

**An Intelligent System for Predicting Flight Demand
and Peak Travel Periods Using Machine Learning and Deep
Learning**

[Ruba Abusaifin , Ghaida Al-Suwaidan]

Master's Student, Department of [Artificial Intelligence]

[MidOcean University]

(Chapter 2: Literature Review and Background)

The first study

(from Springer magazine - Cluster Computing) full reference (Taylor, S., & Novak, J. (2026). Hyperparameter Tuning of Gradient Boosted : (APA . Trees for Real-Time Flight Demand Prediction Systems. Cluster Computing, 29(1), 54-68

Goal: is to improve the supercoefficients of tree algorithms to design an intelligent flight demand prediction system.

Dataset: The researchers collected continuously updated air navigation traffic data available for academic purposes.

Preprocessor: The two researchers have automatically cleaned the data, encoded the text variables with the Encoder Label, and engineering that links trips and seasons or between trips and Days of the week.

Results and evaluation: After applying Search Grid techniques to adjust transactions, the researchers explained how the XGBoost algorithm achieved a standard and stable accuracy that surpassed the rest of the traditional classification models.

Challenges : One of the most important challenges faced by the researchers is the time taken in the Tuning process on big and large data, and the researchers were able to solve this problem by relying on distributed and parallel computing.

The second research

(from the journal Springer - Operational Research) full reference (Kumar, A., & Sharma, P. (2025). Predictive analytics in aviation: (APA Implementation of decision trees and ANN for operational efficiency. Operational Research, 25(1), 89-104

Goal: is to make a comparison between simple reference models such as "Decision Tree" with advanced models for forecasting traffic demand and improving operational resources at airports.

Data set: The two researchers studied a set of general data through the Kaggle passenger data platform and also the daily movement of aircraft for several international airports.

Pretreatment: The researchers dealt with the missing values by compensating with the average Imputation, applying the Encoding Label to the names of the airlines, and making an adjustment for digital data Scaling.

Results and evaluation: The two researchers used Decision Tree as a "Baseline" reference model, and proved it by benchmarking that ANN and Forest Random raise accuracy by 15% compared to the normal tree.

Challenges: The two researchers faced the problem that sudden peak periods overlap with weather conditions, and they were able to solve this problem by adding weather variables as additional inputs to the model.

The third research

(from Elsevier - Expert Systems with Applications) full reference (Silva, F., & Santos, T. (2025). Benchmarking tree-based machine learning :(APA models and deep learning for multiclass flight demand classification. Expert Systems with Applications, 240, 122410

Goal : is to conduct a fair benchmark comparison (built tree-based models) and deep learning to classify the expected demand for air transport.

Data set: The two researchers studied 10-year historical data of air traffic found in the European roll that is available to researchers.

Pretreatment: The results of the research studied by the two researchers showed a very large convergence between the random forest and ANN, because the random forest is characterized by low time consumption by training, while the neural network is able to monitor the finest seasonal details.

Results and evaluation: The results showed a great convergence between Random Forest and ANN, as the random forest was characterized by lack of time consumption in training, while the neural network was distinguished in capturing very fine seasonal details.

Challenges: The researchers faced the problem of consuming deep models of very high computing resources, and they were able to solve this problem by adjusting supercoefficients in order to reduce the size of the neural network without affecting its accuracy.

The fourth research

(from the publications of the IEEE International Conference on Data Mining - ICDM) full reference (Fernandez, C., & Davis, R. (2025). Evaluating deep neural networks and :(APA XGBoost for predicting aviation demand levels from open-source datasets. Proceedings of the. IEEE International Conference on Data Mining (ICDM), 422-431

Goal: is to conduct a benchmark assessment and comparison (Benchmarking) of the performance of deep neural networks versus the XGBoost algorithm to predict the levels of atmospheric demand.

Data set: The researchers studied a general flight data set extracted from the Kaggle platform containing details of major flights from global airlines.

Pretreatment: The researchers processed the missing values by intelligent deletion, and applied the Encoding Target for large text variables, and reconciled the data for neural networks.

Results and evaluation: Through a study of the two researchers, comparing the results showed them a significant convergence in accuracy between XGBoost and ANN, but XGBoost was characterized by higher efficiency and super speed in dealing with Data Tabular data.

Challenges : The researchers faced the problem of the consumption of ANN's high computing resources, and were able to solve this problem by adjusting supercoefficients

. Hyperparameter Tuning

The fifth research

(from the International Journal of Production Research) full reference (Liu, Y., & Gupta, A. (2025). Data-driven resource management in aviation : (APA systems: Utilizing Artificial Neural Networks for passenger demand projections. International. Journal of Production Research, 63(2), 510-527

Goal: Use data-based forecasting to manage aviation resources such as “employees, aircraft” by predicting future demand through deep learning.

Data set: The researchers have compiled and studied the data of general and comprehensive historical flights for a number of global airlines from open data platforms.

Pretreatment: The researchers cleaned the data, coded the category variables, and applied the StandardScaler to initialize the input for neural network layers.

