Intelligent Methods in Engineering Sciences 3(3): 082-090, 2024

INTELLIGENT METHODS
IN ENGINEERING SCIENCES

https://www.imiens.org

International Open Access

September, 2024

e-ISSN 2979-9236

Research Article

https://doi.org/10.58190/imiens.2024.102

# Performance Evaluation of Machine Learning Algorithms in Estimating Taxi Times at Istanbul Airport

***Erkan ÇÖMEZ*** [a], ***Onur INAN*** [b,*]

[a] *Graduate School of Natural and Applied Sciences, Selcuk University, Konya, Türkiye*
[b] *Department of Computer Engineering, Selcuk University, Konya, Türkiye*

ARTICLE INFO

*Article history:*
Received 15 July 2024
Accepted 11 August 2024

*Keywords:*
Istanbul Airport,
Machine Learning,
Regression Algorithms,
Taxi Time Prediction

ABSTRACT

This article evaluates the performance of regression algorithms used to estimate taxi-out times at Istanbul Airport. Artificial neural networks, random forest, gradient boosting, and decision trees algorithms were studied to determine the algorithms with the highest accuracy. Principal Component Analysis (PCA) was used to reduce the data's dimensionality and improve model performance. The findings of the study provide valuable insights for more effective management of airport operations and reduction of flight delays. PCA-applied Artificial Neural Networks (ANN) emerged as the most successful algorithm, demonstrating the highest accuracy ($R^2$: 95.89%) and lowest error margins (MAE: 0.016, MSE: 0.001) in predicting taxi-out times. This superior performance indicates that ANN can effectively capture the complex relationships and variability inherent in airport operational data. Following ANN, the PCA-applied Random Forest algorithm also showed commendable accuracy ($R^2$: 94.89%), providing robust predictions with slightly higher error margins (MAE: 0.157, MSE: 0.044) compared to ANN. These results underline the potential of using advanced machine learning techniques to enhance the efficiency of airport operations, thereby minimizing delays and optimizing resource allocation. Overall, the application of these machine learning models, particularly ANN and Random Forest, offers a significant improvement over traditional methods. The study's outcomes suggest that incorporating these advanced algorithms can lead to more accurate predictions of taxi-out times, supporting better decision-making processes and operational strategies at airports.



## 1. Introduction

The continuous development of the civil aviation transportation industry has necessitated the enhancement of operational efficiency and service capacity at airports. In particular, airport ground services and air traffic management must develop more effective planning and management strategies to cope with increasing aircraft traffic. Given the annual increase in the number of aircraft takeoffs and landings, this situation intensifies the pressure on airports and creates operational challenges.

The low rates of on-time departures can be attributed to factors such as low efficiency in airport ground services technical support and limited airspace capacity. This situation has become a significant bottleneck, restricting the development of civil aviation. The congestion experienced at airports, complex runway configurations, and the integration of information from airlines, airports, and air traffic control departments make it challenging to accurately predict taxi times and, consequently, departure timings. Taxi time is a critical indicator for evaluating the efficiency of airport ground operations support, and its accuracy directly affects the sequencing of aircraft takeoffs and landings.

The current Airport Collaborative Decision Making (A-CDM) system determines the estimated taxi time for all aircraft using only the average airport taxi time. This system does not take into account significant factors such as boarding gates, runway configuration, the number of taxiing aircraft, and weather conditions. These shortcomings reduce the accuracy of taxi and departure time predictions necessary for ensuring on-time departures and optimizing airport operations, leading to flight delays and increased fuel consumption costs. Effective ground movement operations are fundamental to the successful functioning of air transportation networks [1]. Therefore, accurately predicting an aircraft's taxi time is critical for optimizing the arrangement of boarding gates, increasing the efficiency of departure time slots, and enabling airlines to accurately calculate fuel amounts. This also enhances the potential to reduce airport ground emissions, thereby supporting environmental sustainability.

Multiple stakeholders have recommended the necessity

*** Corresponding Author:** oinan@selcuk.edu.tr

of a modernized Air Traffic Control system to reduce congestion and increase capacities at Istanbul Airport. The taxi-out times at Istanbul Airport, which range from as short as 10 minutes to as long as 140 minutes throughout the day, make the taxi-out time prediction problem significantly more challenging compared to other more regulated airports in Turkey.

The application of artificial intelligence in predicting taxi-out times is crucial for enhancing the accuracy and efficiency of airport operations. By leveraging machine learning algorithms, such as artificial neural networks, random forests, gradient boosting, and decision trees, this approach allows for the consideration of numerous variables and complex relationships that traditional methods may overlook. The improved accuracy in predicting taxi times leads to better resource management, reduced delays, optimized fuel consumption, and a significant decrease in airport ground emissions, thereby supporting both operational efficiency and environmental sustainability.

