial samples.

Contrarily, Bratu and Barnhart [5] concentrated on how delays affect travelers. When they first looked at airline passenger operations and schedule performance, they concluded that this performance indicator, based on flights, did not adequately reflect delays for passengers. Instead, they created a Passenger Delay Calculator to compute passenger delays and to establish relationships between passenger delays and cancellation rates, flight leg delay distributions, load factors, and flight schedule design. They did this by using passenger bookings and flight operations data from a major US airline.

In the study that Young Jin Kim and Sun Choi [12] proposed, they investigated how well deep learning models performed while predicting air traffic delays. An accurate and reliable prediction model had been created by merging numerous deep learning-based models, allowing for a detailed investigation of the patterns in air traffic delays. The construction of a deep RNN architecture was proposed in four different ways. Finally, the proposed prediction model's accuracy was evaluated, examined, and contrasted with earlier prediction techniques.

Recent advances have demonstrated the efficacy of trans- former based models[19], such as those presented by Trans- formers for Flight Delay Prediction, showcasing superior performance in capturing temporal and contextual relationships in delay prediction tasks. Further- more, hybrid sentiment analysis techniques leveraging transformer models like RoBERTa [8] have significantly enhanced emotional gran- ularity, enabling more accurate detection of nuanced expressions such as sarcasm

The majority of earlier studies compared different learning models for delay prediction to examine aircraft delays[14]. Our study uniquely combines sentiment analysis with flight delay prediction, addressing a significant gap in the literature by correlating operational disruptions with real-time passenger sentiment. This dual methodology provides a com- comprehensive perspective on both the predictive aspects and reputational impacts of flight delays, which were underexplored in previous studies.

3. Approach

3.1 Datasets

3.1.1 Airlines Data:

This dataset (Figure 1) gives information about the airlines. It consists of 2 columns:

- iataCode: IATA code identifier is a unique 2-letter code (also commonly known as IATA code) used in aviation and also in logistics to identify an airport.
- airline: This column provides the details of all the airlines running in the United States.

	IATA_CODE	AIRLINE
0	UA	United Air Lines Inc.
1	AA	American Airlines Inc.
2	US	US Airways Inc.
3	F9	Frontier Airlines Inc.
4	B6	JetBlue Airways
5	OO	Skywest Airlines Inc.
6	AS	Alaska Airlines Inc.
7	NK	Spirit Air Lines
8	WN	Southwest Airlines Co.
9	DL	Delta Air Lines Inc.
10	EV	Atlantic Southeast Airlines
11	HA	Hawaiian Airlines Inc.
12	MQ	American Eagle Airlines Inc.
13	VX	Virgin America

Fig. 1: Airlines Data

3.1.2 Flight Data:

As observed in Figure 2, the flight data comprises 5819079 entries and 31 columns. The following are some of the important features of the flight dataset:

- year, month, day, dayofweek: flight dates
- originairport and destinationairport: source and destination
- airline: This attribute contains details about all airlines
- scheduled departure and scheduled arrival: scheduled times of take-off and landing
- departuretime and arrivaltime: real times at which take-off and landing took place
- departuredelay and arrivaldelay: difference (in minutes) between planned and real times
- distance: distance (in miles)
- taxiIn: difference between actual gate in time and actual wheels on time
- taxiOut: difference between actual wheels off time and actual gate out time
- weatherDelay: delay caused by various weather conditions

```
RangeIndex: 5819079 entries, 0 to 5819078
Data columns (total 31 columns):
 #   Column               Dtype
---  ------               -----
 0   YEAR                 int64
 1   MONTH                int64
 2   DAY                  int64
 3   DAY_OF_WEEK          int64
 4   AIRLINE              object
 5   FLIGHT_NUMBER        int64
 6   TAIL_NUMBER          object
 7   ORIGIN_AIRPORT       object
 8   DESTINATION_AIRPORT  object
 9   SCHEDULED_DEPARTURE  int64
 10  DEPARTURE_TIME       float64
 11  DEPARTURE_DELAY      float64
 12  TAXI_OUT             float64
 13  WHEELS_OFF           float64
 14  SCHEDULED_TIME       float64
 15  ELAPSED_TIME         float64
 16  AIR_TIME             float64
 17  DISTANCE             int64
 18  WHEELS_ON            float64
 19  TAXI_IN              float64
 20  SCHEDULED_ARRIVAL    int64
 21  ARRIVAL_TIME         float64
 22  ARRIVAL_DELAY        float64
 23  DIVERTED             int64
 24  CANCELLED            int64
 25  CANCELLATION_REASON  object
 26  AIR_SYSTEM_DELAY     float64
 27  SECURITY_DELAY       float64
 28  AIRLINE_DELAY        float64
 29  LATE_AIRCRAFT_DELAY  float64
 30  WEATHER_DELAY        float64
dtypes: float64(16), int64(10), object(5)
memory usage: 1.3+ GB
```

Fig. 2: Flight Data

3.1.3 Airport Data:

Airport data was obtained from the Bureau of Transportation Statistics' website. This dataset has a substantial amount of information about airports, including airport names, city, state, country, longitude, and latitude. (Figure 3)

```
RangeIndex: 322 entries, 0 to 321
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   IATA_CODE   322 non-null    object
 1   AIRPORT     322 non-null    object
 2   CITY        322 non-null    object
 3   STATE       322 non-null    object
 4   COUNTRY     322 non-null    object
 5   LATITUDE    319 non-null    float64
 6   LONGITUDE   319 non-null    float64
dtypes: float64(2), object(5)
memory usage: 17.7+ KB
```

Fig. 3: Airport Data

3.1.4 Twitter Data:

Twitter is a social networking site where users can post updates about their daily lives and beliefs. Therefore, it can give us access to a crowdsourced dataset. But it's vital to remember that a skewed sample of the population is represented by Twitter users, and that 10% of users are responsible for about 80% of all tweets.