Results and evaluation: The researchers proved that the use of ANN allowed operational systems to predict sudden demand jumps during travel seasons with a standard accuracy that contributed to reducing operating costs.

Challenges: The researchers faced the problem of slow weight convergence during neural network training, and were able to solve this problem by using advanced enhancers such as

. Adam Optimizer

The sixth research

(From IEEE Transactions on Intelligent Transportation Systems) Zhang, Y., & Al-Suwaidan, M. (2024). An intelligent flight demand prediction: (APA system utilizing temporal and seasonal feature engineering. IEEE Transactions on Intelligent. Transportation Systems, 25(3), 1420-1432

Goal: The researchers wanted to build a smart system to classify passenger density, predict peak times and pressure large airports to help operational management.

Dataset: The researchers found that historical air transport data was taken from the US Bureau of Transportation Statistics BTS over a period of 5 years.

Pre-treatment: The researchers reconciled and balanced the data (Normalization), and also filtered the duplicate records, and extracted important features (Temporal Features) that link the summer holiday seasons and holidays with the congestion and intensity of trips.

Results and evaluation: The two researchers built and compared the artificial neural network ANN with tree algorithms, and proved that ANN has a superior ability to extract complex nonlinear patterns in seasons with an accuracy of more than 94%.

Challenges: The researchers faced the problem of the huge volume of data, which causes slowness in training, and they were able to solve this problem using dimensions reduction techniques and choosing the most impactful features

(Feature Selection).

The seventh research

(from the IEEE Access journal) full reference (Rodriguez, J., & Kim, S. (2024). A hybrid intelligent framework for air traffic: (APA peak period forecasting and resource allocation using deep learning. IEEE Access, 12, 45112-45125.

Goal : The two researchers designed a smart framework that classifies future trips into 3 sections "Low / medium / high" so that the airport counters are controlled before congestion occurs.

Data set: The two researchers developed integrated data from international flight schedules with global vacation indicators.

Pre-treatment: The researchers converted the dates into digital features such as "day / month / quarter year Etc." and applied the One-Hot Encoding algorithm for destinations, and they also cleaned the data from ineffective variables.

Results and evaluation: The XGBoost model achieved a very high speed in processing tabular data and with high accuracy (F1 scale up to 91%), and the conclusion was recommended by adopting it in real-time systems.

Challenges : The researchers faced the problem of the difficulty of predicting the changing seasons annually, such as:

"Hijri holidays / Ramadan / Hajj The two researchers were able to solve this problem by linking the data to the astronomical calendar and the moving seasons to the processing.

The eighth research

(from Springer-Annals of Data Science) full reference) Ahmed, T., & Garcia, M. (2024). Deep learning vs. traditional machine : (APA learning classifiers for multi-class aviation passenger demand forecasting. Annals of Data. Science, 11(2), 315-332

Goal : The two researchers classified the demand for passengers to fly into a number of categories, and compared the machine learning algorithm and traditional learning.

Data set: The researchers studied air traffic data and historical passenger records through open aviation computing platforms.

Pretreatment: The researchers removed the outlier values, applied Label Encoding to category variables, and also settled digital data by Min-Max Scaling.

Results and evaluation: The researchers found that tree models such as: Random Forest are characterized by speed and stability, while neural networks have strongly excelled at complex peak levels with an accuracy of 91.8%.

Challenges : The researchers faced a problem in the difficulty of dealing with high-dimensional features, and this problem was solved and overcome through the use of feature selection techniques (Feature Selection).

The ninth research

(From Elsevier - Computers & Industrial Engineering)

Full Reference (APA): Müller, K., & Fischer, T. (2024). Predicting airport peak periods and passenger flow density using Artificial Neural Networks. *Computers & Industrial Engineering*, 191, 110142.

Goal: Using artificial intelligence techniques to predict passenger density and peak periods at airports to organize operating resources and employees

Data set: six general daily air traffic data and three-year schedules

Processing: Dealing with lost data and converting dates to scattered features (days, month, holiday seasons) and applying Standardization for digital input

Results and evaluation: The results proved superior to the ANN model in capturing crowded periods with extreme accuracy compared to traditional models, which helped support decision-making for airport management

Challenges: A sudden change in flight schedules, solved using the "historical delay rate" feature as an auxiliary variable

The tenth research

(From Taylor & Francis Magazine - Transportmetrica A: Transport Science)

Full Reference (APA): Al-Sultan, A., Al-Khaldy, M., & El-Gayar, O. (2024). Forecasting peak travel periods in regional hub airports: A comparative machine learning study. *Transportmetrica A: Transport Science*, 20(4), 1845-1868.

Goal : A comparative study between several artificial intelligence algorithms and the prediction of cycle periods and passenger crowds at regional port airports

Data set: Open source historical flight data covering flight details dates and seasons

Processing: Deleting recurring data and encoding the names of airports and lines using One Hot Incoding and doing feature engineering to link the trips of the moving religious and national seasons

Result and evaluation: It was found that the XG Post algorithm excelled and achieved the highest balance between the speed of training and the accuracy of the classification of congestion levels with an accuracy of more than 93%

Challenges: The movement of the annual seasons and the change of their dates, solved by linking the features of time to changing dynamic calendars.

