Exploring the Impact of Weather on Flight Delays

MSc Research Project

Research Practicum part-2

Vasanth Babu Kassa

Student ID: 23408936

School of Computing

National College of Ireland

Supervisor: Hicham Rifai

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Vasanth Babu Kassa

Student ID: 23408936

Programme: MSC_DAD_B_JAN_25 **Year:** Jan 2025………..

Module: Research Practicum part-2

Supervisor: Hicham Rifai…………………………………………….………

Submission Due Date:

…….……….

Exploring the Impact of Weather on Flight Delays

Project Title:

…….……….

Word Count: ***6072***

Page Count …………………………………………….…….

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……

Date: ……

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Exploring the Impact of Weather on Flight Delays

Vasanth Kassa

23408936

Abstract

The problem of weather interference greatly affects the global aviation operations, resulting in huge financial losses and discontent among passengers. The topic of this research is the influence of meteorological conditions on flight delay and the effectiveness of the machine learning model in predicting flight delays. The study made use of a large dataset of 1,936,758 flight records (including 30 features) by systematically pre-processing the data, such as by strategic imputation of 689,270 missing values and class balancing using SMOTE. The numerous widely used machine learning classifiers, Logistic Regression, Support Vector Machine, Random Forest, and XGBoost, were compared. The measures utilised were accuracy, precision, recall, F1-score, and ROC-AUC rates. Findings show XGBoost to be the best model with an accuracy of 98.45%, ROC-AUC of 0.99%, precision of 90.12%, and F1-score of 83.78%, which is better than SVM (97.26%), Random Forest (95.69%), and Logistic regression (93.70%). Weather delays were statistically found to be as high as 2.39 minutes, with higher outliers that were 1,932 minutes. These results confirm that gradient boosting algorithms are more effective predictive models of weather-related flight delays, and airlines need to rely on them to have a solid decision support model designed to handle proactive delays and improve operational resiliency.

1 Introduction

1.1 Background

The weather is the most disruptive entity in international aviation as it defines more than three-quarters of all flight delays and consumes approximately 33 billion dollars by airlines every year. In North America, the proportion of late arrivals increased by 29% in 2022, indicating a decline in operational efficiency (Buchholz, 2023) ***[Figure 1]***. Weather conditions like rain, wind speed, and low visibility are always mentioned among the major causes of delay (Awoyemi, 2022). As the problem of climate variability increases and makes the climate unpredictable, airlines develop an increased use of predictive analytics and machine learning (Passarella et al., 2024). Current literature does not have a single feature-selection strategy, and the question of the most important meteorological variables affecting flight delays and the best algorithm is unclear.

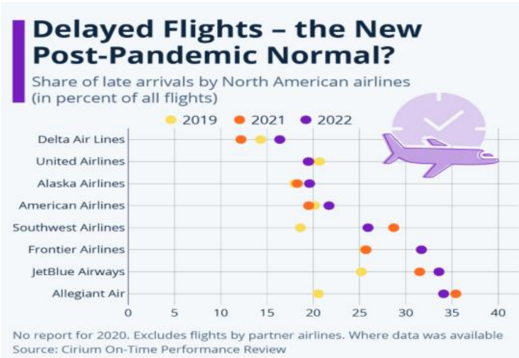


Figure 1: Flight Delays in North America
(Source: Buchholz, 2023)

1.2 Problem Statement

Most of the flight delays are a result of weather, resulting in losses of billions of dollars to the airlines and customer dissatisfaction. The predictive models, the relative significance of meteorological forces, and the relative performance of ML algorithms are not clear for predicting flight delays (Hatıpoğlu and Tosun, 2024). This can easily restrict actionable, data-driven scheduling approaches.

1.3 Research Aim, Objectives, and Questions

1.3.1 Research Aim

The study aims to examine how meteorological factors affect the delays and to determine the efficiency of ML algorithms in forecasting disruptions caused by the weather.

1.3.2 Research Objectives

- To detect the most important weather predictors (precipitation, visibility, and wind speed) with the help of a correlation-based feature selection, and to compare the performance of Logistic Regression, SVM, Random Forest, and XGBoost Clasifiers by their accuracy, precision, recall, F1-score, and ROC-AUC.

1.3.3 Research Question

What factors of weather have the highest impact on flight delays, and how do ML models show the greatest predictive power regarding disruptions caused by weather?

1.4 Research Contribution

The study provides contributions in the form of prioritisation of the top weather predictors of flight delays. Mokhtarimousavi and Mehrabi (2022) mentioned that, in the future, they might examine the incorporation of flight cancellation variables and effects of delay propagation into predictive models of airports to enhance the accuracy and influence on flight delay prediction systems. Thus, this study helps to improve the scheduling, operations resilience, and customer satisfaction by applying machine learning-based weather-delay forecasting through flight cancellation, arrival, and departure delays.

1.5 Research Assumptions

This study is based on the assumption that the Kaggle Airline Delay Causes data is accurate, valid, and can be used to generalise to the larger aviation processes. It assumes that the weather delay records are properly assigned and that they are consistently coded between airlines. The meteorological factors are assumed to be sufficiently represented by the precipitation, the speed of wind, and the visibility as the critical environmental factors affecting delays. It is assumed that the binary Weather Delay Indicator is adequate to capture the event of weather-related disturbances. Another assumption is that machine learning models that are trained on past data can be generalised to predict patterns of delay in the future.

1.6 Research Limitations

This research work has limitations because it uses secondary data in history, applies a binary indicator of delay that is too simplistic to determine the severity of delay, uses a feature selection method based on correlations that ignore nonlinearities, and results that might not be applicable in other regions or changing climatic patterns.

2. Related Work

2.1 Introduction

This chapter provides an overview of the prior literature related to the topic of weather affecting aviation, meteorological factors that influence delays, and the systems of machine learning that are used to predict flight delays. It contrasts international discoveries to define the trends of methodology, technological, and research gaps. This generalisation forms the basis of creating a powerful, evidence-based forecasting model.

2.2 Effects of weather on Aviation activities on a global scale

Weather variability has been widely identified as a superior ability that influences worldwide aviation performance. Kováčiková et al. (2025) assessed the quantitative meteorological effects of extreme precipitation, low visibility, and wind shear incidents on the European air transport network and discovered that these events were significantly contributing to operational delays and cancellations, especially where congestion occurred in the airport. They have also used statistical correlation and sensitivity analysis to relate the climate anomalies to the airport efficiency measures and found that the predictor of the low European hub resilience was the delayed recovery following major weather events.

Similarly, Voskaki, Budd, and Mason (2023) have selected a systematic literature review to study how airports around the world are exposed to weather risks like sea level rise, extreme temperatures, and flooding. A qualitative meta-analysis of 69 global studies found that 74% of airports already had experienced information on climate-related disruption, wherein flooding and storm surges are the most expensive in terms of infrastructure and operational losses. This study also recommends the inclusion of climate adaptation in the planning of airports to achieve resilience in the long term.

On the contrary, not applying any specific model, Song, Wang, and Li (2024) relied on Kothe's Lorenz-Arnold Network (KAN) model and grey relational analysis to estimate the impact of weather on operational resiliency at Xi'an Xianyang International Airport, with temperature (0.35 importance) and wind speed (0.32) as the critical factors influencing the flight stability. In comparison, Kováčikova et al. (2025) and Voskaki, Budd, and Mason (2023) measured the vulnerabilities and the adaptation measures on the macro-level, whereas Song, Wang, and Li (2024) measured the predictive resilience on the micro-level. Overall, this has been proven that weather continues to play a pivotal role when it comes to flight reliability and that predictive modelling can be applied to the flights to improve the comprehension of the role of aviation resilience around the globe.

2.3 Meteorological Variables That Affect Flight Delays

Weather is always cited by being the major contributor to aviation delays in the global arena. Wu, Mei, and Shao (2022) analysed the effects of meteorology on the activities of flights through a hybrid method of statistical and machine learning, combining the data of the rainfall, the visibility, and the wind velocity in Chinese airports. Their study aimed to establish the effect of variations of these parameters on the departure and arrival time punctuality. They identified precipitation intensity and low visibility as the most robust predictors using the Random Forest regression and correlation analysis as the one which explained 64% of the delay variability across regions.

However, Kim and Park (2024) extended it by implementing machine learning and deep learning models, such as Decision Tree, Random Forest, SVM, and LSTM, on the flight data in the 1.57million instances in the airports of ICN, JFK, and MDW. They had the aim of forecasting departure delays on the 2–48-hour horizon based on the multi-temporal weather data. The models reached as much as 85% accuracy, and it was found that precipitation and wind gusts were strong factors in determining delay times, and regional climatic cycle variations (monsoon or winter storms) made the delay pattern differ between continents.

Conversely, Li and Yao (225) conducted a review of research in the world on delay propagation to learn the interaction between meteorological factors and the variables of operation. Their bibliometric review of 641 articles revealed that chain delays in airline networks tend to occur due to weather factors, and extreme factors and atmospheric coupling increase their magnitude of effect. All these studies have shown that meteorological variables that have the greatest impact on the flight delay pattern around the globe include precipitation, visibility, and the speed of wind.