This article evaluates the performance of regression algorithms used to estimate airport taxi-out times (artificial intelligence-supported solutions have been developed). The study is structured as follows:

In the Introduction section, a general overview of the operational efficiency requirements of the civil aviation transportation industry and the importance of airport taxi times is provided. The Literature Review section examines existing research and methods, discussing the advantages and disadvantages of different approaches. The Istanbul Airport section provides detailed information on the characteristics and operational challenges of Istanbul Airport, which was selected as the study area. The Regression Approaches and Techniques for Predicting Taxi Time section explains various regression models in detail, including Artificial Neural Networks, Decision Trees, Random Forests, and Gradient Boosting, along with detailed information on Principal Component Analysis (PCA). The Evaluation Methodologies for Prediction Models section presents the performance metrics and evaluation methods used. The Dataset and Algorithm Schema section explains the process of processing the dataset used to predict taxi-out times at Istanbul Airport and modeling it with various regression algorithms. In the Findings and Discussion section, the performances of the regression models are compared, and the most effective model and the obtained results are discussed. Finally, the Conclusion section summarizes the study's overall findings and provides recommendations for future research.

## 2. Literature Review

Various methods have been developed to estimate taxi-out times. These methods encompass models based on historical data predictions and causal factors such as queuing theory. Shumsky applied dynamic linear models to predict taxi-out times based on aircraft traffic flow and departure demands. By comparing static and dynamic models, it was determined that the dynamic model performed better in predicting taxi-out times for short time intervals. These findings can be considered a significant advancement in the more accurate prediction of taxi-out times [2].

To enhance airport operational efficiencies through advanced predictive methodologies, Balakrishna et al. present a significant approach using a reinforcement learning algorithm specifically to predict taxi-out times at John F. Kennedy International Airport (JFK). The importance of this study lies in the fact that taxi-out delays constitute a substantial portion of overall flight delays, addressing both the economic and operational impacts of these delays. The authors developed a learning-based solution that dynamically adapts to the variability of airport conditions [3].

Srivastava addresses the challenges of predicting taxi times in aircraft departure processes. This research highlights the complexities due to various factors such as airport traffic, runway configurations, and weather conditions. Using high-resolution ASDE-X surveillance data, the study proposes an adaptive model that significantly enhances the accuracy of taxi time predictions by analyzing historical traffic flow data and correlating it with real-time situational variables. The findings have demonstrated substantial improvements in prediction accuracy compared to traditional methods, proving especially effective in managing airport efficiencies and reducing delays at major airports like JFK during peak times. This model offers a sophisticated tool for airports to optimize ground operations and improve overall traffic management [4].

Jordan et al. address the modeling of airport taxi processes during the development of the Tower Flight Data Manager (TFDM) system, which aims to replace existing systems in air traffic control towers with an integrated technology package. Using a statistical learning approach, relatively simple and easily interpretable models were developed to model aircraft taxi times. These models demonstrated remarkable accuracy when tested on real data. The article provides a detailed explanation of the taxi time models developed at Dallas/Fort Worth International Airport and highlights the potential benefits of TFDM [5].

Ravizza et al. combined a statistical approach with a ground movement model to improve the prediction of taxi times at airports. The research was conducted specifically at two major hub airports in Europe: Stockholm-Arlanda and Zurich Airport. The study utilized multiple linear regression techniques to more accurately predict variations in taxi times and incorporated information related to ground movement modeling [6].

Lee et al. explored the use of machine learning techniques and fast-time simulation tools to predict taxi times and departure timings at Charlotte Douglas International Airport. The fast-time simulation tool LINOS and various machine learning models were evaluated using real air traffic data, and the performance of these methods in predicting taxi times was compared. The analyses revealed that LINOS demonstrated similar success in predicting taxi times as the machine learning methods. Additionally, the study focused on the applicability of LINOS in real-time airport operations and the challenges encountered during the adaptation process. Support Vector Regression (SVM) and Linear Regression (LR) methods were noted to perform better than the Dead Reckoning method in terms of prediction performance. Moreover, k-Nearest Neighbors (kNN) and Random Forest (RF) methods were reported to yield better results than all other prediction methods [7].

Idris et al. analyzed various factors affecting taxi-out time using Airline Service Quality Performance (ASQP) data. These factors include runway configuration, airline/terminal, flow direction restrictions, and departure queue size [8-10].

Carr et al. developed a queuing model to predict taxi-out time and concluded that the departure queue size, measured as the number of departures between an aircraft's pushback time and its takeoff time, showed the best correlation with taxi-out time. They proposed a simulation-based study examining queue dynamics and traffic rules. By considering aggregate metrics such as airport capacity and departure density, they predicted taxi-out time [11].

Simaiakis and Balakrishnan proposed a taxi-out time prediction model that includes the estimation of the distributions of unimpeded taxi-out times and the development of a queuing model for the departure runway system, forming an analytical model of the aircraft departure process [12]. Hebert and Dietz developed a multi-stage Markov process model to predict the departure process at LaGuardia Airport, based on five days of data [13].

Lordan et al. proposed a model to predict taxi times for Barcelona-El Prat Airport. This model uses log-linear regression analysis with variables that can be known before operations to predict taxi times, contributing to more efficient management of airport operations [14].