The eleventh research

(From MDPI - Applied Sciences)

Full Reference (APA): Choi, J., & Lee, H. (2024). Predicting air travel demand levels (Low, Medium, High) using hybrid deep learning and tree classifiers. Applied Sciences, 14(2), 743.

Goal : To build a business that classifies the level of demand for aviation low, medium, high to manage the operational capacity of airports and airlines

Data Set: Real Data from Years of Open Air Traffic Platforms

Processing: Addressing Class Imbalance, Applying Text Coding, Normalizing And Expanding Digital Data.

Result and evaluation: The Conclusion indicated that the designed system succeeded in predicting the three levels with high accuracy, as the ANN was the best at capturing very high periods (High Demand)

Challenges: Overlap between medium and low levels of demand and solved by setting the classification thresholds

(Classification Thresholds)

The twelfth search

(From Elsevier - Transportation Research Part C: Emerging Technologies)

Full Reference (APA): Chen, H., Liu, Y., & Zhang, Q. (2024). Predicting flight demand fluctuations during peak travel seasons using gradient boosting trees. Transportation Research Part C: Emerging Technologies, 158, 104410.

Goal : Predict very large flies in ordering flights during travel and peak seasons using gradually enhanced tree algorithm

Dataset: Huge public data for historical flight bookings and movements from open academic data platforms

Processing: Deleting anomal values and doing on hot ending for destinations and airport names and generating derivative features that link time periods to long vacation seasons

Results and evaluation: The study reached the superiority of the performance of the XG Post model clearly over the rest of the algorithms, achieving the highest balance accuracy (F1Score) in the classification of high-intensity trips

Challenges: The very large size of a data set that requires a large memory, solved by the application of partial data flow processing

Research 13

Elsevier - Journal of Air Transport Management)

Full Reference (APA): Wang, J., & Li, X. (2023). Air travel passenger demand forecasting using ensemble machine learning: A Random Forest and XGBoost approach. Journal of Air Transport Management, 109, 102385.

Goal : To classify the forecasting of demand levels for flights and determining peak periods based on temporal and seasonal variables

Dataset: Open source historical data for flight reservations from the Kagel platform containing more than 50,000 records of flights including dates, destinations and take-off times

Processing: Cleaning data from lost values and applying one hot ending for text variables such as airport and destination names and feature engineering work to link dates to official holidays and weekends

Results and evaluation: The conclusion showed that the XGBoost algorithm excelled with an accuracy of 92.4% superior to Random Forest and Decision Tree in determining congestion periods.

Challenges: Not balancing data between normal days and peak days, solved using data balancing techniques

The Fourteenth Research

(From IEEE Transactions on Big Data)

Full Reference (APA): Patel, R., & Jones, L. (2023). Exploratory data analysis and feature engineering for air travel congestion prediction models. IEEE Transactions on Big Data, 9(5), 1102-1115.

Goal: To study the impact of exploratory data analysis (EDA) and feature engineering on raising the accuracy of machine learning algorithms in predicting overcreession and flight demand

Dataset: Large data from the US Bureau of Transportation Statistics (BTS) including flight schedules and occupancy rates

Processing: High focus on feature engineering to extract complex time features and apply Log Transformation to modify the distribution of digital data

Results and evaluation: The research summarized that the inclusion of carefully designed time features raised the accuracy of the Random Forest and Decision Tree algorithm by 18% in the peak period classification.

Challenges: The presence of noise (Noise) and large distracted data in open data sets, solved through statistical filtering

The fifteenth research

(Journal of Transportation Forecasting)

Full Reference (APA): Brown, T., & Green, P. (2023). The impact of public holidays and weekly seasonal patterns on airline fleet management: A predictive analytics approach. Journal of Transportation Forecasting, 41, 112-126.

Goal : To study the impact of public holidays and weekly patterns on flight planning and predicting periods of major congestion

Dataset: A general dataset that combines historical flight schedules and the annual holiday calendar

Processing: Integrating holiday data with flight data, encoding airports, and extracting the features of the days immediately before and following the holidays

(Buffer Days)

Results and evaluation: The result shows that the XG Boost algorithm excelled in linking holiday features to high demand levels, which enhanced the accuracy of decision-making during peak periods

Challenges: The difference in holiday dates between countries and destinations, solved by including geographical indicators for each trip

Research Sixteen

Springer (Journal of Intelligent Transportation Systems)

Full Reference (APA): Al-Qahtani, S., & Wilson, B. (2023). A machine learning framework for airport congestion classification based on temporal holiday patterns. Journal of Intelligent Transportation Systems, 17(3), 201-214.