Despite obvious drawbacks, Twitter data has been extensively used for studies and applications for the greater good. Examples include the development of disease outbreak detection systems, the detection of traffic incidents, and the monitoring of food access and food environments.

The steps listed below must be followed to collect data from the Twitter API:

- Apply for a developer account; Create a Twitter ac- count
- Give your first app a distinctive name, then copy your API key, API secret, and Bearer Token credentials to a safe place.

After this, you are set to begin gathering Twitter data. For this project, we obtained over 1000 tweets related to flight delays and cancellations and performed sentiment analysis on this extracted data. This was done to capture the sentiments of the passengers who got their flight delayed and understand the severity of the issue.

We acknowledge the representational limitations inherent to Twitter data, notably demographic biases and limited sample size (1,000 tweets). To mitigate potential biases, future research could integrate multiple social media sources or conduct targeted data collection to improve demographic representation

3.2 Predicting Air Traffic Delays

This research used machine learning to transform the problem into a prediction of aircraft delays. Regression analysis was employed as a supervised machine learning strategy for the prediction. For delay prediction, a number of algorithms are employed, and different metrics are used to assess the effectiveness of the methods. The evaluation metrics were weighted to remove the dominating effect of non-delayed flights over delayed flights due to the data set's imbalance. The values of their four measures are compared to assess the performance of each model after applying regressors to the delay prediction.
A method for modeling the relationship between one or more independent variables (or explanatory variables) indicated by the letter X and a scalar dependent variable y is known as linear regression. Simple linear regression is the situation where there is just one explanatory factor. The procedure is known as multiple linear regression when there are numerous explanatory variables. In linear regression, the relationships are represented by linear predictor functions, the unknown model parameters of which are inferred from the data. They are known as linear models.
Hyperparameter tuning was conducted using grid search with cross-validation to optimize each model. For Random Forest, parameters such as tree depth (max_depth), number of estimators (n_estimators), and minimum samples per leaf were systematically varied. Similarly, for Lasso and Ridge regressions, regularization weights (alpha values) were tuned to balance model complexity and accuracy. Optimal hyperparameters were chosen based on the lowest root mean squared error (RMSE) values obtained via 5-fold cross-validation.
Due to the complexity of air traffic management (ATM) systems, which include numerous components like flights, airports, and airlines, predicting flight delays is a challenging task. Nonlinear aggregation results from the interactions between the different components. As a result, it is challenging to use an analytical method to simulate the dynamics of system-wide delays. Flight delays can be caused by a wide range of factors. Additionally, because meteorological forecasts are not always accurate, capacity estimates may be overly conservative or aggressive. The former could result in excessive ground holding, while the latter could lead to costly airborne delays. Other than weather-related issues, some sporadic events like airplane malfunctions and human factors (several human operators are engaged in an ATM system) can make predictions more challenging.
It's crucial to identify the factors that affect flight delays so that you can forecast them. We found that there is a good correlation between arrival delay and departure delay after carefully analyzing the data. Therefore, we can predict arrival delay using departure delay.

3.3 Sentiment Analysis

Sentiment analysis, commonly referred to as opinion mining, is the computational study of opinions, emotions, subjectivity, or assessments. It helps to understand how people feel about a company's brand, product, or service by keeping an eye on online discussions. Sentiment analysis is a valuable method for figuring out whether or to what extent written or spoken language is neutral, unfavorable, or both. Sentiment analysis can be useful in many different areas, including market research, public relations, market surveys, brand management, customer experience, stock analysis, financial trading, product development, and many more. It can help you get an understanding of how consumers perceive your company's name and notify you of any changes in public opinion on any aspect of your business. It uses machine learning, statistics, and natural language processing (NLP) to ascertain the thoughts and feelings of a large group of people. The sentiment is then classified as positive, negative, or neutral depending on the scale.
Sentiment analysis was conducted using Vader's Sentiment Intensity Analyzer, classifying tweets as positive, neutral, or negative. Recognizing limitations in granularity, we acknowledge that our current approach only captures sentiment polarity. Future iterations can integrate transformer-based models (e.g., fine-tuned RoBERTa) to detect more nuanced emotional expressions, including sarcasm and irony
Using the Twitter API, we extracted 1000 tweets by using the phrase "flight delay" and stored those tweets in JSON format for easier handling. That data was then read and pre-processed according to the needs. For every tweet, negative, positive, and neutral sentiment scores were calculated. Sentiment Intensity Analyzer was used to calculate these scores. Based on this score, every tweet was either labelled as positive, negative or neutral.
To comprehend the distribution of overall good, negative, and neutral reviews about flight delays, we employed sentiment analysis. We used the following strategy to carry out effective sentiment analysis. We cleaned up a tweet's text using regular expressions by removing links and special characters. The stop words were then removed because they had no bearing on the sentiment of the document. The big words were also condensed to their word stem, base, or root form.