Chen et al. proposed a model for predicting aircraft taxi times and quantifying the uncertainties associated with these times using multi-objective fuzzy rule-based systems to better manage the uncertainties encountered in air traffic management (ATM). The study aims to mitigate the effects of uncertainties arising from factors such as variable weather conditions, operational scenarios, and pilot behaviors. Based on historical aircraft taxi data, the research offers a new approach that more informatively captures these uncertainties [15].

Diana conducted a comparison of different machine learning models for predicting taxi-out times at Seattle/Tacoma International Airport. The study evaluates the performance of Ensemble learning, Ordinary Least Squares (OLS), and penalized regression algorithms across two different periods in which NextGen capabilities were implemented. During the pre-sample period, the OLS and ridge models outperformed other ensemble learning models, while in the post-sample period, the gradient boosting model provided the lowest root mean square error. The study suggests that there is no single algorithm that provides the best fit in all cases and recommends selecting the most well-balanced model [16].

In a study conducted by Yin et al., a taxi-out time prediction model based on a large network topology was presented using machine learning techniques. The study was built on historical data analyses conducted at Shanghai Pudong International Airport, examining the factors affecting taxi-out times and their relationships. The research comprehensively evaluated various machine learning methods, including linear regression, support vector machines, and random forests, and compared the training performances of these models. The results showed that the random forest model, trained over a period of one month, significantly outperformed other models in terms of prediction accuracy [17].

## 3. Istanbul Airport

Istanbul Airport stands out as Turkey's largest and most modern airport. Located on the European side of Istanbul, between the villages of Tayakadın and Akpınar in the Arnavutköy district, it officially opened on October 29, 2018. Upon reaching full capacity, it is expected to be among the largest airports in the world with an annual capacity of 200 million passengers.

### *3.1. General Overview*

Istanbul Airport is built on a massive area of 76.5 million square meters. Initially, the airport had a capacity of 90 million passengers, and once all phases are completed, it will reach a capacity to handle 200 million passengers annually. Currently, three independent runways are operational, allowing an average of 700 aircraft to take off daily. Once all six planned runways are in operation, Istanbul Airport will be one of the few airports globally with such a capacity.

### *3.2. Terminal Features*

The main terminal building, covering an area of 1.3 million square meters, is one of the largest terminal buildings in the world under one roof. It features a total of 189 gates, with 151 equipped with passenger bridges and 38 served by buses. Additionally, the airport offers 375 check-in counters and numerous self check-in kiosks, providing passengers with a fast and comfortable experience. The 42-kilometer long automated baggage handling system allows for highly efficient baggage processing.

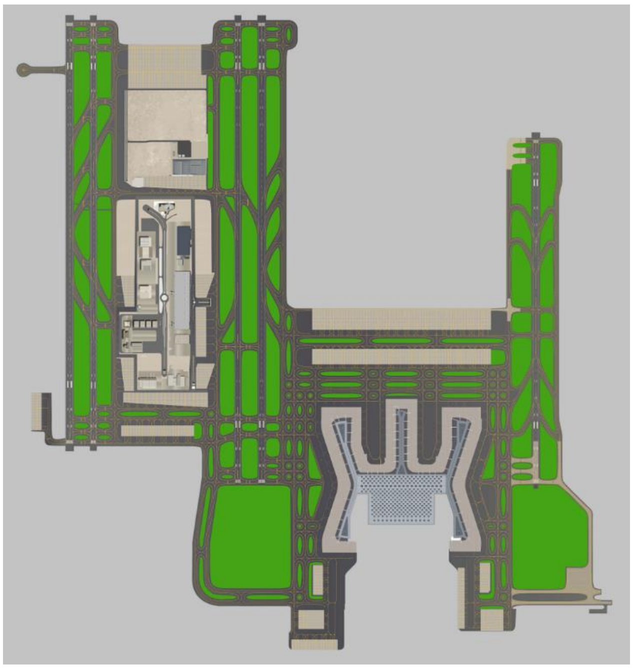

**Figure 1.** Istanbul Airport Map

## 4. Regression Approaches and Techniques for Predicting Taxi Time

### *4.1. Neural Network-Based Regression Models*

Nowadays, learning using artificial neural networks (ANNs) or deep learning has become a dominant approach in machine learning [18]. Neural network-based regression models hold a significant place among advanced artificial intelligence systems, especially in the prediction of continuous values, and have broad applications. These models process multiple variables in the input data set to produce a continuous output value based on these variables. They can be used in various fields, such as real estate valuation and vehicle speed prediction.

The foundation of the model consists of layered artificial neural networks. Each layer processes the information received from the previous layer with a specific activation function and passes it on to the next layer. This process continues until the final output layer. During the model's training, the difference between the predictions made using real-world data and the actual values is minimized. This process is typically performed using the backpropagation algorithm, which iteratively updates the model's weights to minimize the error rate.

Neural network regression models are used to model complex relationships in data and make accurate predictions on new, similar data. These models are preferred in many fields due to their high adaptability and prediction accuracy.