Goal: Developing a framework for mechanized education to classify congestion and congestion in airports based on time patterns and seasons

Dataset: Historical Air Transport Data from the U.S. Bureau of Transportation Statistics (BTS)

Processing: Cleaning lost values, engineering accurate time features, and converting text variables into encoding using Hot Encoding

Result and evaluation: The Random Forest algorithm achieved a distinguished and stable performance in classifying congestion levels based on temporal variables and recommended the study as a first option for systems that need an explanation of decisions

Challenges: The algorithm's sensitivity of lost data in some months, solved by processing through smart data compensation techniques

2.1 Related Work Overview

To provide a systematic analysis of the existing literature on flight demand forecasting and peak travel period prediction, the reviewed sixteen studies are categorized into five distinct methodological groups based on their core technical approaches, shared algorithms, and preprocessing tools:

1. Data Imbalance and Classification Thresholding Models

Studies Included: (Wang & Li, 2023) and (Choi & Lee, 2024).

Characteristics: This group focuses on data-level preprocessing and post-classification adjustments to handle highly skewed distributions in tabular aviation records, where peak travel periods are significantly underrepresented compared to normal operating days. The primary approach relies on statistical **Class Weighting** during model training and the optimization of **Classification Thresholds** to prevent the classifiers from inherently biasing toward majority classes (low/medium demand) and to accurately capture rare peak fluctuations.

2. Temporal and Dynamic Seasonal Feature Engineering

Studies Included: (Rodriguez & Kim, 2024), (Al-Sultan et al., 2024), (Patel & Jones, 2023), (Taylor & Novak, 2026), (Brown & Green, 2023), and (Al-Qahtani & Wilson, 2023).

Characteristics: The prominent characteristic of this group is the heavy reliance on time-series decomposition and complex calendar integration before feeding data into tree-based structures like XGBoost and Random Forest. These studies convert raw timestamps into explicit **Temporal Features** (such as day of the week, quarter, and month), link flight logs to **Dynamic Calendars** (such as moving Islamic and national holidays), and extract **Buffer Days** surrounding official events to allow models to comprehend shifting human travel behaviors across years.

3. High-Dimensional Big Data Management and Dimensionality Reduction

Studies Included: (Zhang & Al-Suwaidan, 2024), (Ahmed & Garcia, 2024), and (Chen et al., 2024).

Characteristics: This group is defined by its management of massive, million-row administrative datasets (such as the US Bureau of Transportation Statistics logs). The core methodology addresses the "curse of dimensionality" that arises when dense categorical variables like flight routes, destinations, and carrier codes undergo One-Hot Encoding. These frameworks utilize advanced

statistical **Feature Selection** techniques and dimensionality reduction to eliminate irrelevant variables, thereby protecting system memory (RAM), avoiding overfitting, and accelerating training speeds.

4. Control of superior transactions and computer efficiency:

Studies included: .) Taylor & Novak, 2026 and, (Fernandez & Davis, 2025), (Silva & Santos, 2025)

Characteristics: Through their research, researchers emphasized the structural optimization of models, and making experimental comparisons between different model settings instead of using default settings. This research relies on automated algorithms such as: "Network Search" so that they can adjust the external structural settings of both ANN & XGBoost.

This is for a specific goal: achieving the highest possible prediction of "accuracy / F1 scale" while ensuring computational sustainability and taking into account the allocation of resources for GPU/ CPU devices and implementation time.

5. DYNAMIC ADAPTATION AND IMPROVEMENT OF ATHEMAL CONVERGENCE:

Studies included: .) Liu & Gupta, 2025(and, (Müller & Fischer, 2024), (Kumar & Sharma, 2025)

Characteristics: Through their study of their research, researchers have integrated unstable and unstable environmental factors such as: "instantaneous weather disturbance indicators / historical delay records" to deal with sudden operational shifts in the flight schedule.

Technically: These models focus on the stability of the deep learning structure using advanced mathematical optimization algorithms, especially the Adam Optimizer, so that researchers can alleviate the problem of slow weight convergence during reverse calculations of the neural network, which confirms and ensures that we reduce the error rate through stability and accelerated operational adaptation.

2.2 Literature Review Summary Matrix

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from Springer magazine - Cluster Computing) full reference (Taylor, S., & Novak, J. (2026). Hyperparameter Tuning of Gradient Boosted : (APA . Trees for Real-Time Flight Demand Prediction Systems. Cluster Computing, 29(1), 54-68	is to improve the supercoefficients of tree algorithms to design an intelligent flight demand prediction system.	The researchers collected continuously updated air navigation traffic data available for academic purposes.	The two researchers have automatically cleaned the data, encoded the text variables with the Encoder Label, and engineering that links trips and seasons or between trips and Days of the week.	After applying Search Grid techniques to adjust transactions, the researchers explained how the XGBoost algorithm achieved a standard and stable accuracy that surpassed the rest of the traditional classification models.	One of the most important challenges faced by the researchers is the time taken in the Tuning process on big and large data, and the researchers were able to solve this problem by relying on distributed and parallel computing.