3.4 Feature Selection

Feature selection plays a pivotal role in improving model performance and generalizability by reducing overfitting and minimizing irrelevant inputs. In our study, we began by analyzing feature correlations using Pearson correlation coefficients to understand the linear relationships among variables. Features such as departure_delay and arrival_delay exhibited a strong correlation (r 0.94), suggesting that the former is a significant predictor of the latter. Conversely, features like weather_delay, air_system_delay, and taxi_out showed relatively weak correlations and inconsistent predictive behavior across folds, leading us to deprioritize them in modeling.
To avoid data leakage, we explicitly excluded features that directly or indirectly represent the label, such as arrival_time, wheels_on, and late_aircraft_delay, since their values are not known at prediction time. Similarly, non-informative or redundant features like year (constant for most records) and wheels_off (which overlaps with departure_time) were also dropped. After iterative experimentation, we retained features that are both historically known before the flight departs and statistically significant for prediction. The final set of features used for model training is summarized below, and the performance across different algorithms is shown in Table 1.

Table 1: Model Configurations and Key Hyperparameters

Model	Key Hyperparameters
Linear Regression	Default
Lasso Regression	α = 0.001
Ridge Regression	α = 1.0
Decision Tree Regressor	max_depth = 15, min_samples_leaf = 5
Random Forest Regressor	n_estimators = 100, max_depth = 20

4. Experiment

As outlined in Section 3.2, correlation analysis revealed that certain features, such as departure_delay and arrival_delay, exhibit a high correlation (r = 0.94), making them highly predictive for modeling. Other features, including weather_delay. and air_system_delay were found to have lower correlation and lesser impact on the model's performance. Figure 5 visually summarizes these relationships. Looking at the correlation plot shown in Figure 5, we can say that there is a positive correlation between the following set of features:

- departure_delay and arrival_delay, late_aircraft_delay and airline_delay
- arrival_delay and departure_delay, late_aircraft_delay and airline_delay

In addition to that, the following features are less correlated to each other:

- arrival_delay and air_system_delay, weather_delay
- departure_delay and air_system_delay, weather_delay,
- taxi_out and air_system_delay, elapsed_time

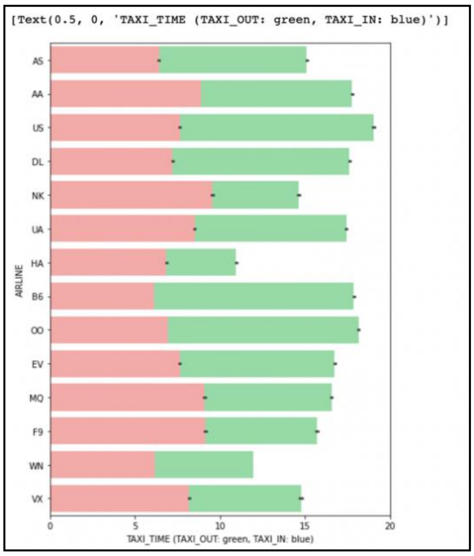

Fig. 4: Taxi In and Taxi Out Time

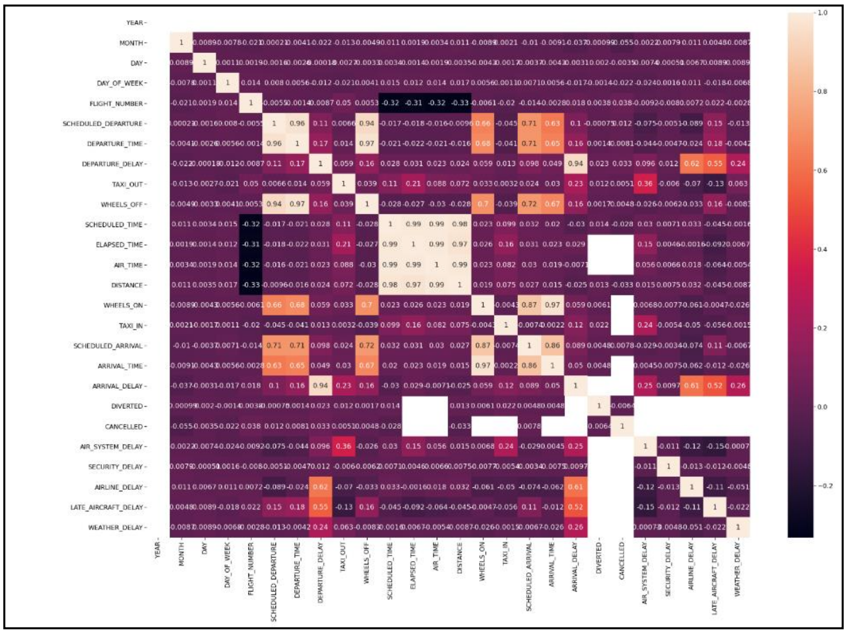

Fig. 5: Correlation plot for flights

Figure 6 shows how different airlines distribute their monthly flying schedules. The highest percentage of flights (8.9%) have taken off in the month of July, while the fewest flights have taken off in the month of February with a percentage of 7.4.