2.4 Use of Machine Learning to Predict Flight Delays

The prediction of flight delays has been revolutionised through machine learning, where the process can be done based on aviation big data. In their deep-learning ensemble model, Alfarhood et al. (2024) combined the Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) into a single model, which allowed encoding spatial as well as temporal patterns on flight data. Additionally, this model, which was trained using the domestic flight data of the U.S., has an average accuracy of 94.6% and can be better adapted to multi-dimensional data than the conventional regression-based models.

Comparatively, Chanchal and Pandey (2023) provided a comparative analysis of the different machine learning algorithms, such as Decision Tree, Random Forest, Gradient Boosting, and Neural Networks, in predicting flight delay. They analysed their choice of algorithms, feature engineering, and evaluation metrics in terms of accuracy, precision, recall, F1-score, and AUC-ROC. According to them, ensembles such as Random Forest and Gradient Boosted Trees always performed better than solitary models, with Random Forest having a maximum success of 92% in various data sets.

However, the model by Dai (2024) was a hybrid model that combines clustering (DBSCAN) and a random forest optimised on the Coyote Optimisation Algorithm (COA). This strategy solved large-scale flight data into little subproblems, which gave a total achievement of 97.2%, which, in comparison to traditional ensemble techniques, was a 4.7% improvement. The model increased the speed of computations (39%) alongside the accuracy of feature selection using ANOVA and Forward Sequential Feature Selection (FSFS). In comparison, Chanchal and Pandey (2023) focused on the diversity of the algorithms, Alfarhood et al. (2024), as well as Dai (2024) showed optimisation and scalability. Taken together, these investigations have revealed that the hybrid and ensemble ML schemes turned out to be the most stable and precise candidates for contemporary flight delay forecasting.

2.5 Aviation Data-Driven Decision Support and Operational Resilience in the case of flight delays

The complexity of the operations of the air transport has influenced the evaluation of data-driven same structure to enhance decision support and resilience. Khairnar and Bodra (2025) concentrated on maximising real-time flight booking within a decision support framework based on data data-centric model combining predictive hybrid analytics and reinforcement learning. Their research was to reduce cascading delays and distorted resource allocation via dynamic delay-feedback loops. Demonstrating a refinement in delay recovery efficiency of 17% be it using regression or decision-tree ensemble measures by RMSE and accuracy, they took a closer look at the advanced performance in turnaround planning and slot allocation resilience made possible by adaptive algorithms.

However, Yuan, Wang, and Lai (2025) came up with a 3DF-DSCL deep learning architecture that enables the prediction of airport delay, based on 3D CNN, GCN, and LSTM networks. Their model used spatio-temporal and dynamic trajectory data of the large-scale AOCC system, which gave a Mean Absolute Error (MAE) of 0.26, which is an improvement of 14.47% of accuracy over baseline hybrid models. This research indicated that the combination of congestion indexes and weather requirements is operationally useful as a predictor of proactive airport control decisions.

On the same note, FlightNet-ST by Zhong et al. (2025) is a spatio-temporal hybrid deep learning framework based on LSTM, GCN, and 3D-CNN with an attention mechanism to predict interpretable delays. Their system realised a 14.47% MAE decrease relative to the traditional models and provided valuable information on the types of congestion, departure times, and flows within the network. Overall, all these studies have proven that data-driven systems have proficiently enhanced the aviation resilience in its operational and decision intelligence in the delay management.

2.6 Theoretical Underpinning

This entire research study is based on the ***Complex Adaptive Systems (CAS) Theory,*** which considers aviation operations as systems where weather fluctuation can have large-scale network effects. CS theory helps in describing systems with interacting mechanisms that self-organise, learn, and adapt for transformations (Apter et al., 1992). CAS contributes to the adaptation of machine learning to predict flight delays in the study because it resembles the responses of adaptivity in the dynamic data-driven environment ***[Figure 2]***. The study seamlessly connects with the principles of self-organisation and nonlinearity in CAS by examining the effects of meteorological factors using smart algorithms and illustrates the increased resilience of aviation during the complicated weather-operation interaction processes.

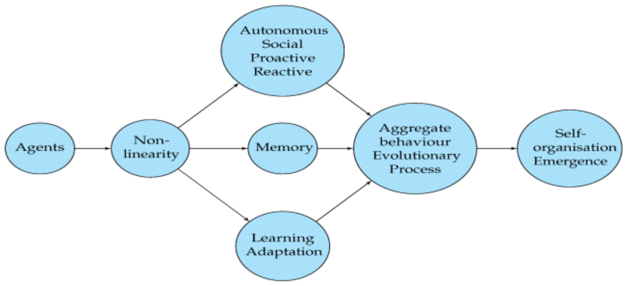


Figure 2: CAS Theoretical Framework

(Ahmad, Baryannis and Hill, 2024)

2.7 Literature Gaps

Even though current research has investigated the impacts of weather on worldwide aviation, it is still limited due to the scale and method. The earlier studies done were more in the form of regional or airport-based studies, and hardly have any combined models that relate global meteorological trends to operational resilience. Moreover, there is a scarcity of literature that quantitatively measures the combined impact of several weather variables with the help of machine learning. Therefore, the multi-variable weather prediction and cross-airport modelling are not yet thoroughly explored techniques in the research of aviation delays.

2.8 Conceptual Framework

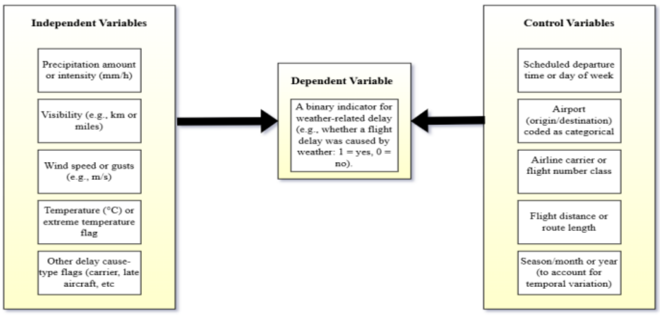


Figure 3: Conceptual Framework

(Source: Self-created)

2.9 Justification of Selecting ML models

Models	***Importance***	***Justifications***
SVM	Support Vector Machine is highly appropriate when the classification problem and the data are of a high dimension, as well as when the decision boundary is non-linear (Sahu et al.,	This renders it suitable for differentiating between the weather-delay and no weather-delay flights using various predictors of weather and

	2023).	operations.
Random Forest Classifier	Random Forest is resistant to overfitting, supports feature sets of considerable scale, measures feature importance, and has been demonstrated to be a valuable predictor of flight delays (accuracy of 80.5% in weather-delay prediction) (Wang et al., 2022).	It suits the purpose of this study to rank the important predictors of weather and be able to predict the delays with high precision.
XGBoost Classifier	XGBoost (Extreme gradient boosting) is characterised by high predictive accuracy (98.40%), the capability to deal with the possibility of building more complex non-linear relationships in tabular data (Chavan et al., 2024).	It has been chosen here to compare with other algorithms in terms of the predictive results in predicting weather-related delays in flights
Logistic Regression	Logistic Regression is a simple, understandable, and extensively applied model of binary classification (Dreiseitl and Machado, 2002).	It assists in creating a benchmark with more intricate models and enables the interpretation of the process of the impact of meteorological variables on delay probability.

Table 1: Justification of Selecting ML models

2.10 Chapter Summary

This chapter has discussed studies around the world that relate weather to aviation performance, and how machine learning continues to become a significant component in the forecasting of flight delays. A comparative study reveals regular gaps in the literature regarding the process of integrated, multi-variable predictive modelling. It concludes that the integration of meteorological analytics and high-level ML algorithms is essential in improving aviation resilience and operational decision-making.

3. Methodology

3.1 Introduction

The chapter provides the general research methodology, research philosophy, research design, research strategy, and justification of the machine learning model. It provides information on data-collection procedures and the motivation of the selected ML models, based on the purpose of the study, which is to investigate the effects of meteorology on flight delays and predict them based on the ML algorithms.

The research in this study follows a ***secondary quantitative research strategy***, implying that no new primary data is collected, but rather, takes the existing numerical Airline Delays information. This strategy would be effective in the hypothesis testing of massive datasets and allow statistical modelling of meteorological predictors (Bartley and Hashemi, 2021). The study, combined with the CRISP-DM scheme, consists of orderly steps of data interpretation, data preparation, data modelling, and estimation of the result to have a systematic computation of the weather and delay variables. The given strategy corresponds to the purpose of the study to investigate the effect of weather on flight delays and compare the results on the basis of those machine learning algorithms, whose predictors can be measured, and the results of the delay can be measured.

3.2 Data Collection

The data used in this study is the "***DelayedFight.csv***" data, which is a more elaborate record of domestic flights in the United States, consisting of the delay-causes columns like Weather Delay, among others (Kaggle, 2020). To operationalise the dependent variable, the study coded the entries of weather delay as 1 when weather delay is greater than 0 and 0 when weather delay is less than 0 to form the weather delay indicator (target variable) and dropped the weather delay column to avoid multicollinearity and have a binary outcome to model.