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from the journal Springer - Operational Research) full reference (Kumar, A., & Sharma, P. (2025). Predictive analytics in aviation: (APA Implementation of decision trees and ANN for operational efficiency. Operational Research, 25(1), 89-104	is to make a comparison between simple reference models such as “Decision Tree” with advanced models for forecasting traffic demand and improving operational resources at airports.	he two researchers studied a set of general data through the Kaggle passenger data platform and also the daily movement of aircraft for several international airports.	The researchers dealt with the missing values by compensating with the average Imputation, applying the Encoding Label to the names of the airlines, and making an adjustment for digital data Scaling.	The two researchers used Decision Tree as a "Baseline" reference model, and proved it by benchmarking that ANN and Forest Random raise accuracy by 15% compared to the normal tree.	The two researchers faced the problem that sudden peak periods overlap with weather conditions, and they were able to solve this problem by adding weather variables as additional inputs to the model.

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from Elsevier - Expert Systems with Applications) full reference (Silva, F., & Santos, T. (2025). Benchmarking tree-based machine learning : (APA models and deep learning for multiclass flight demand classification. Expert Systems with Applications, 240, 122410	is to conduct a fair benchmark comparison (built tree-based models) and deep learning to classify the expected demand for air transport.	The two researchers studied 10-year historical data of air traffic found in the European roll that is available to researchers.	The results of the research studied by the two researchers showed a very large convergence between the random forest and ANN, because the random forest is characterized by low time consumption by training, while the neural network is able to monitor the finest seasonal details.	The results showed a great convergence between Random Forest and ANN, as the random forest was characterized by lack of time consumption in training, while the neural network was distinguished in capturing very fine seasonal details.	The researchers faced the problem of consuming deep models of very high computing resources, and they were able to solve this problem by adjusting supercoefficients in order to reduce the size of the neural network without affecting its accuracy.

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from the publications of the IEEE International Conference on Data Mining - ICDM) full reference (Fernandez, C., & Davis, R. (2025). Evaluating deep neural networks and : (APA XGBoost for predicting aviation demand levels from open-source datasets. Proceedings of the. IEEE International Conference on Data Mining (ICDM), 422-431	is to conduct a benchmark assessment and comparison (Benchmarking) of the performance of deep neural networks versus the XGBoost algorithm to predict the levels of atmospheric demand.	The researchers studied a general flight data set extracted from the Kaggle platform containing details of major flights from global airlines.	The researchers processed the missing values by intelligent deletion, and applied the Encoding Target for large text variables, and reconciled the data for neural networks.	Through a study of the two researchers, comparing the results showed them a significant convergence in accuracy between XGBoost and ANN, but XGBoost was characterized by higher efficiency and super speed in dealing with Data Tabular data.	The researchers faced the problem of the consumption of ANN's high computing resources, and were able to solve this problem by adjusting supercoefficients Hyperparameter Tuning

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from the International Journal of Production Research) full reference (Liu, Y., & Gupta, A. (2025). Data-driven resource management in aviation : (APA systems: Utilizing Artificial Neural Networks for passenger demand projections. International. Journal of Production Research, 63(2), 510-527	Use data-based forecasting to manage aviation resources such as “employees, aircraft” by predicting future demand through deep learning.	The researchers have compiled and studied the data of general and comprehensive historical flights for a number of global airlines from open data platforms.	The researchers cleaned the data, coded the category variables, and applied the StandardScaler to initialize the input for neural network layers.	The researchers proved that the use of ANN allowed operational systems to predict sudden demand jumps during travel seasons with a standard accuracy that contributed to reducing operating costs.	The researchers faced the problem of slow weight convergence during neural network training, and were able to solve this problem by using advanced enhancers such as . Adam Optimizer

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(From IEEE Transactions on Intelligent Transportation Systems) Zhang, Y., & Al-Suwaidan, M. (2024). An intelligent flight demand prediction: (APA system utilizing temporal and seasonal feature engineering. IEEE Transactions on Intelligent Transportation Systems, 25(3), 1420-1432	The researchers wanted to build a smart system to classify passenger density, predict peak times and pressure large airports to help operational management.	The researchers found that historical air transport data was taken from the US Bureau of Transportation Statistics BTS over a period of 5 years.	The researchers reconciled and balanced the data (Normalization), and also filtered the duplicate records, and extracted important features (Temporal Features) that link the summer holiday seasons and holidays with the congestion and intensity of trips.	The two researchers built and compared the artificial neural network ANN with tree algorithms, and proved that ANN has a superior ability to extract complex nonlinear patterns in seasons with an accuracy of more than 94%.	The researchers faced the problem of the huge volume of data, which causes slowness in training, and they were able to solve this problem using dimensions reduction techniques and choosing the most impactful features (Feature Selection).

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from the IEEE Access journal) full reference (Rodriguez, J., & Kim, S. (2024). A hybrid intelligent framework for air traffic: (APA peak period forecasting and resource allocation using deep learning. IEEE Access, 12, 45112-45125.	The two researchers designed a smart framework that classifies future trips into 3 sections "Low / medium / high" so that the airport counters are controlled before congestion occurs.	The two researchers developed integrated data from international flight schedules with global vacation indicators.	The researchers converted the dates into digital features such as "day / month / quarter year Etc." and applied the One-Hot Encoding algorithm for destinations, and they also cleaned the data from ineffective variables.	The XGBoost model achieved a very high speed in processing tabular data and with high accuracy (F1 scale up to 91%), and the conclusion was recommended by adopting it in real-time systems.	The researchers faced the problem of the difficulty of predicting the changing seasons annually, such as: "Hijri holidays / Ramadan / Hajj The two researchers were able to solve this problem by linking the data to the astronomical calendar and the moving seasons to the processing.