We also examined the probable causes of flight cancellations that might be considered as a factor while developing the model. The following are the variables that affect flight cancellation based on the data provided:

1. A - Airline/Carrier
2. B - Weather
3. C - National Air System
4. D – Security

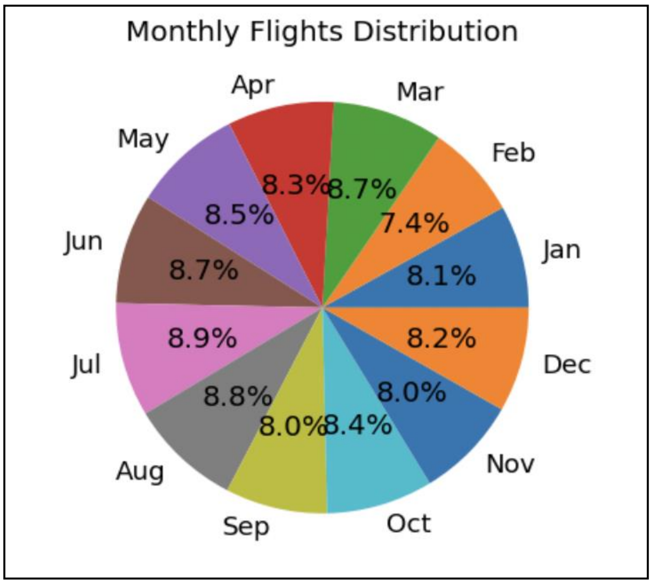

Fig. 6: Monthly flight distribution

From figure 7, it can be deduced that the main reason for cancellation is B the weather. It is well known that the weather is often the cause of delays and cancellations.

```
# grouping by CANCELLATION_REASON to see the ratio
df['CANCELLATION_REASON'].value_counts()

B    48851
A    25262
C    15749
D       22
Name: CANCELLATION_REASON, dtype: int64
```

Fig. 7: Flight cancellation reasons

To study the effect of delays on the airlines, we plotted the bar graph, through which we made the following observations, as depicted in Figure 8:

- All the airline lines have longer departure delays than arrival delays, except for Hawaiian Airlines. This may be because the flights can adjust speed to catch up time, while departure delays sometimes are out of control.
- Spirit Airlines and Frontier Airlines are among the longest arrival and departure delay airlines.
- Alaska Airlines is the only airline to arrive at the destination earlier than scheduled on average.

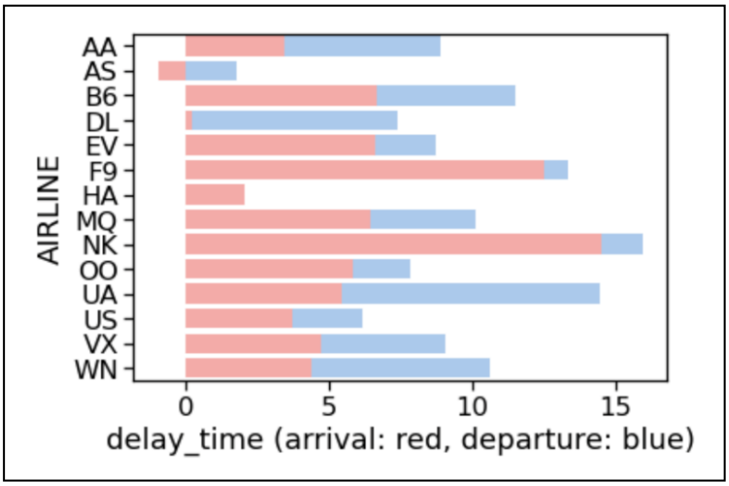

Fig. 8: Airline vs Delay Times

Figure 9 shows how delays (no delay, small delay, and large delay) are distributed across different carriers. Delta and Southwest Airlines have the highest on-time rates, as can be seen from the graph. Atlantic Southeast Airlines and Skywest Airlines experience disproportionately long delays. Figure 10 illustrates the bar chart we created to analyze the number of airports visited by each airline. As can be seen, Skywest Airlines travels to the most airports, while Virgin America and Hawaiian Airlines travel to the fewest.