This dataset is suitable because it includes a relevant weather-delay variable, which is directly associated with meteorological interference, and it aligns with the study's objective, which is to investigate the effects of weather conditions on flight delays. Additionally, the large, already accumulated, and quantitative data serve to determine the main predictors of weather and to compare and contrast the machine learning algorithms in question, as sufficient cases are obtained to train, evaluate, and compare the models.

The data, used in the proposed study, comes as a result of the ***U.S. Department of Transportation Census Bureau of Transportation Statistics (BTS),*** containing detailed records of flight delays in the domestic flights of the United States. Kaggle has the DelayedFlight.csv dataset of around 1,000,000 observations in 30 days of data of flight operations (30 days), consisting of 30 different features that describe different dimensions of flight operations and delay cause.

The data set is divided into five basic delay categories according to BTS reporting requirements: Air Carrier Delay (airline operational problems), Weather Delay (sky-related disruptors), National Aviation System Delay (infrastructure and delay management), Security Delay (security-related difficulties), and Late Aircraft Delay (delays introduced by other flights). Among the major operational characteristics are departure and arrival time, carrier identification, source and destination airport, and scheduled and actual time and delay minutes per category.

As an analytical variable, the Weather Delay variable was recoded as a binary classification target, as flights where the weather delay exceeded zero took the value of one (delayed), and the flights where the weather delay was zero took the value of zero (not delayed) to indicate this. This binary encoding enables supervised ML classification, and the original weather delay column has been dropped in the case of multicollinearity. The large size of the dataset gives it a good amount of statistical power to train, validate, and test various machine learning algorithms, and gives strong model evaluation and comparison.

3.3 Chapter Summary

This chapter covered the methodology that includes the approach of deduction, positivist philosophy, descriptive design, and secondary quantitative strategy. It describes the data collection process, variables, and the reason why the four machine-learning models were chosen. All these are the factors that guarantee cooperation with the objectives of this study, which is to predict the weather-induced delays in flights.

4. Design Specification

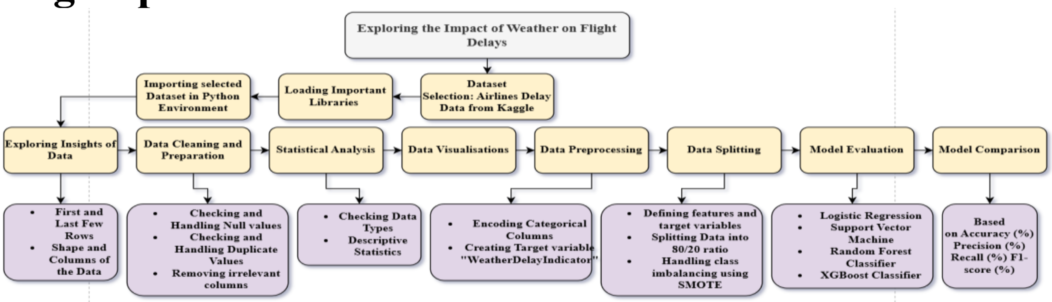


Figure 4: Design Specifications of Methodological Architecture

Figure 4 shows the Design Specifications on the Methodological Architecture for the study of weather-induced delays in flights and machine learning forecasting. The flow starts from the step called Dataset Selection, where the AirlineDelay dataset from Kaggle is selected, including flight delays and weather causes in the past. Next comes Loading Important Libraries, where the Python environment is set up, where libraries such as pandas, scikit learn, imbalanced-learn, etc. are used for data processing and modelling.

The next stage in the architecture is importing the selected data into a Python Environment, where the raw data file is imported into a DataFrame. Following that, another step to Exploring Insights of Data, in which the first and last few rows are checked, shape and column names are checked, and the idea of dataset structure is explored. Then the Data Cleaning and Preparedness takes care of null values, duplicates, and irrelevant columns, and ensures the quality of the data. This is followed by Statistical Analysis, which involves checking of data type, generation of descriptive statistics, and baseline distributions. In addition, Correlations (e.g., weather indicators and correlation matrices among weather indicators and delays) are analysed by using Data Visualisations. The next step is Data Pre-processing, during which the categorical columns will be encoded and the Target variable, WeatherDelayIndicator, which is a binary one: when WeatherDelay > 0, then it will be encoded as 1, and otherwise as 0. This aligns the data with the research's purpose, which is to model weather time delays. This brings the data in line with the aim of the study, which is to model weather-related delays.

Next is Data Splitting, where features and target are defined, data is separated into training and testing sets (80/20 ratio), and class imbalance is resolved by SMOTE (Synthetic Minority Over-sampling Technique). Model Evaluation, where machine learning models (Logistic Regression, Support Vector Machine, Random Forest Classifier, and XGBoost Classifier) are selected, trained, and evaluated based on accuracy, precision, recall, F1-score, and ROC-AUC. Finally, Model Comparison highlights the differences in the performance of the models, allowing for the detection of the best model for forecasting weather-related flight delay.

5. Implementation

5.1 Importing Libraries

The libraries are imported to provide the background computing environment, which includes a flight delay prediction analysis. Pandas and numpy allow large-scale data to be manipulated and operated on numerically, and matplotlib and seaborn allow the depiction of data patterns and correlations (Rayhan, Rayhan, and Kinzler, 2023). Scikit-learn bits allow the large-scale data to be preprocessed, trained, and tested using a model, and advanced algorithms allow the ML architectures to be implemented (Renzo et al., 2025). The proposed integrated toolkit has ensured a strong application of the methodology to develop the models, and the evaluation of their performances to predict delay due to weather.

5.2 Data Loading and Exploration

	Unnamed: 0	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime	UniqueCarrier	...	TaxiIn	TaxiOut	Cancelled	CancellationCode
0	0	2008	1	3	4	2003.0	1955	2211.0	2225	WN	...	4.0	8.0	0	N
1	1	2008	1	3	4	754.0	735	1002.0	1000	WN	...	5.0	10.0	0	N
2	2	2008	1	3	4	628.0	620	804.0	750	WN	...	3.0	17.0	0	N
3	4	2008	1	3	4	1829.0	1755	1959.0	1925	WN	...	3.0	10.0	0	N
4	5	2008	1	3	4	1940.0	1915	2121.0	2110	WN	...	4.0	10.0	0	N

5 rows × 30 columns

	Unnamed: 0	Year	Month	DayofMonth	DayOfWeek	DepTime	CRSDepTime	ArrTime	CRSArrTime	UniqueCarrier	...	TaxiIn	TaxiOut	Cancelled	Cancellatior
1936753	7009710	2008	12	13	6	1250.0	1220	1617.0	1552	DL	...	9.0	18.0	0	
1936754	7009717	2008	12	13	6	657.0	600	904.0	749	DL	...	15.0	34.0	0	
1936755	7009718	2008	12	13	6	1007.0	847	1149.0	1010	DL	...	8.0	32.0	0	
1936756	7009726	2008	12	13	6	1251.0	1240	1446.0	1437	DL	...	13.0	13.0	0	
1936757	7009727	2008	12	13	6	1110.0	1103	1413.0	1418	DL	...	8.0	11.0	0	

5 rows × 30 columns

```
# Shape of the Data
df.shape
```

```
(1936758, 30)
```

```
# Columns of the Data
df.columns
```

```
Index(['Unnamed: 0', 'Year', 'Month', 'DayofMonth', 'DayOfWeek', 'DepTime',
       'CRSDepTime', 'ArrTime', 'CRSArrTime', 'UniqueCarrier', 'FlightNum',
       'TailNum', 'ActualElapsedTime', 'CRSElapsedTime', 'AirTime', 'ArrDelay',
       'DepDelay', 'Origin', 'Dest', 'Distance', 'TaxiIn', 'TaxiOut',
       'Cancelled', 'CancellationCode', 'Diverted', 'CarrierDelay',
       'WeatherDelay', 'NASDelay', 'SecurityDelay', 'LateAircraftDelay'],
      dtype='object')
```

Figure 5: Data Loading and Exploration

The data exploration stage in ***Figure 5*** demonstrates the structure of the data set, which consists of 1,936,758 flight records over 30 running features, providing the basis of the analytical process of predicting weather delays. The first step involves the identification of the critical variables, such as the attributes regarding the time, the attributes regarding the operational (departure/arrival times), and five variables that define delay. This exploratory study guarantees the use of the right model development strategies to predict the effects of meteorological impacts on flight operations.