### *4.2. Decision Tree Regression Model*

Decision tree regression models are an effective machine learning method used to produce continuous value outputs. A decision tree is widely used in machine learning and data mining [19]. This method analyzes the relationships between features in the data set and constructs a tree structure. Each node represents a feature in the data set, and decisions are made based on these features. These decisions create branches according to a determined threshold value, and this process continues until the end of the model. The leaf nodes represent the outcomes predicted by the model. The regression decision tree is particularly successful in explaining variations in the data by partitioning and representing subsets of the data.

During the model training process, the data set is split by selecting the best feature and threshold value. This splitting process typically aims to minimize the sum of squared errors. Parameters such as the depth of the decision tree and the number of branches directly affect the complexity and generalization ability of the model. Deep trees can lead to overfitting, while very shallow trees can cause underfitting. Therefore, techniques like cross-validation are used to improve the model's accuracy and generalization capacity.

Decision tree regression models are frequently preferred in science and industry not only for their strong predictive capabilities but also for the ease of interpreting the model's results. By uncovering hidden structures in data sets, these models play a crucial role in complex decision-making processes.

### *4.3. Random Forest Regression Model*

Random Forest is a powerful ensemble-learning method proposed by Breiman that can be used for both classification and regression tasks [20]. This model consists of a large number of decision trees, and the predictions from each tree are combined to obtain an overall prediction. This aggregation process makes the model stronger and more accurate.

The core principle of Random Forest is to train each decision tree independently using slightly different subsets of the data. This approach allows each tree to grow independently, enabling the model to generalize better overall. Additionally, the trees are split using randomly selected features, which increases diversity and contributes to the robustness of the model.

Random Forest models offer high accuracy and are less prone to overfitting compared to a single decision tree, making them ideal for applications with complex data structures. Furthermore, they provide powerful tools for determining which features have the most impact on predictions, making them valuable for variable importance and feature selection tasks. Random Forest performs particularly well with large data sets and complex problem sets. Therefore, the flexibility and high accuracy of the model make it a preferred choice for many researchers and practitioners.

### *4.4. Gradient Boosting Regression Model*

The Gradient Boosting [21], offers high accuracy in solving complex regression and classification problems by improving weak predictors. This method works through an iterative process, adding new models in such a way that they reduce the errors made in previous steps. In each iteration, the model focuses specifically on the erroneous predictions and adjusts the weights to correct these errors, thus continually improving the model with each step.

One of the main advantages of this model is its high accuracy rate and resistance to overfitting. Gradient Boosting is particularly robust against noise and outliers in the data, making it ideal for modeling the complexities inherent in data sets. The Gradient Boosting Regression Model has a wide range of applications, leading to its frequent use in both theoretical research and practical applications.

This resilience and adaptability make Gradient Boosting a preferred method for many researchers and practitioners dealing with intricate data structures and requiring precise predictions.

### *4.5. Principal Component Analysis (PCA)*

Principal Component Analysis (PCA) is frequently used as a statistical method for data analysis, allowing the reduction of data set sizes with numerous interrelated features by expressing the data with fewer variables [22]. By reducing the number of variables in the original data set, PCA aims to represent the data in a simpler way. This process creates a new set of variables that best explain the variance in the data. PCA enables data analysis with fewer components by eliminating correlations in high-dimensional data. This method is particularly useful in the following situations:

When there are a large number of variables in the data set

When there is high correlation among variables

When there is a need to reduce computational costs

The main steps of PCA are:

Standardization of Data: Data is standardized to enable comparison between variables.

Calculation of the Covariance Matrix: The variance and correlation of the data are determined.

Calculation of Eigenvalues and Eigenvectors: The eigenvalues and eigenvectors of the covariance matrix are calculated to create new components that best represent the data.

Selection of Components: Components that explain most of the variance in the data are selected.

Creation of a New Data Set: A new, lower-dimensional data set is created using the selected components.

PCA can be used as a preprocessing step in machine learning algorithms such as Neural Networks (ANN), Gradient Boosting (GB), Decision Trees (DT), and Random Forest (RF), enhancing model performance and reducing computation time. PCA is particularly helpful in high-dimensional data sets, facilitating faster and more effective model operation.

## 5. Evaluation Methodologies for Prediction Models

In our study, the performance metrics used to evaluate the regression models are the coefficient of determination ($R^2$), mean absolute error (MAE), and mean squared error (MSE). The $R^2$ coefficient is a statistical measure that indicates the proportion of the variance in the dependent variable that is predictable from the independent variables, as defined by Equation (1). This metric determines how well the model's predictions match the actual values; values close to 1 indicate that the model makes predictions with a high degree of accuracy. These metrics objectively assess how well the algorithms perform on the training data and how close their predictions on the test data are to the actual values.

$$\boldsymbol{r^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}} \tag{1}$$

There are numerous methods for measuring error; however, mean square error (MSE) is the most commonly favored [23]. The mean squared error (MSE) calculates the average of the squares of the differences between the predicted values and the actual results (Equation (2)). A value approaching 0 indicates that the model's error is at a minimum level, thereby signifying the accuracy of the predictions.

$$\boldsymbol{\frac{1}{n}\sum_{j=1}^{n} \left(y_j - \hat{y}_i\right)^2} \tag{2}$$

In evaluating model performance, additional metrics such as Mean Absolute Error (MAE) and processing time were also considered. MAE calculates the average of the absolute differences between predicted and actual values (Equation (3)). This metric is useful for understanding the magnitude of prediction errors.

$$\boldsymbol{\frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|} \tag{3}$$

In the research, the performance of the regression models was evaluated using metrics such as the coefficient of determination ($R^2$), mean absolute error (MAE), mean squared error (MSE), and processing time (in miliseconds). These values are ranked from highest to lowest and presented in Table 1.