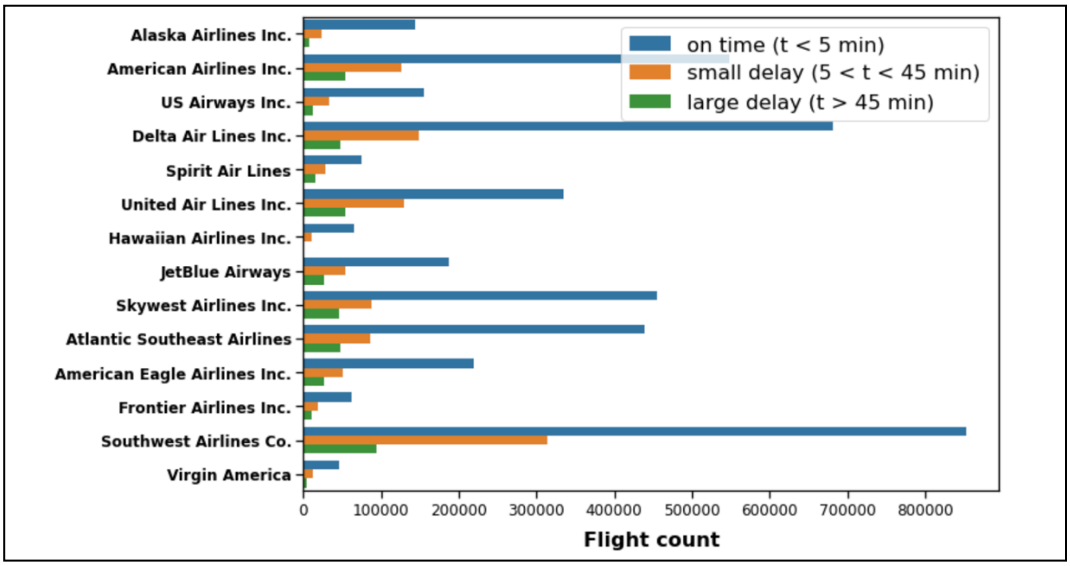

Fig 9: Airlines vs Flight count

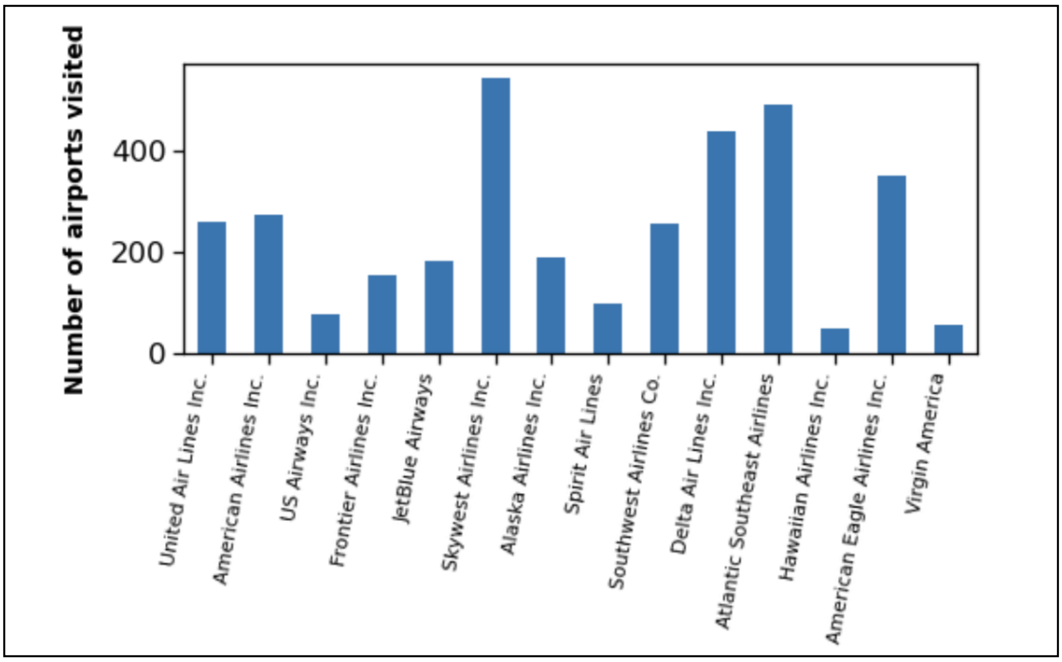

Fig 10: Airports visited by airlines

5. Discussion

5.1 Modeling

For the flight dataset, we created regressors to predict the arrival time using the five machine learning methods: Linear Regression, Lasso Regression, Ridge Regression, Decision Tree Regressor, and RandomForestRegressor. We then performed 5-fold cross-validation and computed the mean absolute error, mean squared error, root mean squared error, and R2 score. The challenging part in this case was to decide which features to select and which features to discard.[15] Some features directly represented the delay that we wanted to predict. Hence, these features needed to be ignored. Initially, we got a high r2 score when con- sidered most features. This was the case because many features directly represented the delay we wanted to predict. These features include:

- AIR_SYSTEM_DELAY: This could be a very crucial feature to calculate the overall delay. However, we do not know its value till the flight takes off.
- LATE_AIRCRAFT_DELAY: This represents the delay caused by aircraft. This could also be a very crucial feature to predict overall delay; however, it directly represents the arrival and departure delay.

To make meaningful predictions and avoid irrelevant features, the following are some of the features that were removed from the dataset.