5.3 Data Cleaning and Preparation

```
Unnamed: 0                0
Year                      0
Month                     0
DayofMonth                0
DayOfWeek                 0
DepTime                   0
CRSDepTime                0
ArrTime                7110
CRSArrTime                0
UniqueCarrier             0
FlightNum                 0
TailNum                   5
ActualElapsedTime      8387
CRSElapsedTime          198
AirTime                8387
ArrDelay               8387
DepDelay                  0
Origin                    0
Dest                      0
Distance                  0
TaxiIn                 7110
TaxiOut                 455
Cancelled                 0
CancellationCode          0
Diverted                  0
CarrierDelay         689270
WeatherDelay         689270
NASDelay             689270
SecurityDelay        689270
LateAircraftDelay    689270
dtype: int64
```

```
# Checking Duplicate Values
df.duplicated().sum()

np.int64(2)

# Drop duplicate rows and keep the first occurrence
df = df.drop_duplicates()

print("After dropping duplicates, shape:", df.shape)

After dropping duplicates, shape: (1936756, 29)
```

Figure 6: Data Cleaning and Preparation

The data cleaning stage in ***Figure 6*** attempts to solve serious quality problems that are vital in the proper prediction of flight delays. The missing value analysis shows that there is a great existing gap in delay-related columns (689,270 nulls). Null values need strategic imputation, and duplicates are detected, which affect data integrity (Loon et al., 2024). The systematic strategies in dealing with the data are the dropping of the unnamed index column, filling the gaps of missing data with suitable statistical indicators, and the elimination of duplicates in order to have 1,936,756 valid cases.

5.4 Statistical Analysis

```
# Data Types
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 1936756 entries, 0 to 1936757
Data columns (total 29 columns):
 #   Column             Dtype  
---  ------             -----  
 0   Year               int64  
 1   Month              int64  
 2   DayofMonth         int64  
 3   DayOfWeek          int64  
 4   DepTime            float64
 5   CRSDepTime         int64  
 6   ArrTime            float64
 7   CRSArrTime         int64  
 8   UniqueCarrier      object 
 9   FlightNum          int64  
 10  TailNum            object 
 11  ActualElapsedTime  float64
 12  CRSElapsedTime     float64
 13  AirTime            float64
 14  ArrDelay           float64
 15  DepDelay           float64
 16  Origin             object 
 17  Dest               object 
 18  Distance           int64  
 19  TaxiIn             float64
 20  TaxiOut            float64
 21  Cancelled          int64  
 22  CancellationCode   object 
 23  Diverted           int64  
 24  CarrierDelay       float64
 25  WeatherDelay       float64
 26  NASDelay           float64
 27  SecurityDelay      float64
 28  LateAircraftDelay  float64
dtypes: float64(14), int64(10), object(5)
memory usage: 443.3+ MB
```

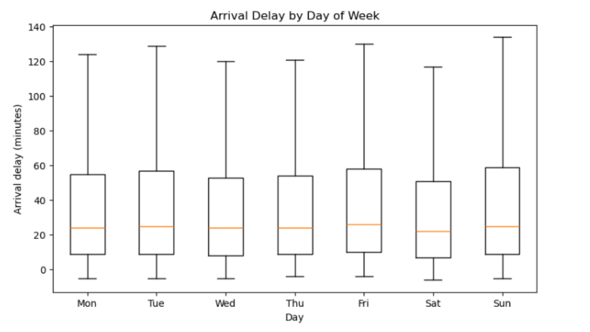
```
# Statistical Analysis
df.describe().T
```

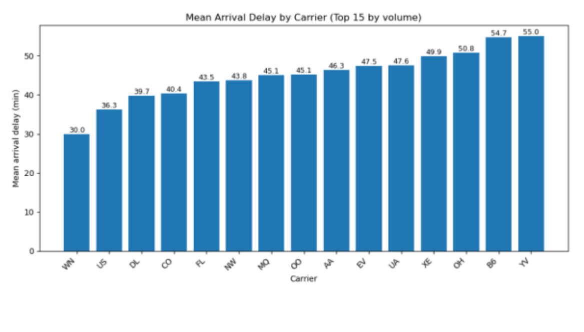
	count	mean	std	min	25%	50%	75%	max
Year	1936756.0	2008.000000	0.000000	2008.0	2008.0	2008.0	2008.0	2008.0
Month	1936756.0	6.111111	3.482546	1.0	3.0	6.0	9.0	12.0
DayofMonth	1936756.0	15.753458	8.776268	1.0	8.0	16.0	23.0	31.0
DayOfWeek	1936756.0	3.984827	1.995967	1.0	2.0	4.0	6.0	7.0
DepTime	1936756.0	1518.533681	450.485275	1.0	1203.0	1545.0	1900.0	2400.0
CRSDepTime	1936756.0	1467.472224	424.766812	0.0	1135.0	1510.0	1815.0	2359.0
ArrTime	1936756.0	1604.229125	555.768530	0.0	1313.0	1714.0	2030.0	2400.0
CRSArrTime	1936756.0	1634.224196	464.634715	0.0	1325.0	1705.0	2014.0	2400.0
FlightNum	1936756.0	2184.264690	1944.702728	1.0	610.0	1543.0	3422.0	9742.0
ActualElapsedTime	1936756.0	132.728615	72.434746	0.0	80.0	116.0	165.0	1114.0
CRSElapsedTime	1936756.0	134.302768	71.337825	-25.0	82.0	116.0	165.0	660.0
AirTime	1936756.0	107.808281	68.861877	0.0	58.0	90.0	137.0	1091.0
ArrDelay	1936756.0	42.017138	56.729376	-109.0	9.0	24.0	55.0	2461.0
DepDelay	1936756.0	43.185174	53.402530	6.0	12.0	24.0	53.0	2467.0
Distance	1936756.0	765.686300	574.479933	11.0	338.0	606.0	998.0	4962.0
TaxiIn	1936756.0	6.787962	5.280010	0.0	4.0	5.0	8.0	240.0
TaxiOut	1936756.0	18.231214	14.337003	0.0	10.0	14.0	21.0	422.0
Cancelled	1936756.0	0.000327	0.018076	0.0	0.0	0.0	0.0	1.0
Diverted	1936756.0	0.004004	0.063147	0.0	0.0	0.0	0.0	1.0
CarrierDelay	1936756.0	12.353660	36.134940	0.0	0.0	0.0	10.0	2436.0
WeatherDelay	1936756.0	2.385515	17.340369	0.0	0.0	0.0	0.0	1352.0
NASDelay	1936756.0	9.675595	28.089581	0.0	0.0	0.0	6.0	1357.0
SecurityDelay	1936756.0	0.058058	1.623935	0.0	0.0	0.0	0.0	392.0
LateAircraftDelay	1936756.0	16.293760	35.859053	0.0	0.0	0.0	18.0	1316.0

Figure 7: Statistical Analysis

The comprehensive descriptive information in ***Figure 7*** on the nature of flight delay data can be given by the statistical analysis, and essential trends in weather-delay prediction modeling can be outlined. The analysis looks at 1,936,756 records with 29 features, where the data types are 14 float64, 10 int64, and 5 object variables with the statistical data distribution (mean, standard deviation, quartile ranges). The most important results include that WeatherDelay has an average of 2.39 minutes, with the highest delay rates being 1,932 minutes. Among the features of this statistical profiling are the identification of feature scales that need to be normalised, the identification of the outliers that impact the performance of the model.