## 6. Dataset and Algorithm Schema

The dataset used to predict the taxi times and fuel consumption of aircraft at Istanbul Airport includes various operational and meteorological parameters. A unique identifier for each flight, the flight number, is used to track and analyze flights, with a minimum value of 1 and a maximum value of 9909. The stand ID, which is the identifier of the parking location of the aircraft, is essential for understanding the ground movements and taxi times of the aircraft, with a minimum value of 1 and a maximum value of 432. The runway ID, which is the identifier of the runway used for landing or takeoff, can have different lengths and configurations that affect taxi times, with a minimum value of 1 and a maximum value of 11.

The distance from the parking location to the runway, known as the runway exit distance, is a critical factor that directly affects taxi time, with a minimum value of 490 meters and a maximum value of 11514 meters. Wind speed, encountered during landing or takeoff, has a significant impact on flight safety and performance, with a minimum value of 0 and a maximum value of 30. Wind direction indicates the direction from which the wind is coming and can affect the landing and takeoff routes of aircraft, with a minimum value of 0 degrees and a maximum value of 360 degrees. Temperature is the temperature value of the day, which can affect engine performance and fuel consumption, with a minimum value of 0 degrees and a maximum value of 29 degrees.

Cloud amount is a measure used to determine weather conditions and can affect flight operations and visibility, with a minimum value of 1 and a maximum value of 4. Cloud base height, which indicates the height of the cloud base from the ground, plays an important role in flight operations and landing-takeoff processes, with a minimum value of 200 meters and a maximum value of 40000 meters. CAVOK (Ceiling and Visibility OK) indicates whether a certain meteorological condition is present, specifying that certain criteria for visibility and cloud base height are met, with a minimum value of 0 and a maximum value of 1.

The METAR code, which is the routine aviation weather report for a specific airport, summarizes the current weather conditions at the airport, with a minimum value of 5 and a maximum value of 14. Visibility is critical for flight safety, with a minimum value of 3500 meters and a maximum value of 9999 meters. The departure time, indicating the time of landing or takeoff of the aircraft, is a category that expresses the time period of the flight during the day and is used to evaluate operational density, with a minimum value of 0 and a maximum value of 7. Taxi time, the time taken for the aircraft to reach the parking stand after landing or to reach the runway from the parking stand for takeoff, is the key metric optimized in this study, with a minimum value of 60 seconds and a maximum value of 2340 seconds.

This dataset allows for a detailed analysis of flight operations at Istanbul Airport. The variables in the dataset comprehensively cover the operational and environmental factors that affect the taxi times and fuel consumption of aircraft.

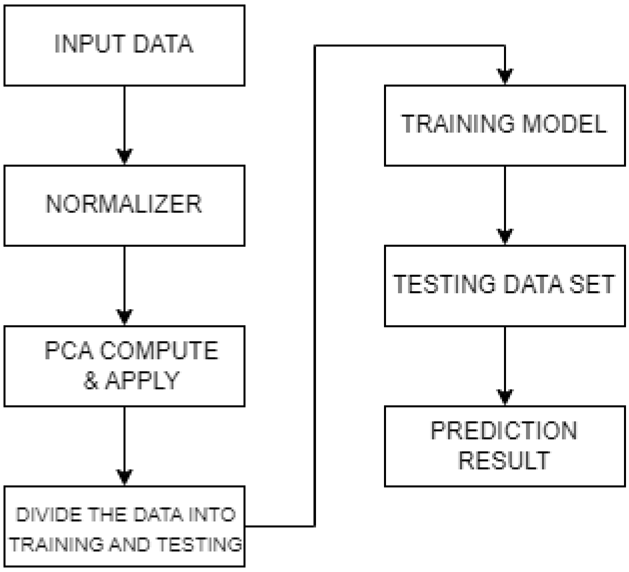


**Figure 2.** Flow Diagram

The dataset contains various operational and meteorological parameters used to predict the taxi times and fuel consumption of aircraft at Istanbul Airport. Figure 2 shows the steps of the machine learning process. In the first step, the dataset is loaded into the system, which includes various parameters such as flight number, stand ID, runway ID, runway exit distance, wind speed, wind direction, temperature, cloud amount, cloud base height, CAVOK, METAR code, visibility, departure time, and taxi time.

After the data is loaded into the system, normalization is performed. This step adjusts the numerical values in the dataset to a common scale and ensures that features with different value ranges contribute equally to the model. Next, Principal Component Analysis (PCA) is applied to transform the dataset into a different space, capturing the most significant variance in the data and making the model more efficient and faster.