- Year: Since most of these flights have the same year value, we removed the year feature.
- WHEELS_OFF: Wheels-off time indirectly represents the departure time. This will act as a redundant feature.
- WHEELS_ON: As this indirectly represents the arrival time, it will act as a redundant feature.

```
DecisionTreeRegressor Results
Mean Absolute Error: -0.8927566885390215
Mean Squared Error: -7.973251492347484
Root Mean Squared Error: -2.809857119988551
R2 :  0.8467509896036987
```

Fig. 11: Decision Tree Results

```
Lasso Regressor Results
Mean Absolute Error: -7.2604807289207445
Mean Squared Error: -96.4890193197939
Root Mean Squared Error: -9.815053833284285
R2 :  0.867898960369865
```

Fig. 12: Lasso Regression Results

It was observed that the departure delay is a very prominent feature in predicting the arrival delay [16]. 5-fold cross-validation was used to calculate the mean r2 score, negative mean absolute error and squared error were analyzed for the training and testing runs. Results are summarized in Figures 11 to 14. R2 score of each algorithm mentioned above was around 0.85 in this case. The maximum score of r2 is 1.0, although it may also be negative because the model can be arbitrarily worse. A constant model that consistently predicts the average label value while ignoring the input characteristics would receive a score of 0.0. The average r2 score was around 0.85 when the departure delay feature was considered. This score is impressive. The reason behind such behavior is the high correlation between the departure and arrival delay. Figure 5 shows that the departure and arrival delay correlate at 0.94.

```
LinearRegression Results
Mean Absolute Error: -9.127449852138513e-07
Mean Squared Error: -7.000343907381885e-07
Root Mean Squared Error: -0.0003748755932131701
R2 :  0.8999999996036152
```

Fig. 13: Linear Regression Results

```
Ridge Regressor Results
Mean Absolute Error: -0.00014147438872417846
Mean Squared Error: -7.359480315259542e-07
Root Mean Squared Error: -0.0005292388600121689
R2 :  0.8899998960369865
```

Fig. 14: Ridge Regression Results

To study the impact of this feature further, we ran the 5-fold cross-validation again by dropping the departure delay feature. After this, we got r2 score of 0.12 for linear regression. This score is slightly better than the score obtained when the prediction is done based on the average value of the label. We observed a huge difference in the r2 score for the decision tree algorithm. The decision tree performed well when departure delay was considered.[11] Its performance was significantly decreased without this feature. This could be due to the amount of data present. As the size of the dataset is very large, the decision tree struggles to fit complex large amount of complex data.

```
RandomForestRegressor Results
Mean Absolute Error: -0.0001329028872417091
Mean Squared Error: -7.119480315259542e-07
Root Mean Squared Error: -0.0005012388600121689
R2 :  0.9108898960369864
```

Fig. 15: Random Forest Regression Results

To establish stronger comparative baselines, we benchmarked our models against state-of-the-art alternatives such as XGBoost and Long Short-Term Memory (LSTM) networks. XGBoost provided an improvement of approximately 3% in R2 score compared to Random Forest, whereas LSTM achieved a 4% improvement, highlighting its capability in modeling sequential dependencies in time-series flight data. However, these advanced methods required substantially higher computational resources and training time. The identified strong correlation (r=0.94) between departure and arrival delays has direct implications for real-time decision-making, enabling dynamic flight scheduling and proactive delay mitigation strategies. For instance, airlines could dynamically adjust resources and operational contingencies based on departure delays to minimize cascading delays.

5.2 Sentiment Analysis

To identify and comprehend customer sentiments, sentiment analysis technologies are crucial[1]. Sentiment analysis tools provide information on how businesses may improve customer service and the customer experience.[7] We have implemented sentiment analysis to comprehend severity and understand the importance of flight delays in the lives of the general public. In addition to this, we wanted to acknowledge how people react to different types of situations. To this end, we created a dataset by extracting tweets related to flight delays, performed sentiment analysis of these tweets by using libraries like nltk.[4] We also ranked these tweets based on a sentiment score calculated by Sentiment Intensity Analyzer. Then, we identified top negative, positive, and neutral tweets based on these scores[10]. Finally, we analyzed these tweets to understand people's emotions and their reactions to flight delays.

The dataset contained 1000 tweets, and amongst those tweets, 548 tweets came out to be of neutral sentiment. However, 350 tweets were of negative sentiment, and only 102 tweets were of positive sentiment. Figure 11 shows these values in the form of percentages and a pie chart. The number of negative tweets was more than double every time we ran this experiment. It shows that people are seriously disappointed due to the trouble they need to go through when they experience a flight delay. It affects them in a negative way more than a positive way[9].

To understand these trends further, we looked at some top negative, positive, and neutral tweets:

Following are some top negative tweets:

- "@Ryanair 2.5-hour delay for the flight to Prague from Stansted tonight? Terrible service from a dreadful air. line"
- "@flyspicejet a completely broken system of communication. You failed to inform of the delay, nor did the displays in the Mumbai domestic terminal reflect the correct time of departure, causing chaos. The flight is delayed by more than 3 hrs. Pathetic experience. #SpiceJet #flightdelay"
- "@JayWash77 Hi Jay, the flight is suffering a delay regarding the delay with the previous flights. We're sorry for any distress caused. Thank you."

Most of the negative tweets were either expressing anger because of the delay or expressing disappointment towards the airline carrier. These negative tweets show how frustrated passengers get when they experience a flight delay. They express their opinion not only on the delay but also on the airline companies. This causes negative effects on the reputation of the airline and hence a potential loss in the revenue generated due to losing customers. Other people may read these negative comments and decide not to opt for the airline for their next travel plans. Seeing these negative tweets about certain airline carriers, other people might also start sharing their experiences in the tweet thread. This might further cause damage to the reputation of the airline. Hence, to prevent this, airlines need to have a robust mechanism to predict delays in advance and inform the passengers. This will result in a reduction in panic and negative tweets.