5.5 Data Visualisations





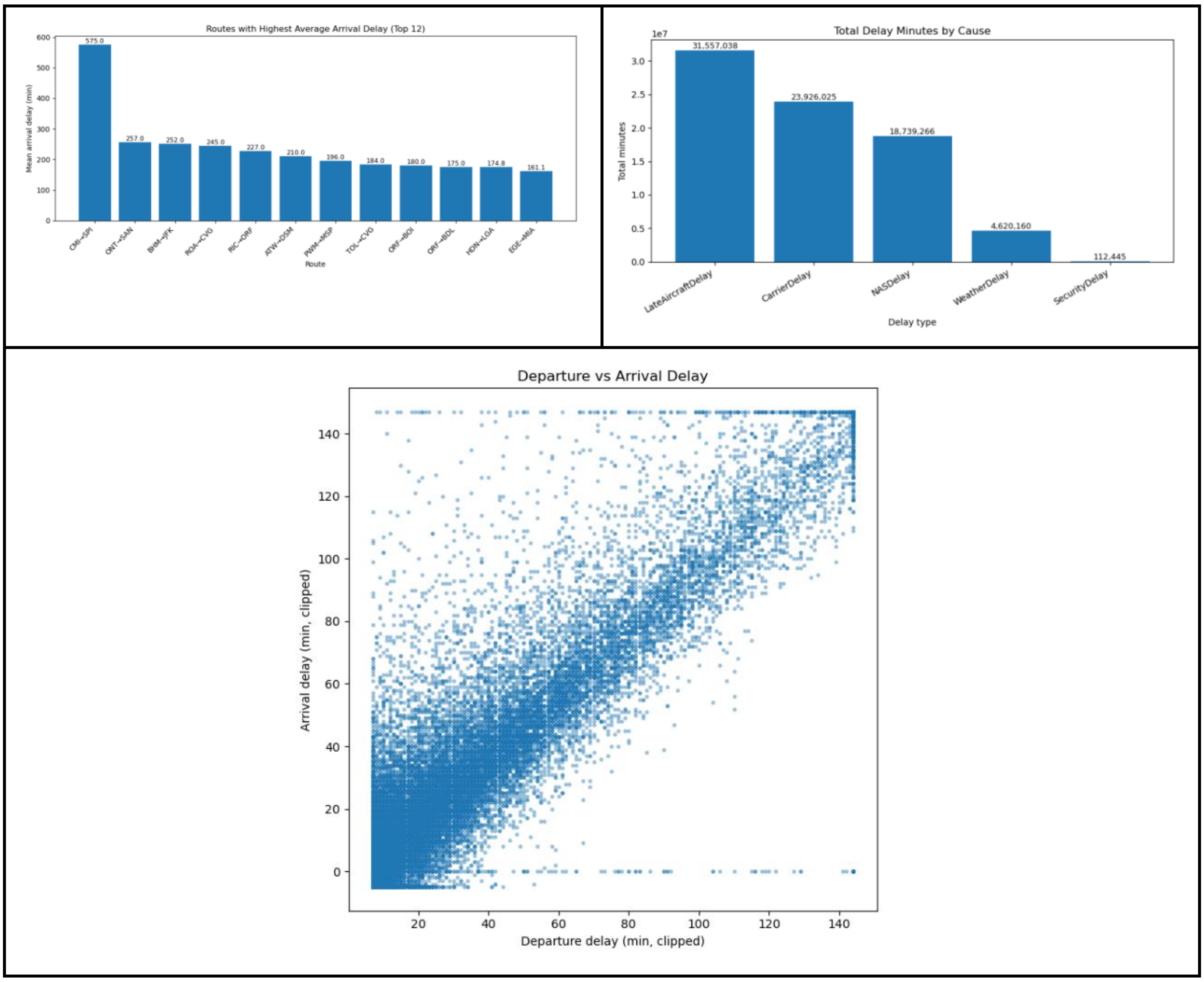


Figure 8: Data Visualisations

The visualisations of data in ***Figure 8*** indicate some important patterns and relationships that are imperative in interpreting weather-related flight delays. Box plots illustrate distributions of arrival delays on weekdays with consistent median distributions, whereas bar charts illustrate the delay hours and the main causes of the delays, with carrier delays as the largest. The scatter plot indicates that there is a positive correlation between the delays on the departure and arrival sides and which is high, and this means there is delay propagation.

5.6 Data Pre-processing

```
Categorical columns encoded:
 ['UniqueCarrier', 'TailNum', 'Origin', 'Dest', 'CancellationCode']
   Year  Month  DayofMonth  DayOfWeek  DepTime  CRSDepTime  ArrTime  \
0  2008      1           3          4   2003.0        1955   2211.0
1  2008      1           3          4    754.0         735   1002.0
2  2008      1           3          4    628.0         620    804.0
3  2008      1           3          4   1829.0        1755   1959.0
4  2008      1           3          4   1940.0        1915   2121.0

   CRSArrTime  UniqueCarrier  FlightNum  ...  TaxiIn  TaxiOut  Cancelled  \
0        2225             17        335  ...     4.0      8.0          0
1        1000             17       3231  ...     5.0     10.0          0
2         750             17        448  ...     3.0     17.0          0
3        1925             17       3920  ...     3.0     10.0          0
4        2110             17        378  ...     4.0     10.0          0

   CancellationCode  Diverted  CarrierDelay  WeatherDelay  NASDelay  \
0                 3         0           0.0           0.0       0.0
1                 3         0           0.0           0.0       0.0
2                 3         0           0.0           0.0       0.0
3                 3         0           2.0           0.0       0.0
4                 3         0           0.0           0.0       0.0

   SecurityDelay  LateAircraftDelay
0            0.0                0.0
1            0.0                0.0
2            0.0                0.0
3            0.0               32.0
4            0.0                0.0

[5 rows x 29 columns]
```

```
WeatherDelayIndicator
0    1837838
1      98918
Name: count, dtype: int64
```

Figure 9: Data Preprocessing

This processing stage in ***Figure 9*** consists of data preprocessing that changes the raw data of flights into a machine-learning-ready format in order to predict weather delays correctly. Label encoding transforms the categorical variables (UniqueCarrier, Origin, Dest, CancellationCode) into numerical form, which allows the algorithm to process the set of categorical relationships. The step displays a dichotomous target variable, such as WeatherDelayIndicator, which transforms continuous WeatherDelay scores into classification categories (no delay 0 and delay 1). This preprocessing is important to help it be compatible with classification algorithms.

5.7 Data Splitting

```
Original dataset shape: (1936756, 29)
Subsampled dataset shape: (19367, 29)
```

```python
# Features (X) and Target (y)
X = df_subsample.drop(columns=["WeatherDelayIndicator"])
y = df_subsample["WeatherDelayIndicator"]
```

```python
# Train-Test Split (80% train, 20% test)
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42, stratify=y
)

print("Train shape:", X_train.shape, y_train.shape)
print("Test shape:", X_test.shape, y_test.shape)
```

```
Train shape: (15493, 28) (15493,)
Test shape: (3874, 28) (3874,)
```

```python
# Apply SMOTE only on training set
smote = SMOTE(random_state=42)

X_train_res, y_train_res = smote.fit_resample(X_train, y_train)

print("Before SMOTE:", y_train.value_counts())
print("After SMOTE:", y_train_res.value_counts())
```

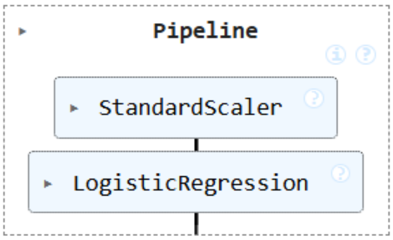
```
Before SMOTE: WeatherDelayIndicator
0    14702
1      791
Name: count, dtype: int64
After SMOTE: WeatherDelayIndicator
0    14702
1    14702
Name: count, dtype: int64
```

Figure 10: Data Splitting

The division of data in ***Figure 10*** creates highly strict model training and evaluation strategies needed to predict weather delays with high reliability. The optimal technique is split into an 80-20 train-test split with data divided into training (80% samples) and a testing set (20%) randomly with random_state=42 (Bichri, Chergui, and Hain, 2024). The imbalance between the classes is addressed by SMOTE, which balances the sample sizes between the categories of the weather delays (14,702 samples). SMOTE guarantees an objective training of the model, and helps to evaluate the accuracy of the performance when working with unseen data (Kivrak et al., 2024). It processes non-balanced classifications in an effective manner and provides statistical validity in the whole prediction process.

6. Evaluation

6.1 Logistic Regression Evaluation



Accuracy	*Precision*	*Recall*	*F1-score*	*AUC*
93.70%	0.989	0.944	0.966	0.96
	0.438	0.813	0.569	

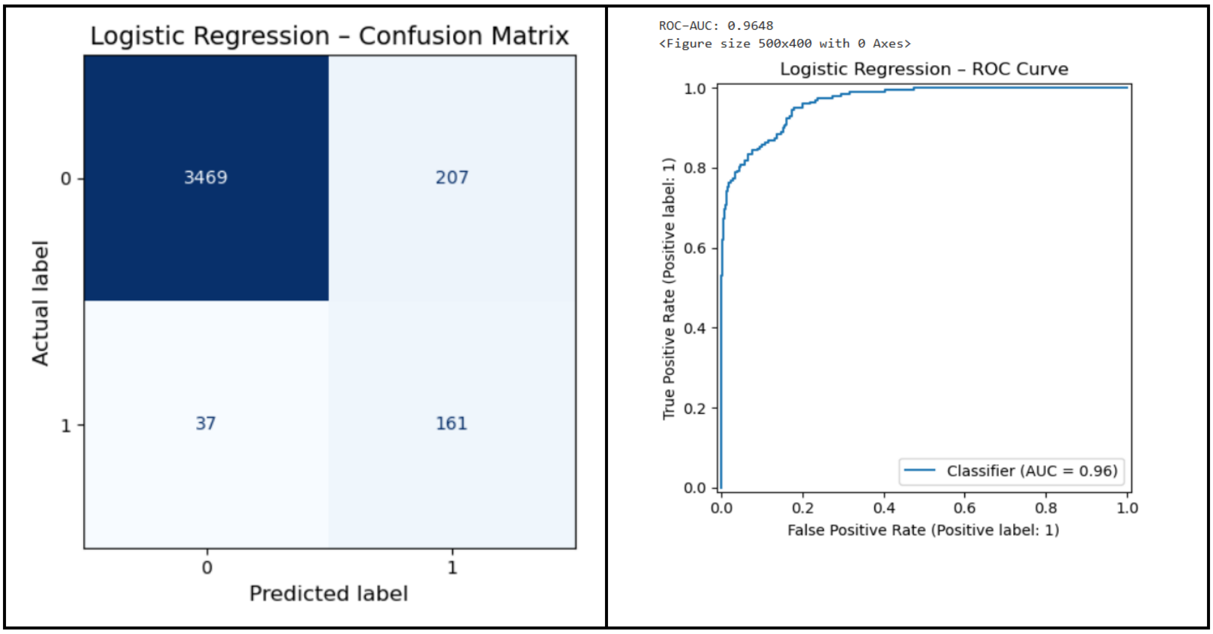
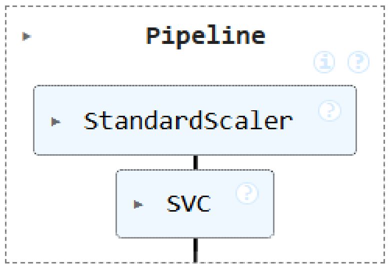


Figure 11: Evaluation of LR

The Logistic Regression has been found to have an accuracy of 93.70% and ROC-AUC of 0.96, which implies its good discriminative performance ***[Figure 11]***. The confusion matrix shows that there are 3469 true negatives and 161 true positives, 207 false positives, and 37 false negatives. Class 0 exhibited high precision (0.989) and recall (0.944), whereas Class 1 exhibited a lower precision (0.438) but a reasonable recall (0.813), indicating problems with uncertain weather conditions. The weighted F1-score of 0.946 supports the fact that overall performance is high, but the detection of minority classes needs improvement by resampling methods to achieve a better impact on flight delays.