After PCA is applied, the dataset is divided into training (80%) and testing (20%) sets. This split is used to evaluate the model's performance. The normalized and PCA-transformed dataset is then used to train the regression algorithms. The algorithms used in this study include Artificial Neural Networks (ANN), Random Forest, Gradient Boosting, and Decision Tree. For the Artificial Neural Network (ANN) model, a maximum of 100 iterations was selected, with 2 hidden layers and 10 neurons per layer. The Random Forest model was configured with 100 trees. For the Gradient Boosted model, 100 trees were used with a tree depth of 4 and a learning rate of 0.1.

Once the model is trained, its performance is evaluated on the test dataset. This step measures how well the model generalizes to unseen data. Finally, the model makes predictions on the test dataset, and these predictions are evaluated to determine the model's accuracy and effectiveness. This process ensures that the model is thoroughly trained, tested, and validated, achieving high accuracy in predicting taxi times and fuel consumption.

## 7. Findings and Discussion

In our study, the performance metrics of various regression algorithms were evaluated and ranked in Table 1 from highest to lowest in terms of the coefficient of determination ($R^2$), mean absolute error (MAE), mean squared error (MSE), and execution time (in milliseconds). Each model was run 10 times, and the average was taken. These metrics allow us to compare the predictive accuracy and computational efficiency of the algorithms.

**Table 1.** Performance Metrics of Regression Algorithms with PCA

| Regression Algorithm | $R^2$ | MAE | MSE | TIME (ms) |
|---|---|---|---|---|
| Artificial Neural Networks | 95,89 | 0,016 | 0,001 | 5590 |
| Random Forest | 94,89 | 0,157 | 0,044 | 5351 |
| Gradient Boosting | 91,54 | 0,205 | 0,08 | 4992 |
| Decision Tree | 91,01 | 0,165 | 0,087 | 1644 |

**Table 2.** Performance Metrics of Regression Algorithms without PCA

| Regression Algorithm | $R^2$ | MAE | MSE | TIME (ms) |
|---|---|---|---|---|
| Artificial Neural Networks | 71,7 | 0,06 | 0,007 | 3717 |
| Random Forest | 86,8 | 0,26 | 0,13 | 4537 |
| Gradient Boosting | 79,3 | 0,33 | 0,20 | 3197 |
| Decision Tree | 81,4 | 0,206 | 0,184 | 1146 |

Tables 1 and 2 present the $R^2$, MAE, MSE, and computation times of various regression algorithms. This comparison is made to evaluate the impact of PCA on model performance.

In the regression models with PCA, Artificial Neural Networks (ANN) achieved the highest $R^2$ value of 95.89%, whereas, in the models without PCA, this value significantly dropped to 71.7%. Similarly, a decline in $R^2$ values was observed for other algorithms as well. When examining the MAE values, models with PCA exhibited lower errors. Specifically, the MAE for Artificial Neural Networks decreased from 0.06 before PCA to 0.016 with PCA. This indicates that PCA can enhance model accuracy. A similar trend is observed in the MSE values, where models with PCA generally showed lower MSE values. For instance, the MSE for the Random Forest algorithm dropped from 0.13 without PCA to 0.044 with PCA, suggesting that PCA can reduce prediction errors.

However, in terms of computation times, PCA negatively impacted model performance. The computation time for Artificial Neural Networks increased from 3717 ms without PCA to 5590 ms with PCA. Likewise, other algorithms also showed increased computation times. This indicates that PCA introduces additional computational overhead, thereby extending the total processing time.

In conclusion, the performance comparison of regression algorithms with and without PCA reveals that PCA can enhance model accuracy and reduce prediction errors. However, PCA also introduces additional computational costs, increasing the overall runtime of the models. The use of PCA can be particularly beneficial in scenarios where model accuracy is critical and longer computation times are acceptable. Therefore, it is recommended to thoroughly analyze model performance when using PCA and consider alternative methods if necessary.

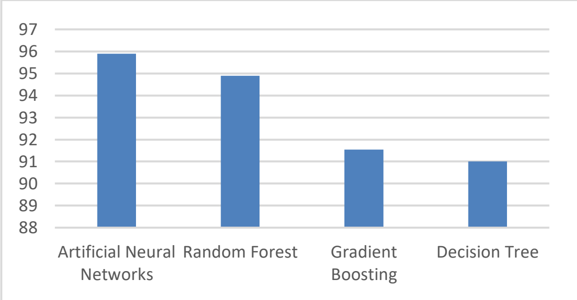


**Figure 3.** Comparison of Coefficients of Determination ($R^2$) of Regression Algorithms

Figure 3 shows a comparative display of the coefficients of determination ($R^2$) of various regression algorithms. An $R^2$ value approaching 1 indicates a high level of model accuracy. According to the analysis, the most successful regression algorithm is identified as the artificial neural networks regression, with the highest $R^2$ value of 95.89%.

Figure 4 presents a comparative display of the mean absolute error (MAE) values of different regression algorithms. An MAE value close to 0 indicates a high level of model accuracy. Based on the analysis results, the algorithm with the lowest MAE value is considered the most successful.