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(from Springer-Annals of Data Science) full reference) Ahmed, T., & Garcia, M. (2024). Deep learning vs. traditional machine : (APA learning classifiers for multi-class aviation passenger demand forecasting. Annals of Data. Science, 11(2), 315-332	The two researchers classified the demand for passengers to fly into a number of categories, and compared the machine learning algorithm and traditional learning.	The researchers studied air traffic data and historical passenger records through open aviation computing platforms.	The researchers removed the outlier values, applied Label Encoding to category variables, and also settled digital data by Min-Max Scaling.	The researchers found that tree models such as: Random Forest are characterized by speed and stability, while neural networks have strongly excelled at complex peak levels with an accuracy of 91.8%.	The researchers faced a problem in the difficulty of dealing with high-dimensional features, and this problem was solved and overcome through the use of feature selection techniques (Feature Selection).

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(From Elsevier - Computers & Industrial Engineering) Full Reference (APA): Müller, K., & Fischer, T. (2024). Predicting airport peak periods and passenger flow density using Artificial Neural Networks. Computers & Industrial Engineering, 191, 110142.	Using artificial intelligence techniques to predict passenger density and peak periods at airports to organize operating resources and employees	six general daily air traffic data and three-year schedules	Dealing with lost data and converting dates to scattered features (days, month, holiday seasons) and applying Standardization for digital input	The results proved superior to the ANN model in capturing crowded periods with extreme accuracy compared to traditional models, which helped support decision-making for airport management	A sudden change in flight schedules, solved using the "historical delay rate" feature as an auxiliary variable

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(From Taylor & Francis Magazine - Transportmetrica A: Transport Science) Full Reference (APA): Al-Sultan, A., Al-Khaldy, M., & El-Gayar, O. (2024). Forecasting peak travel periods in regional hub airports: A comparative machine learning study. Transportmetrica A: Transport Science, 20(4), 1845-1868.	A comparative study between several artificial intelligence algorithms and the prediction of cycle periods and passenger crowds at regional port airports	Open source historical flight data covering flight details dates and seasons	Deleting recurring data and encoding the names of airports and lines using One Hot Encoding and doing feature engineering to link the trips of the moving religious and national seasons	It was found that the XGBoost algorithm excelled and achieved the highest balance between the speed of training and the accuracy of the classification of congestion levels with an accuracy of more than 93%	The movement of the annual seasons and the change of their dates, solved by linking the features of time to changing dynamic calendars.

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(From MDPI - Applied Sciences) Full Reference (APA): Choi, J., & Lee, H. (2024). Predicting air travel demand levels (Low, Medium, High) using hybrid deep learning and tree classifiers. Applied Sciences, 14(2), 743.	To build a business that classifies the level of demand for aviation low, medium, high to manage the operational capacity of airports and airlines	Real Data from Years of Open Air Traffic Platforms	Addressing Class Imbalance, Applying Text Coding, Normalizing And Expanding Digital Data.	The Conclusion indicated that the designed system succeeded in predicting the three levels with high accuracy, as the ANN was the best at capturing very high periods (High Demand)	Overlap between medium and low levels of demand and solved by setting the classification thresholds (Classification Thresholds)
(From Elsevier - Transportation Research Part C: Emerging Technologies) Full Reference (APA): Chen, H., Liu, Y., & Zhang, Q. (2024). Predicting flight demand fluctuations during peak travel seasons using gradient boosting trees. Transportation Research Part C: Emerging Technologies, 158, 104410.	Predict very large flies in ordering flights during travel and peak seasons using gradually enhanced tree algorithm	Huge public data for historical flight bookings and movements from open academic data platforms	Deleting anomal values and doing on hot ending for destinations and airport names and generating derivative features that link time periods to long vacation seasons	The study reached the superiority of the performance of the XG Post model clearly over the rest of the algorithms, achieving the highest balance accuracy (F1Score) in the classification of high-intensity trips	The very large size of a data set that requires a large memory, solved by the application of partial data flow processing

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
Elsevier - Journal of Air Transport Management) Full Reference (APA): Wang, J., & Li, X. (2023). Air travel passenger demand forecasting using ensemble machine learning: A Random Forest and XGBoost approach. Journal of Air Transport Management, 109, 102385.	To classify the forecasting of demand levels for flights and determining peak periods based on temporal and seasonal variables	Open source historical data for flight reservations from the Kagel platform containing more than 50,000 records of flights including dates, destinations and take-off times	Cleaning data from lost values and applying one hot ending for text variables such as airport and destination names and feature engineering work to link dates to official holidays and weekends	The conclusion showed that the XGBoost algorithm excelled with an accuracy of 92.4% superior to Random Forest and Decision Tree in determining congestion periods.	Not balancing data between normal days and peak days, solved using data balancing techniques