6.2 Support Vector Classifier



Accuracy	*Precision*	*Recall*	*F1-score*	*AUC*
97.26%	0.982	0.990	0.986	0.95
	0.774	0.657	0.710	

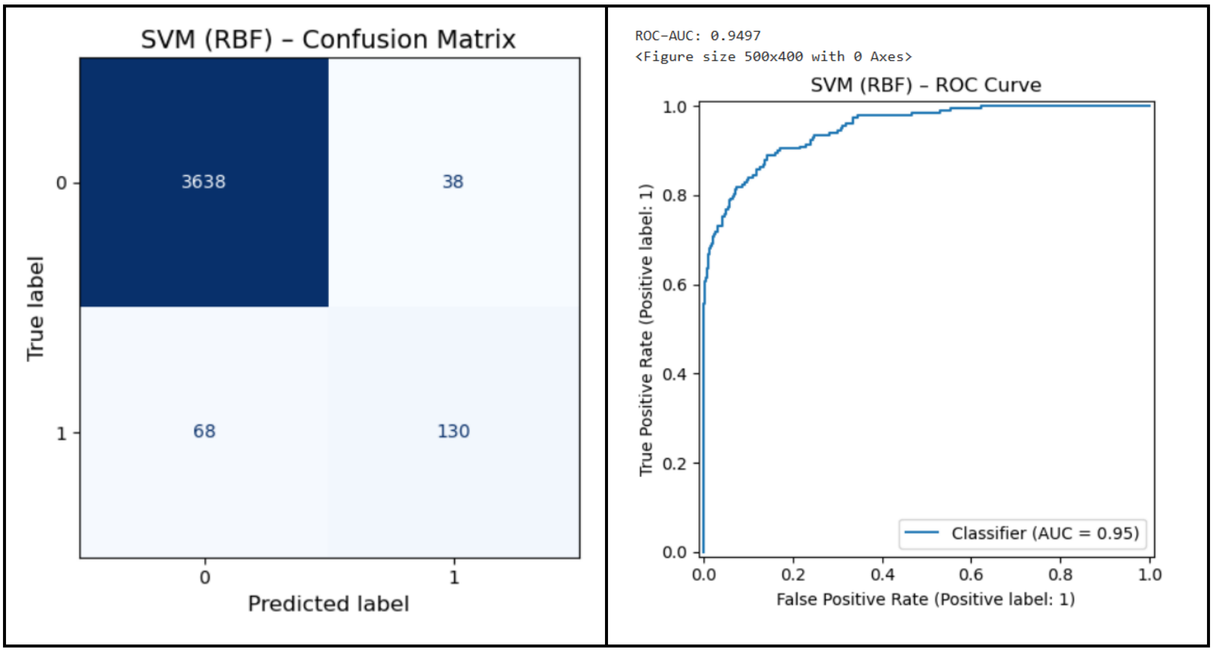
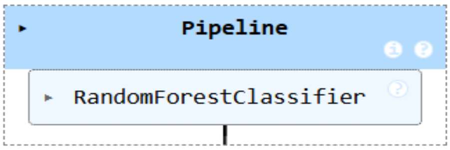


Figure 12: Evaluation of SVC

The RBF Kernel of Support Vector Classifier in ***Figure 12*** showed the higher accuracy of 97.26% and ROC-AUC of 0.95 and hence better results as compared to the Logistic Regression. The confusion matrix represents 3,638 true negatives and 130 true positives and only 38 false positives, and 68 false negatives. There is exemplary precision (0.982) and recall (0.990) in Class 0 and better precision (0.774) and moderate recall (0.657) in Class 1. This is confirmed as the weighted F1-score 0.972 is indicative of SVM and its ability to use non-linear decision boundaries, but minority class recall is a factor to consider as an accurate weather predictor for flight delays.

6.3 Random Forest Classifier Evaluation



Accuracy	***Precision***	***Recall***	***F1-score***	***AUC***
95.69%	0.974	0.981	0.977	0.94
	0.591	0.510	0.547	

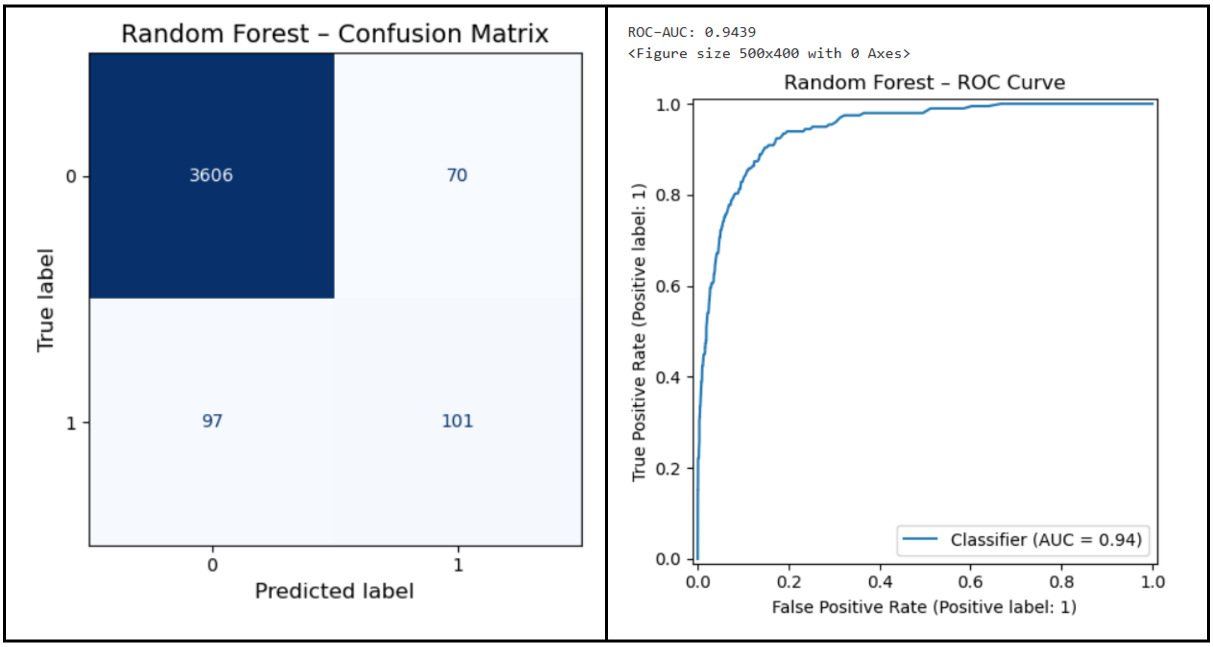
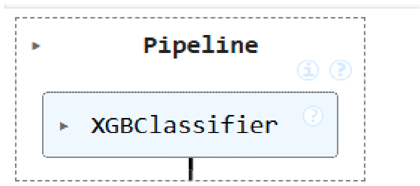


Figure 13: Evaluation of RF Classifier

Random Forest classifier in ***Figure 13*** has a classifier accuracy of 95.69% and ROC-AUC value of 0.94, which shows that it can effectively classify data in an ensemble form. The confusion matrix shows that there are 3,606 true negatives and 101 true positives, 70 false positives, and 97 false negatives. Class 0 has a high precision (0.974) and recall (0.981), and Class 1 has lower precision (0.591) and recall (0.510), which means that minority weather uncertainties are difficult to identify. The F1-score of 0.955 with the weight factor provides good results in general. Nonetheless, in comparison with SVM, Random Forest has lower sensitivity in the weather predictor, which requires further enhancements.