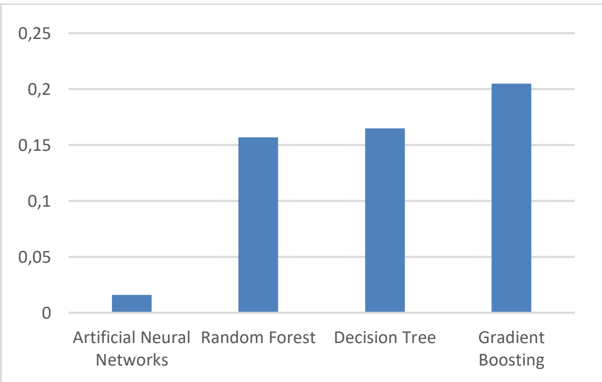


**Figure 4.** Comparison of Mean Absolute Error (MAE) Values of Regression Algorithms

Figure 5 presents a graph comparing the mean squared error (MSE) values of various regression algorithms. An MSE value approaching 0 indicates high algorithm accuracy. According to the obtained results, the algorithm with the lowest MSE value is considered the most successful.

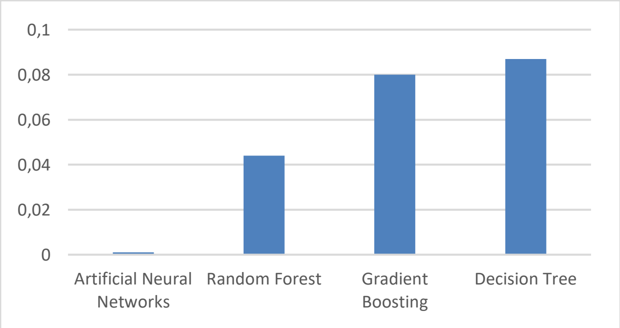


**Figure 5.** Comparison of Mean Squared Error (MSE) Values of Regression Algorithms

Figure 6 presents a graph comparing the computation times of various regression algorithms in milliseconds. All algorithms, except for the artificial neural networks algorithm, produced results relatively quickly.

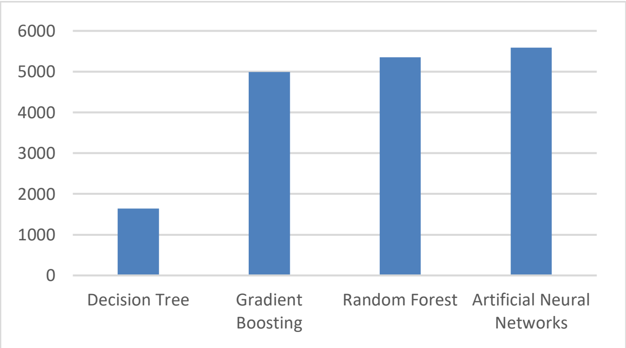


**Figure 6.** Comparison of Time Values of Regression Algorithms

## 8. Conclusion

In this study, airport taxi-out times were predicted using various regression algorithms. Through analyses and comparisons, the performance metrics such as the coefficient of determination ($R^2$), mean absolute error (MAE), mean squared error (MSE), and computation time were evaluated.

Among the regression algorithms compared, Artificial Neural Networks Regression demonstrated the highest performance. Artificial Neural Networks achieved the highest $R^2$ value of 95.89%, indicating that the model's predictions are highly accurate.

In the evaluation of mean absolute error (MAE), Artificial Neural Networks also showed the best performance with the lowest MAE value, indicating that the predicted values are very close to the actual values.

In the mean squared error (MSE) analyses, Artificial Neural Networks again achieved the lowest value, demonstrating that the overall prediction error of the model is minimal. The MSE value represents the average of the squares of the differences between the predicted and actual values, and a lower value indicates higher model accuracy.

Regarding computation times, it was observed that Artificial Neural Networks produced results in longer durations compared to other algorithms. In contrast, algorithms such as Random Forest, Gradient Boosting, and Decision Tree produced faster results.

Overall, Artificial Neural Networks were identified as the most successful algorithm for predicting taxi-out times, with high accuracy but longer computation times. These findings contribute significantly to the more effective management of airport operations and the reduction of flight delays.

The results of this study offer critical insights for the strategic selection and advancement of predictive models aimed at optimizing airport operations. The notable enhancements in prediction accuracy and efficiency underscore the powerful impact of integrating PCA with advanced machine learning algorithms. Future research can build on these findings by incorporating larger and more diverse datasets and exploring innovative algorithms. Such endeavors could dramatically boost predictive performance and operational efficiency, setting new standards for airport management, minimizing delays, optimizing resource use, and paving the way for groundbreaking advancements in the field.

## References

[1] J. a D. Atkin, E.K. Burke, S. Ravizza, The airport ground movement problem: Past and current research and future directions, Research in Air Transportation. (2010) 131–138. http://www.cs.nott.ac.uk/~smr/share/10_ICRAT_Ravizza.pdf.