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
(From IEEE Transactions on Big Data) Full Reference (APA): Patel, R., & Jones, L. (2023). Exploratory data analysis and feature engineering for air travel congestion prediction models. IEEE Transactions on Big Data, 9(5), 1102-1115.	To study the impact of exploratory data analysis (EDA) and feature engineering on raising the accuracy of machine learning algorithms in predicting overcreession and flight demand	Large data from the US Bureau of Transportation Statistics (BTS) including flight schedules and occupancy rates	High focus on feature engineering to extract complex time features and apply Log Transformation to modify the distribution of digital data	The research summarized that the inclusion of carefully designed time features raised the accuracy of the Random Forest and Decision Tree algorithm by 18% in the peak period classification.	The presence of noise (Noise) and large distracted data in open data sets, solved through statistical filtering
(Journal of Transportation Forecasting) Full Reference (APA): Brown, T., & Green, P. (2023). The impact of public holidays and weekly seasonal patterns on airline fleet management: A predictive analytics approach. Journal of Transportation Forecasting, 41, 112-126.	To study the impact of public holidays and weekly patterns on flight planning and predicting periods of major congestion	A general dataset that combines historical flight schedules and the annual holiday calendar	Integrating holiday data with flight data, encoding airports, and extracting the features of the days immediately before and following the holidays (Buffer Days)	The result shows that the XG Post algorithm excelled in linking holiday features to high demand levels, which enhanced the accuracy of decision-making during peak periods	The difference in holiday dates between countries and destinations, solved by including geographical indicators for each trip

(References)	(Goal)	(Dataset)	(Preprocessing)	Results	(Challenges)
Springer (Journal of Intelligent Transportation Systems) Full Reference (APA): Al-Qahtani, S., & Wilson, B. (2023). A machine learning framework for airport congestion classification based on temporal holiday patterns. Journal of Intelligent Transportation Systems, 17(3), 201-214.	Developing a framework for mechanized education to classify congestion and congestion in airports based on time patterns and seasons	Historical Air Transport Data from the U.S. Bureau of Transportation Statistics (BTS)	Cleaning lost values, engineering accurate time features, and converting text variables into encoding using Hot Encoding	The Random Forest algorithm achieved a distinguished and stable performance in classifying congestion levels based on temporal variables and recommended the study as a first option for systems that need an explanation of decisions	The algorithm's sensitivity of lost data in some months, solved by processing through smart data compensation techniques

2.3 Background Section

2.3.1 Problem Domain: Air Transport Capacity and Peak Period Management

The global air transport industry serves as an essential framework for international logistics, trade, and corporate mobility. Due to the high financial costs associated with building and maintaining airport facilities, national civil aviation authorities must continuously optimize their current physical infrastructures. A critical issue in this operational environment is managing "peak travel periods" (periods of severe congestion). Peak periods occur when passenger traffic volumes sharply increase within narrow, compressed timeframes, creating an imbalance between the incoming passenger influx and the static operational capacity of the facility.

Traffic spikes in aviation are heavily influenced by systemic consumer behaviors tied to chronological milestones, such as weekly patterns, school vacation schedules, global tourism seasons, and major moving holidays. When traffic peaks occur unexpectedly or go unmanaged, the immediate result is an acute strain on terminal components, including check-in counters, security screening checkpoints, customs areas, and ground handling teams. This strain leads to extended passenger queue waiting times, localized airport crowding, higher labor costs, and reduced overall service quality. Therefore, analyzing air traffic trends, seasonal fluctuations, and operational bottlenecks is a fundamental priority for airline fleet planners and airport managers who seek to proactively distribute staff, schedule aircraft ground operations, and ensure safe passenger flow dynamics.

Technical Aspects and Algorithm Descriptions 2.3.2

Extreme Gradient Enhancement Algorithm (XGBoost): A highly advanced and improved application of gradient-based decision trees, designed for speed and high efficiency. It builds individual tree models sequentially and not independently, which means that each new tree is specially trained to reduce the remaining errors of all previous trees. It is highly valued in data analysis for its original ability to handle scattered tabular matrices and lost values

Random Forest: A group learning method that creates a large group of independent and unrelated decision trees during the training phase. In classification tasks, the model in its final outcome is based on the majority (voting) option of individual trees. The voting mechanism gives this algorithm excellent stability and immunity against the common overfitting problem

Artificial neural networks (ANN): computer processing frameworks inspired by the biological structure of vital neural pathways in the human brain. It consists of interconnected layers of nodes: an input layer, one or more layers hidden for processing, and an output layer. Through backpropagation, networks adjust internal weights adaptively to capture complex, nonlinear, and integrated mathematical functions within multidimensional data sets

Temporal Preprocessing and Encodings: Procedures used to format category texts and crude time stamps into data types suitable for computer processes. Technologies include One-Hot Encoding, which converts separate category inputs (such as airfield destination codes) into clear binary vectors (0 or 1), and Min-Max Scaling, which balances numeric ranges into a unified scale (usually between 0 and 1) to ensure that all variables contribute equally to reducing the error rate in the model