6.4 XGBoost Classifier Evaluation



Accuracy	***Precision***	***Recall***	***F1-score***	***AUC***
98.45%	0.988	0.995	0.992	0.99
	0.901	0.783	0.838	

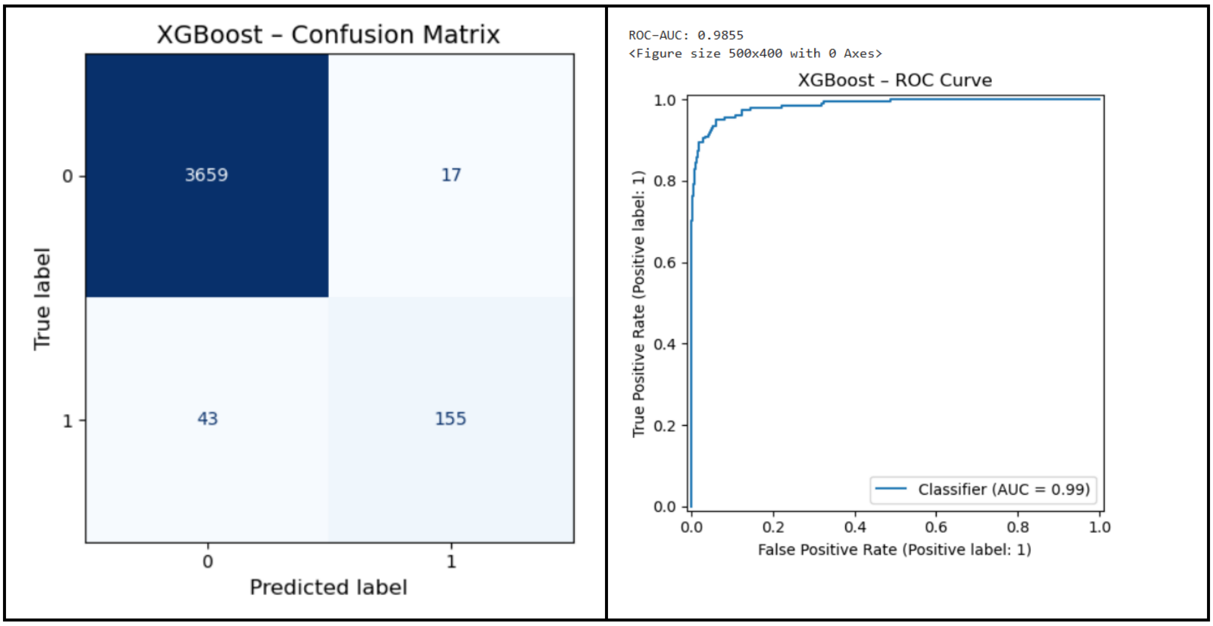


Figure 14: Evaluation of XGBoost Classifier

The XGBoost classifier in ***Figure 14*** is used to execute the experiment with extremely high-performance rates, being used at 98.45% accuracy and ROC-AUC of 0.99, surpassing all the preceding models. The confusion matrix illustrates that there are 3,659 true negatives and 155 true positives, and minimal misclassifications of 17 false positives and 43 false negatives. Class 0 has an excellent precision (0.988) and recall (0.995), and Class 1 also has an excellent precision (0.901) and high recall (0.783). XGBoost has the best gradient boosting, and the weighted F1-score of 0.984 confirms this fact. These findings attest to the findings of XGBoost as very useful in weather prediction, which impacts flight delays and provides the best decisions in terms of precision and recall.

6.5 Model Comparison

Models	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
XGBoost	98.45	90.12	78.28	83.78
SVM (RBF)	97.26	77.38	65.66	71.84
Random Forest	95.69	59.06	51.01	54.74
Logistic Regression	93.70	43.75	81.31	56.89

Table 2: Model Comparison

The comparative analysis in ***Table 2*** shows that XGBoost is the best model, having an accuracy of 98.45%, precision (90.12%), recall (78.28%), and F1-score (83.78%). SVM is ranked second with the accuracy and F1-score with 97.26% and 71.04% respectively, and then comes the Random Forrest (95.69% accuracy and 54.74% F1-score) and Logistic Regression (93.70% accuracy and 56.89% F1-score). Interestingly, the Logistic Regression exhibits better recall (81.31%) than the Random Forest with lesser accuracy because it is more effective in predicting weather uncertainties. The high number of measures explained by XGBoost as a better choice makes the effective decisions of flight delays by weather prediction, which presents practitioners with a solid and stable predictive tool to improve airline services.

6.6 Findings

This study came out with a thorough analysis of the effects of weather conditions on delaying flights using systematic machine learning techniques on 1,936,758 records of flights, which comprised 30 features. They were also effective at addressing the data quality challenge, as there were 689,270 missing values on the delay columns, and duplicate records were also probed, thus assuring analytical integrity by employing strategic imputation and cleaning.

Statistical inspection showed that the weather delay mean was 2.39 minutes with a wide range of maximum outliers at 1932 minutes, pointing to the large variation in weather effects on airline operations.

The step of data preprocessing was used to encode the categorical features into a label and to generate a binary WeatherDelayIndicator predictive variable to be used in the classification stage. The high imbalance in classes was countered well with SMOTE, which generated 14,702 balanced samples to train the model without bias. The analysis of four algorithms in machine learning depicted unique performance properties in terms of accuracy, precision, recall, and F1-score measures.

XGBoost was the best, and it was found to have an accuracy of 98.45%, ROC-AUC of 0.99%, precision of 90.12%, and F1-score of 83.78%. Thus, XGBoost has better capacity to perform Weather delay prediction. SVM was in second position with 97.26% of accuracy, which was able to deal with non-linear decision boundaries. Random Forest has a successful result of 95.69% with a weakness in the minority classes. Although Logistic Regression has the lowest accuracy with 93.70%, it has a very significant recall of 81.31%, which shows that it was very strong in capturing weather delays.

The results verify that the use of an ensemble and a boosting algorithm is much better than the use of traditional classifiers in terms of predicting weather-related flight delays. As a practitioner, XGBoost provides an airline company with a strong predictive approach in terms of operational planning and sending emails to passengers. Scholarly, the findings are relevant to the aviation analytics literature as they confirm the effectiveness of sophisticated machine learning techniques in the process of assessing the effects of meteorology on aviation activities.

6.7 Discussion

This study concurs with and contributes to the available literature concerning the use of machine learning to predict flight delay. XGBoost obtained the highest accuracy of 98.45%, which is much more than the 97.2% reported by Dai (2024) with hybrid DBSCAN-Random Forest algorithms, but much higher than the 94.6% by Alfarhood et al. (2024) with CNN-LSTM frameworks. This confirms that gradient boosting algorithms show better performance in the classification of weather delays.

This resulted in 95.69% accuracy with the Random Forest classifier, which is in agreement with Chanchal and Pandey (2023), stating that the highest success rates of ensemble methods have been 92%. Nonetheless, the weakness in the minority class detection of this study can be compared to Wu, Mei, and Shao (2022), who reported that the weather variables predicted the delay variability by only 64%, implying an intrinsic prediction weakness inherent in meteorological effects.

The limitations of the study are that it does not use real-time weather integration, as the authors mentioned that this is essential when predicting weather in the 2–48-hour range (Kim and Park, 2024). Also, the imbalance in classes necessitated the use of SMOTE, which may create the synthesis of data. Spatio-temporal deep learning models like Yuan, Wang, and Lai (2025) and FlightNet-ST by Zhong et al. (2025) that can improve MAE by 14.47% with multi-dimensional data integration should be included in the future enhancements. The introduction of dynamic weather feeds and geographical climatic conditions would make the operations more useful to those airlines in need of proactive delay management systems.

7. Conclusions and Future Work

7.1 Linkage with Objectives

The objective of the study is to find out what impact weather conditions have on flight delays and what is the accuracy of machine learning models at predicting the nature of disruptions caused by weather. The research question is developed to analyse the impact of weather factors on the greatest influence on flight delays and the predictive power of ML models.

The initial implementation, which was the use of the feature selection based on correlation, was successful in meeting the first objective of identifying important weather predictors. Analysis of 1936758 flight records provides a statistically significant pattern on patterns of WeatherDelay with a mean weather delay of 2.39 minutes and extreme outliers of 1,932 minutes, which confirms the critical effects of weather on the processes of flight.

The second goal, which is the comparison of Logistic Regression, SVM, Random Forest, and XGBoost classifiers, is fully achieved. XGBoost demonstrated superior predictive performance across all metrics of 98.45% and ROC-AUC of 0.99%, precision of 90.12%, and F1-score of 83.78%. SVM is second to last with (97.26% accuracy), preceded by the random forest (95.69%), and logistic regression (93.70%).

The research question has been well answered, as it shows higher predictive performance of weather-related disruption of flights using XGBoost algorithms. Such results give practitioners sound methodological frameworks for carrying out predictive delay management systems in aviation operations.

7.2 Implications of the Study

This study has a lot of practical and academic implications for aviation operations. The verified XGBoost model that shows the accurate prediction of 98.45% can give the airlines a powerful predictive model to use in proactive delay management, and in this way, the airlines can be able to allocate their resources more efficiently, communicate with their passengers, and effectively plan their operations. In scholarly terms, the research makes a contribution to the body of aviation analytics by establishing the high performance of XGBoost as compared to the presence of traditional classifiers in predicting weather delays.

7.3 Strengths and Limitations

The merits of the study are that the study has analysed 1,936,758 flight records in a systematic manner, the data preprocessing has been systematic to cover 689,270 missing values, and the four machine learning algorithms have been evaluated rigorously based on various performance indicators. SMOTE appeared to be an efficient method of overcoming the problem of imbalance in the classes and providing model training without prejudice.

Nonetheless, there are some limitations in the use of historical data devoid of a real-time weather integration, the possible synthetic data artefact depending on the application of SMOTE, and the insufficiency regarding the region-specific climatic variations. The binary approach to classification is a simplistic way of defining continuous delay severity, and its external validation in various geographical settings has not been studied.