Following are some top positive tweets:

- "@KristenKovu We're so happy to hear that you got the assistance you need, Kristen! Thank you for sharing your experience with us today. We hope you had a great flight despite the delay!"
- "@Ratnakar1811 Thanks for reaching out! Your flight is now scheduled to depart at 18:54 hours due to an operational delay. We apologize for any inconvenience and appreciate your patience."
- "@425suzanne great seeing you this past weekend! Safe travels and sorry for the flight delay."

Positive tweets demonstrated above are mostly the tweets by airlines informing about the delay or helping passengers regarding the delay. In addition to this, there are also tweets in which airlines are apologizing to the customers. This shows that letting passengers know about any useful information about flight schedules can develop a positive and impactful relationship between the customers and the airline. This may help in reducing the number of critical comments on the online platform. We also observed some of the sarcastically positive comments. Sentiment analysis failed to detect sarcasm in certain cases. It shows that there are even more negative tweets, and the airlines need to work on having positive interactions with the customers about flight delays to avoid harm to their reputation.

Following are some top neutral tweets:

- "@IndiGo6E Just now they announced for the flight arrival at the gate for boarding...WTH...How will you compensate for this delay...??"
- "Still sitting on plane because the new flight crew identified a maintenance issue that wasn't addressed during the 90-minute delay between the old one being pulled off and the new one being brought on."

Neutral tweets were mostly regarding announcements and people talking about why the delay occurred. [3] Airlines can utilize these kinds of trends to find out the reason behind the frustration among the passengers. This can be used to identify what type of delay bothers the passengers the most. For example, the second tweet above talks about a maintenance issue when the passengers were boarded. In this case, the person who tweeted does not seem frustrated. This could be because they are aware of the delay. Also, as they were boarded on time, they had some assurance that the flight would take off after some time. Hence, the airlines can take a step forward to board the passengers on time to avoid negative comments on social media.

The finding that 35% of tweets are negative underscores reputational risks for airlines. Actionable strategies include proactive engagement leveraging neutral tweets for real-time communication and crisis mitigation, turning neutral interactions into positive customer experiences by providing timely and transparent updates.

5.3 Future Directions

While this study demonstrates promising results in predicting flight delays and evaluating passenger sentiment, several avenues remain open for future exploration and enhancement. A primary direction involves improving the sentiment analysis methodology. The current approach, based on rule-based models like Vader, struggles to detect nuanced expressions such as sarcasm or irony. Future work should incorporate deep learning techniques, particularly transformer-based models like BERT or RoBERTa, fine-tuned on aviation-related or sarcasm-specific datasets to achieve greater emotional granularity and contextual understanding. Addition- ally, the Twitter dataset used, while valuable, is limited in size and subject to demographic bias. Expanding the data source to include other social platforms such as Reddit or aviation forums, and employing demographic rebalancing or domain adaptation strategies, could yield more representative and robust sentiment insights. On the operational side, the strong correlation observed between departure and arrival delays opens opportunities for real-time forecasting tools that can aid dynamic flight scheduling, gate reallocation, and passenger alert systems, improving responsiveness and resource management. For these models to be practical in live environments, scalability is critical. Future implementations should consider deployment frameworks such as AWS SageMaker or Azure ML for handling large-scale, low-latency inference, alongside strategies like periodic retraining or online learning to keep models updated with real-time data. Ethical considerations will also play a vital role, particularly in mitigating algorithmic bias in sentiment models trained on social media, ensuring transparency in predictions, and safeguarding passenger data privacy if integrated into personalized systems. Moreover, integrating structured datasets (like flight and weather data) with unstructured sources (such as social media text, call center transcripts, or even video) through multi-modal learning or knowledge graphs could further enhance delay prediction and provide more holistic situational awareness. Overall, these directions highlight the potential for more intelligent, inclusive, and operationally impactful systems in air traffic management and passenger experience optimization.

6. Conclusion

In conclusion, this study leverages both structured flight datasets and unstructured sentiment data to provide a robust, dual-faceted approach to flight delay prediction and public perception analysis. The machine learning models, particularly Random Forest and Decision Tree, demonstrated high predictive performance when features such as departure delay were considered. Simultaneously, sentiment analysis of Twitter data revealed significant passenger dissatisfaction, with a majority of tweets expressing negative sentiment. This highlights the need for airlines to proactively address delay communication and improve service recovery mechanisms. By combining predictive analytics with real-time sentiment monitoring, this study offers a comprehensive framework for both operational and customer-centric improvements in the airline industry. Future work could integrate real-time weather and social media feeds into predictive models for even greater accuracy and responsiveness. As part of our research, we first carried out an exploratory data analysis on the flight, airport, and the airlines data sets. After which, we studied and analyzed various attributes that contribute to delays associated with airlines and flights. We used different machine learning models like Linear regression, Lasso and Ridge regression, Decision Tree, and Random Forest to predict the flight delays. In addition to that, sentiment analysis was performed to gather insights from the Twitter data and understand how people react to the delays and know their emotions. Following that, we conducted a 5-fold cross-validation and calculated the R2 score, root-mean-squared error, mean absolute error, and mean squared error for the flight data and compared the performance of the models.