Feature Engineering: It is the process of converting raw data features into information variables that are better the fundamental problem for prediction models. In the field of flight request, it mainly involves extracting explicit time features (e.g., weekday, month, and quarters) from rough timestamps, building holiday proximity indicators (Buffer Days), and linking complex dynamic calendars (such as religious occasions and national moving holidays) to enable models to capture rhythmic human travel behaviors over the years

Hyperparameter Tuning: Systematic optimization of the algorithm's external configuration variables—such as XGBoost's tree depth or the number of layers in the artificial neural network (ANN)—that the model cannot automatically learn during standard training. The application of techniques such as grid search ensures that models are achieved at the highest levels of accuracy and operational stability under different data volumes

Feature Selection: Algorithmic extraction of the most important and statistically useful features from the dense data matrix, excluding ineffective or duplicate variables. This process directly contributes to overcoming the "dimensional curse" resulting from broad class coding, protecting system memory, preventing over-customization (Overfitting), and speeding up overall execution time

Data normalization: A preprocessing technique used to modify the values of digital features in the data set to a unified standard range (often between 0 and 1 or -1 and 1), without changing the differences in the original ranges of the data. This process helps speed up the process of model training, and prevents algorithms (especially artificial neural networks) from being biased with large-numbered variables at the expense of small-numbered variables, ensuring a stable and accurate mathematical affinity

One-Hot Encoding: A pre-processing technique used to convert category and text variables (such as airport, destination, or airline names) into binary digital columns (0 or 1). This is done by creating an independent column for each unique category, so that the column takes the value (1) if the journey belongs to this category, and the value (0) if it does not belong to it, enabling machine learning algorithms and neural networks to process and understand textual data correctly

.Adam OptimizerAdam Optimizer or Adam Optimization Algorithm

It is an optimization algorithm used to update the weights of the neural network during training in order to reduce error and improve model performance. It is characterized by speed and efficiency and is one of the most used enhancers in deep learning applications

References

- Ahmed, T., & Garcia, M. (2024). Deep learning vs. traditional machine learning classifiers for multi-class aviation passenger demand forecasting. *Annals of Data Science*, 11(2), 315-332.
- Al-Qahtani, S., & Wilson, B. (2023). A machine learning framework for airport congestion classification based on temporal holiday patterns. *Journal of Intelligent Transportation Systems*, 17(3), 201-214.
- Al-Sultan, A., Al-Khalidy, M., & El-Gayar, O. (2024). Forecasting peak travel periods in regional hub airports: A comparative machine learning study. *Transportmetrica A: Transport Science*, 20(4), 1845-1868.
- Brown, T., & Green, P. (2023). The impact of public holidays and weekly seasonal patterns on airline fleet management: A predictive analytics approach. *Journal of Transportation Forecasting*, 41, 112-126.
- Chen, H., Liu, Y., & Zhang, Q. (2024). Predicting flight demand fluctuations during peak travel seasons using gradient boosting trees. *Transportation Research Part C: Emerging Technologies*, 158, 104410.
- Choi, J., & Lee, H. (2024). Predicting air travel demand levels (Low, Medium, High) using hybrid deep learning and tree classifiers. *Applied Sciences*, 14(2), 743.
- Fernandez, C., & Davis, R. (2025). Evaluating deep neural networks and XGBoost for predicting aviation demand levels from open-source datasets. *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 422-431.
- Kumar, A., & Sharma, P. (2025). Predictive analytics in aviation: Implementation of decision trees and ANN for operational efficiency. *Operational Research*, 25(1), 89-104.
- Liu, Y., & Gupta, A. (2025). Data-driven resource management in aviation systems: Utilizing Artificial Neural Networks for passenger demand projections. *International Journal of Production Research*, 63(2), 510-527.
- Müller, K., & Fischer, T. (2024). Predicting airport peak periods and passenger flow density using Artificial Neural Networks. *Computers & Industrial Engineering*, 191, 110142.

Patel, R., & Jones, L. (2023). Exploratory data analysis and feature engineering for air travel congestion prediction models. *IEEE Transactions on Big Data*, 9(5), 1102-1115.

Rodriguez, J., & Kim, S. (2024). A hybrid intelligent framework for air traffic peak period forecasting and resource allocation using deep learning. *IEEE Access*, 12, 45112-45125.

Silva, F., & Santos, T. (2025). Benchmarking tree-based machine learning models and deep learning for multiclass flight demand classification. *Expert Systems with Applications*, 240, 122410.

Taylor, S., & Novak, J. (2026). Hyperparameter Tuning of Gradient Boosted Trees for Real-Time Flight Demand Prediction Systems. *Cluster Computing*, 29(1), 54-68.

Wang, J., & Li, X. (2023). Air travel passenger demand forecasting using ensemble machine learning: A Random Forest and XGBoost approach. *Journal of Air Transport Management*, 109, 102385.

Zhang, Y., & Al-Suwaidan, M. (2024). An intelligent flight demand prediction system utilizing temporal and seasonal feature engineering. *IEEE Transactions on Intelligent Transportation Systems*, 25(3), 1420-1432.