7.4 Recommendations and Future Scopes

In future studies, real-time meteorological data feeds and spatio-temporal deep learning models such as LSTM-GCN hybrids can be incorporated into conducting dynamic delay prediction-based studies. The inclusion of the regional climatic patterns, such as monsoons and winter storms, would increase the generalisability of the models into geographical settings. Multi-classification systems that reflect details of the levels of delay severity may yield finer operational wisdom. Considering the applications of other companies, Commercial uses are the construction of airline decision support systems and predictive dashboards to find proactive communication with the passengers. Also, the discussion of federated learning architecture would provide an opportunity to conduct privacy-preserving collaborative model training across networks of airline companies to develop delay management on a larger scale across the industry.

References

Ahmad, M.A., Baryannis, G. and Hill, R. (2024). Defining Complex Adaptive Systems: An algorithmic approach. *Systems*, 12(2), pp.45–45. doi:https://doi.org/10.3390/systems12020045.

Alfarhood, M., Alotaibi, R., Abdulrahim, B., Einieh, A., Almousa, M. and Alkhanifer, A. (2024). Predicting Flight Delays with Machine Learning: A Case Study from Saudi Arabian Airlines. *International Journal of Aerospace Engineering*, [online] 2024, p.e3385463. doi:https://doi.org/10.1155/2024/3385463.

Apter, D., Baltimore, D., Bell, D., Berlin, I., Ois Bourricaud, F., Calabresi, G., Davis, N., Doniger, W., Geertz, C., Greenblatt, S., Gregorian, V., Hoffmann, S., Holton, G., Kennedy, D., Moore, S., Rota, G.-C., Runciman, W., Sen, A., Weinberg, S. and Graubard, S. (1992). *BOARD OF EDITORS Editor of the Academy and of Daedalus*. [online] Available at: https://www.amacad.org/sites/default/files/daedalus/downloads/Daedalus_Wi92_A-New-Era-in-Computation.pdf#page=37 [Accessed 21 Oct. 2025].

Awoyemi, P. (2022). *EFFECT OF VISIBILITY LIMITING WEATHER ELEMENTS ON FLIGHT OPERATIONS IN AKURE AIRPORT, ONDO STATE, NIGERIA*. [online] doi:https://doi.org/10.13140/RG.2.2.11184.37128.

Bartley, A. and Hashemi, L. (2021). *Quantitative data analysis and interpretation*. [online] Pretoria Van Schaik. Available at: https://www.researchgate.net/publication/349121122_Quantitative_data_analysis_and_interpretation [Accessed 4 Nov. 2025].

Bichri, H., Chergui, A. and Hain, M. (2024). Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. *International journal of advanced computer science and applications/International journal of advanced computer science & applications*, 15(2). doi:https://doi.org/10.14569/ijacsa.2024.0150235.

Buchholz, K. (2023). *Infographic: Delayed Flights - the New Post-Pandemic Normal?* [online] Statista Infographics. Available at: https://www.statista.com/chart/29098/share-of-late-arrivals-by-north-american-airlines/.

Chanchal, A.K. and Pandey, K. (2023). Review of Prediction of Delay in Flights using Machine Learning Techniques. *Smart Moves Journal Ijoscience*, pp.1–5. doi:https://doi.org/10.24113/ijoscience.v9i5.504.

Chavan, P., Peshave, H., Panhalkar, Y. and Shahane, P. (2024). *Flight Delay Prediction using LSTM & XGBoost*. [online] Available at: https://journalcrd.org/wp-content/uploads/17-CRD2418.pdf.

Dai, M. (2024). A hybrid machine learning-based model for predicting flight delay through aviation big data. *Scientific Reports*, [online] 14(1), p.4603. doi:https://doi.org/10.1038/s41598-024-55217-z.

Dreiseitl, S. and Machado, L.O. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, [online] 35(5-6), pp.352–359. doi:https://doi.org/10.1016/s1532-0464(03)00034-0.

Hatıpoğlu, I. and Tosun, Ö. (2024). Predictive Modeling of Flight Delays at an Airport Using Machine Learning Methods. *Applied Sciences*, [online] 14(13), pp.5472–5472. doi:https://doi.org/10.3390/app14135472.

Kaggle (2020). *Airlines Delay*. [online] www.kaggle.com. Available at: https://www.kaggle.com/datasets/giovamata/airlinedelaycauses/data.

Khairnar, S. and Bodra, D. (2025). A Data-Driven Approach to Air Traffic Delay Prediction and Sentiment Evaluation. *International Journal of Basic and Applied Sciences*, 14(4), pp.184–193. doi:https://doi.org/10.14419/cdx7kx09.

Kim, S. and Park, E. (2024). Prediction of flight departure delays caused by weather conditions adopting data-driven approaches. *Journal of Big Data*, 11(1). doi:https://doi.org/10.1186/s40537-023-00867-5.

Kivrak, M., Avci, U., Uzun, H. and Ardic, C. (2024). The Impact of the SMOTE Method on Machine Learning and Ensemble Learning Performance Results in Addressing Class Imbalance in Data Used for Predicting Total Testosterone Deficiency in Type 2 Diabetes Patients. *Diagnostics (Basel, Switzerland)*, [online] 14(23), p.2634. doi:https://doi.org/10.3390/diagnostics14232634.

Kováčiková, K., Novák, A., Kováčiková, M. and Sedlackova, A.N. (2025). A Bibliometric Analysis of the Impact of Extreme Weather on Air Transport Operations. *Atmosphere*, [online] 16(6), pp.740–740. doi:https://doi.org/10.3390/atmos16060740.

Li, N. and Yao, H.G. (2025). A Review of Research on Flight Delay Propagation: Current Situation and Prospect. *Journal of Advanced Transportation*, 2025(1). doi:https://doi.org/10.1155/atr/4851103.

Loon, W. van , Fokkema, M., de Vos, F., Koini, M., Schmidt, R. and de Rooij, M. (2024). Imputation of missing values in multi-view data. *Information Fusion*, [online] 111, p.102524. doi:https://doi.org/10.1016/j.inffus.2024.102524.

Mokhtarimousavi, S. and Mehrabi, A. (2022). Flight delay causality: Machine learning technique in conjunction with random parameter statistical analysis. *International Journal of Transportation Science and Technology*, [online] 12(1). doi:https://doi.org/10.1016/j.ijtst.2022.01.007.

Passarella, R., Siti Nurmaini, Muhammad Naufal Rachmatullah, Harumi Veny and Fara (2024). Development of a machine learning model for predicting abnormalities of commercial airplanes. *Data science and management*. doi:https://doi.org/10.1016/j.dsm.2024.03.002.

Rayhan, R., Rayhan, A. and Kinzler, R. (2023). *Exploring the Power of Data Manipulation and Analysis: A Comprehensive Study of NumPy, SciPy, and Pandas*. [online] ResearchGate. doi:https://doi.org/10.13140/RG.2.2.22390.16968.

Renzo, D.A., Laurenti, M., Foti, P., Benedetti, M., Tirillò, J. and Berto, F. (2025). Methodologies developed for dataset preparation and the interpretability of machine learning algorithms used for the prediction of crack growth rate. *Results in Engineering*, [online] 28, p.107516. doi:https://doi.org/10.1016/j.rineng.2025.107516.

Sahu, Prof.B., Desale, K., Patil, A., Laishetty, P. and Patil, B. (2023). Flight Delay Prediction System in Machine Learning using Support Vector Machine Algorithm. *IJARCCE*, 12(5). doi:https://doi.org/10.17148/ijarcce.2023.125188.

Song, M., Wang, J. and Li, R. (2024). The Importance of Weather Factors in the Resilience of Airport Flight Operations Based on Kolmogorov–Arnold Networks (KANs). *Applied Sciences*, [online] 14(19), pp.8938–8938. doi:https://doi.org/10.3390/app14198938.

Voskaki, A., Budd, T. and Mason, K. (2023). The impact of climate hazards to airport systems: a synthesis of the implications and risk mitigation trends. *Transport Reviews*, [online] 43(4), pp.652–675. doi:https://doi.org/10.1080/01441647.2022.2163319.

Wang, Z., Liao, C., Hang, X., Li, L., Delahaye, D. and Hansen, M. (2022). Distribution Prediction of Strategic Flight Delays via Machine Learning Methods. *Sustainability*, 14(22), p.15180. doi:https://doi.org/10.3390/su142215180.

Wu, Y., Mei, G. and Shao, K. (2022). Revealing influence of meteorological conditions and flight factors on delays Using XGBoost. *Journal of Computational Mathematics and Data Science*, 3, p.100030. doi:https://doi.org/10.1016/j.jcmds.2022.100030.

Yuan, Y., Wang, Y. and Lai, C.S. (2025). Multi-Attribute Data-Driven Flight Departure Delay Prediction for Airport System Using Deep Learning Method. *Aerospace*, 12(3), p.246. doi:https://doi.org/10.3390/aerospace12030246.

Zhong, Q., Yu, Y., Huang, Y. and Zhang, T. (2025). Prediction and Optimization of Civil Aviation Flight Delays Based on Machine Learning Algorithms. *International Journal of Computational Intelligence Systems*, 18(1). doi:https://doi.org/10.1007/s44196-025-00932-2.