[2] R.A. Shumsky, Dynamic Statistical Models for the prediction of aircraft take-off times, PhD Thesis. (1995) 1–214. papers2://publication/uuid/95251702-60DD-48BC-AE91-042F332CC19B (erişim 27 Nisan 2024).

[3] P. Balakrishna, R. Ganesan, L. Sherry, B.S. Levy, Estimating Taxi-out times with a Reinforcement Learning algorithm, AIAA/IEEE Digital Avionics Systems Conference - Proceedings. (2008). doi:10.1109/DASC.2008.4702812.

[4] A. Srivastava, Improving departure taxi time predictions using ASDE-X surveillance data, AIAA/IEEE Digital Avionics Systems Conference - Proceedings. (2011) 2B5-1-2B5-14. doi:10.1109/DASC.2011.6095989.

[5] R. Jordan, M.A. Ishutkina, T.G. Reynolds, A statistical learning approach to the modeling of aircraft taxi time, AIAA/IEEE Digital Avionics Systems Conference - Proceedings. (2010). doi:10.1109/DASC.2010.5655532.

[6] S. Ravizza, J.A.D. Atkin, M.H. Maathuis, E.K. Burke, A combined statistical approach and ground movement model for improving taxi time estimations at airports, Journal of the Operational Research Society. 64 (2013) 1347–1360. doi:10.1057/jors.2012.123.

[7] H. Lee, W. Malik, Taxi Time Prediction at Charlotte Airport Using Fast - Time Simulation and Machine Learning Techniques, (2015) 1–11.

[8] H.R. Idris, I. Anagnostakis, B. Delcaire, R.J. Hansman, J.-P. Clarke, E. Feron, A.R. Odoni, Observations of Departure Processes at Logan Airport to Support the Development of Departure Planning Tools, https://doi.org/10.2514/atcq.7.4.229. 7 (2016) 229–257. doi:10.2514/ATCQ.7.4.229.

[9] H. Idris, Observation and analysis of departure operations at Boston Logan International Airport [Ph.D. thesis], MIT, 2001.

[10] H. Idris, J.-P. Clarke, R. Bhuva, L. Kang, Queuing Model for Taxi-Out Time Estimation, https://doi.org/10.2514/atcq.10.1.1. 10 (2016) 1–22. doi:10.2514/ATCQ.10.1.1.
[11] F. Carr, A. Evans, J.P. Clarke, E. Feron, Modeling and control of airport queueing dynamics under severe flow restrictions, Proceedings of the American Control Conference. 2 (2002) 1314–1317. doi:10.1109/ACC.2002.1023202.
[12] I. Simaiakis, H. Balakrishnan, Queuing models of airport departure processes for emissions reduction, AIAA Guidance, Navigation, and Control Conference and Exhibit. (2009). doi:10.2514/6.2009-5650.
[13] J.E. Hebert, D.C. Dietz, Modeling and Analysis of an Airport Departure Process, https://doi.org/10.2514/2.2133. 34 (2012) 43–47. doi:10.2514/2.2133.
[14] O. Lordan, J.M. Sallan, M. Valenzuela-Arroyo, Forecasting of taxi times: The case of Barcelona-El Prat airport, Journal of Air Transport Management. 56 (2016) 118–122. doi:10.1016/J.JAIRTRAMAN.2016.04.015.
[15] J. Chen, M. Weiszer, E. Zareian, M. Mahfouf, O. Obajemu, Multi-objective fuzzy rule-based prediction and uncertainty quantification of aircraft taxi time, IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. 2018-March (2017) 1–5. doi:10.1109/ITSC.2017.8317826.
[16] T. Diana, Can machines learn how to forecast taxi-out time? A comparison of predictive models applied to the case of Seattle/Tacoma International Airport, Transportation Research Part E: Logistics and Transportation Review. 119 (2018) 149–164. doi:10.1016/J.TRE.2018.10.003.
[17] J. Yin, Y. Hu, Y. Ma, Y. Xu, K. Han, D. Chen, Machine learning techniques for taxi-out time prediction with a macroscopic network topology, AIAA/IEEE Digital Avionics Systems Conference - Proceedings. 2018-September (2018). doi:10.1109/DASC.2018.8569664.
[18] LeCun, Y., Bengio, Y., Hinton, G., LeCun, Y., Bengio, Y., & Hinton, G. (2015-05-27). Deep learning. Nature 2015 521:7553, 521(7553). https://doi.org/10.1038/nature14539
[19] Sun, H., & Hu, X. (2017). Attribute selection for decision tree learning with class constraint. Chemometrics and Intelligent Laboratory Systems, 163, 16-23. https://doi.org/10.1016/j.chemolab.2017.02.004
[20] Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/a:1010933404324
[21] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.
[22] Inan, O., Uzer, M. S., & Yılmaz, N. (2013). A new hybrid feature selection method based on association rules and PCA for detection of breast cancer. International Journal of Innovative Computing, Information and Control, 9(2), 727-729.
[23] Karakoyun, M. (2024). Artificial neural network training using a multi selection artificial algae algorithm. Engineering Science and Technology, an International Journal, 53. https://doi.org/10.1016/j.jestch.2024.101684