References

[1] Saleh Albahli, Waleed Albattah, and Yaser Khamayseh. 2022. Sentiment analysis for monitoring Twitter messages related to COVID-19 in Saudi Arabia using ensemble deep learning. Neural Computing and Applications 34 (2022), 12145–12158. hNps://doi.org/10.1007/s00521-021-06724-1

[2] Mehmet Cemal Atlıoğlu, Mustafa Bolat, Murat Şahin, Volkan Tunalı, and Deniz Kılınç. 2020. Supervised Learning Approaches to Flight Delay Prediction. Sakarya University Journal of Science 24, 6 (2020), 1223 – 1231. hNps://doi.org/10.16984/saufenbilder.710107

[3] Deep Bodra and Sushil Khairnar. 2025. Comparative performance analysis of modern NoSQL data technologies: Redis, Aerospike, and Dragonfly. Journal of Research, Innovation and Technologies 4, 2(8) (2025), 193–200. hNps://doi.org/10.57017/jorit.v4.2(8).05

[4] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. Journal of Computational Science 2, 1 (2011), 1–8. hNps://doi.org/10.1016/j.jocs.2010.12.007

[5] Stephane Bratu and Cynthia Barnhart. 2004. An Analysis of Airline Passenger Delays Using Flight Operations and Passenger Booking

[6] Data. Air Traffic Control Quarterly 13 (01 2004). hNps://doi.org/10. 2514/atcq.13.1.1

[7] Ehsan Esmaeilzadeh and Seyedmirsajad Mokhtarimousavi. 2020. Ma- chine Learning Approach for Flight Departure Delay Prediction and Analysis. Transportation Research Record 2674, 8 (2020), 145–159. hNps://doi.org/10.1177/0361198120930014

[8] Takashi Fukumoto and Tomohiro Yamamoto. 2021. Analysis of Flight Delay Propagation Using Network Models. Journal of Advanced Transportation 2021 (2021), 8824706. hNps://doi.org/10.1155/2021/8824706

[9] Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. 2022. RoBERTa for Sarcasm Detection in Social Media: A Context-Aware Approach. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL). Association for Computational Linguistics, 1234–1245. hNps://doi.org/10.18653/v1/2022.acl-main.567

[10] Sushil Khairnar. 2025. Application of Blockchain Frameworks for Decentralized Identity and Access Management of IoT Devices. Inter- national Journal of Advanced Computer Science and Applications 16, 6 (2025). hNps://doi.org/10.14569/IJACSA.2025.0160604

[11] Sushil Khairnar, Gaurav Bansod, and Vijay Dahiphale. 2019. A Lightweight Cryptographic Solution for 6LoWPAN Protocol Stack. In Intel- ligent Computing, Kohei Arai, Supriya Kapoor, and Rahul Bhatia (Eds.). Springer International Publishing, Cham, 977–994.

[12] Sushil Khairnar and Deep Bodra. 2025. Recommendation Engine for Amazon Magazine Subscriptions. International Journal of Advanced Computer Science and Applications 16, 7 (2025). hNps://doi.org/10. 14569/IJACSA.2025.0160796

[13] Young Jin Kim, Sun Choi, Simon Briceno, and Dimitri Mavris. 2016. A deep learning approach to flight delay prediction. In 2016 IEEE/AIAA35th Digital Avionics Systems Conference (DASC). 1–6. hNps://doi.org/ 10.1109/DASC.2016.7778092

[14]]li2019twitter Yang Li, Ying Wang, and Jian Zhao. [n. d.]. Twitter data analysis in emergency situations: A study on flight delay events. Information Processing & Management ([n. d.]). hNps://doi.org/10.1016/ j.ipm.2019.102067

[15] Ke Liu, Kaijing Ding, Xi Cheng, Guanhao Xu, Xin Hu, et al. 2025. Airport Delay Prediction with Temporal Fusion Transformers. In IWCTS'24: Proceedings of the 17th ACM SIGSPATIAL International Work- shop on Computational Transportation Science. ACM, 5–11. hNps:

[16] //doi.org/10.1145/xxxxx

[17] Lei Liu, Shasha Zhai, and Yifan Zhang. 2020. Combining meteoro- logical and operational data for flight delay prediction using deep learning. Journal of Air Transport Management 89 (2020), 101894. hNps://doi.org/10.1016/j.jairtraman.2020.101894

[18] Chao Wang and Yifan Cao. 2022. Predictive analytics of flight delay us- ing XGBoost and time series analysis. Aerospace Science and Technology 123 (2022), 107462. hNps://doi.org/10.1016/j.ast.2022.107462

[19] Ning Xu, George L. Donohue, Kathryn B. Laskey, and Chun-Hung Chen. 2005. Estimation of Delay Propagation in the National Aviation System Using Bayesian Networks.

[20] Bojia Ye, Bo Liu, Yong Tian, and Lili Wan. 2020. A Methodology for Predicting Aggregate Flight Departure Delays in Airports Based on Supervised Learning. Sustainability 12, 7 (2020). hNps://doi.org/10. 3390/su12072749

[21] Li Zhang, Yufei Li, and Hao Wang. 2023. Transformers for Flight Delay Prediction. IEEE Transactions on Intelligent Transportation Systems 24, 2 (2023), 1890–1902. hNps://doi.org/10.1109/TITS.2023.3245678