International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 1, February 2025, pp. 1038~1050
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i1.pp1038-1050

Airport infrastructure and runway precision aids for forecasting flight arrival delays

Hajar Alla, Youssef Balouki
Laboratory of Mathematics, Computer and Engineering Sciences, Department of Mathematics and Computer Science, Faculty of Sciences and Techniques, Hassan First University of Settat, Settat, Morocco

Article Info

Article history:

Received May 25, 2024
Revised Sep 5, 2024
Accepted Oct 1, 2024

Keywords:

Airport infrastructure
Bayesian optimization
Flight delay prediction
Multilayer perceptron
Runway precision aids

ABSTRACT

Recent research has concentrated on using machine learning approaches to forecast flight delays. The majority of prior prediction algorithms were based on simple and standard attributes collected from the database from which the data were pulled. This article is the first attempt to propose novel features linked to airport capacity and infrastructure. The total runways, the total runway intersections, the longest runway length, the shortest runway length, the runway precision rate, the total terminals, and the total gates were all examined. In this paper, we suggest an optimized multilayer perceptron to predict flight arrival retards implementing data for domestic flights operated in United States airports. We employed data normalization, sampling techniques, and hyper-parameter tuning to strengthen the reliability of the suggested model. The experimental findings demonstrated that data normalization, sampling approaches, and Bayesian optimization produced the most accurate model with 92.49% accuracy. The achievements of the study were compared to other benchmark research from literature. The time complexity for the proposed model was computed and presented at the end of the investigation.

Corresponding Author:

Hajar Alla
Laboratory of Mathematics, Computer and Engineering Sciences, Department of Mathematics and Computer Science, Faculty of Sciences and Techniques, Hassan First University of Settat
Settat, 26000, Morocco
Email: h.alla@uhp.ac.ma

1. INTRODUCTION

Flight delays can arise from multiple sources, comprising extreme weather, traffic jams, pilot experiences and qualifications, maintenance issues or repairs needed on the aircraft, late-arriving inbound aircraft or passengers, and crew scheduling. In some cases, delays may also be caused by strikes or labor disputes, security concerns or airport construction, delays due to overbooking, or issues with connecting flights. Airport or runway closures or construction, infrastructure, and capacity are also considered for traffic delays. Airport capacity can be considered as the highest number of aircraft and passengers that an airport can handle at any given time. Factors that can affect airport capacity include the number and size of runways, taxiways, gates, and terminals, as well as the efficiency of airport operations. The infrastructure of an airport, if it is not dedicated to handling a lot of aircraft at the same time, can lead to flight delays.

To enhance the safety and efficiency of aircraft operations on runways, precision aids are installed at airports in order to help pilots align their aircraft with the runway and descend to a safe landing. These aids can include equipment like the instrument landing systems (ILS) that help pilots land an aircraft safely, especially in poor visibility conditions. Furthermore, the microwave landing systems (MLS) provide precise guidance to the runway using radio signals for landing in low-visibility conditions, and runway lighting systems, which help pilots identify the location of the runway and its boundaries. Other examples of runway precision aids include visual glide slope indicators, runway alignment indicator lights, and runway threshold lights. These systems are designed to help pilots make safe and accurate landings, even in challenging weather conditions. The area navigation (RNAV) procedure, which stands for area navigation, is also a precision tool used by pilots to fly more direct routes. RNAV is also known as a global positioning system (GPS) navigation, as it commonly uses GPS data to define the position of the aircraft and guide it toward its final point. It is another type of navigation mode available in the flight management system (FMS) installed on modern airplanes. All this precision equipment helps reduce flight times and fuel consumption by avoiding the possibility of going around. In fact, if the runway is equipped with non-precision aids, the aircraft will not be able to land in bad weather and will perform a missed approach instead or divert to an alternate airport.

Following the International Civil Aviation Organization (ICAO) [1], a missed approach or a go-around is an operation performed by an aircraft by stopping and interrupting the approach if the visual reference necessary and the minimum needed for landing has not been established or reached. Figure 1 represents an aircraft go-around procedure. Poor airport infrastructure, such as inadequate runway capacity, outdated or limited terminal facilities, and non-precision aids, can contribute to traffic density and congestion, which leads to flight delays. For this end and as far as we know, we offered novel attributes that are related to the infrastructure of the airport and the precision of its aids, which have never been considered in previous studies, namely, number of runways, runway intersections, longest runway length, shortest runway length, runway precision rate, number of terminals, and number of gates. Other relevant features, such as: airport name, day of week, airline, tail number (registration), flight number, airport of origin, airport of destination, arrival time, departure time, arrival delay (binary), and departure delay (binary) have been taken from the Bureau of Transportation Statistics database (BTS) [2].

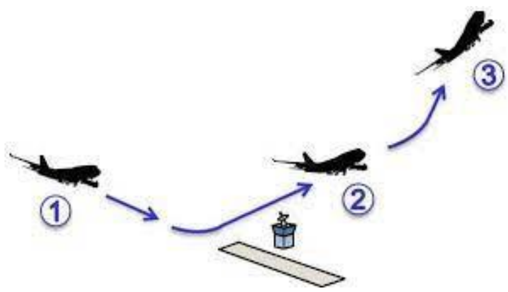

Figure 1. A missed approach operation (source: [3])

This research is intended to deliver an analytical predictive framework that minimizes the impact of delays and cancellations on passengers, airlines, and airport authorities by predicting the arrival delay of a particular flight based on new delay-contributing factors that were not studied in previous research. So as to boost the robustness and efficiency of the model, data normalization was adopted to transform features to be on a similar scale. To better understand the patterns, correlations, and associations between the features and the target variable, which can lead to better prediction performance, a data balancing technique was executed. To boost the performance, robustness, and generalization of the model, we applied an optimization of the hyperparameters using Bayesian optimization in such a way that the model performs better on unseen data.

Numerous studies and research have been conducted on flight delays. The analysis axes addressed the causes of flight delays, their consequences, and measures to avoid them from different perspectives. The investigations comprised the study of statistics, network of things, probability theories, operational research, and machine learning. Liu *et al.* [4] applied an econometric model in order to perform an empirical analysis of flight actual airborne time (AAT) in the US and China. Borsky and Unterberger [5] suggested a difference-in-difference framework based on an econometric analysis to study the impact of sudden changes in meteorological conditions on departure delays using United States data between January 2012 and September 2017. Chen *et al*. [6] aimed to study aircraft delay distribution patterns in one area and demonstrate the impact of delays happening at several airports using a variety of visualization techniques. In order to forecast the likelihood of airline delays during take-off and landing operations, the kernel density function has been utilized by the authors in [7]. Zeng *et al*. [8] focused on complex network theory and the causal inference method to study the propagation of delays and their influence on air traffic control systems. To examine how extreme weather conditions impact punctuality in high-speed rail and aviation services, Chen and Wang [9] utilized both data visualization and statistical analysis. for anticipating and assessing the functional condition of the airport arrival system, Rodríguez-Sanz [10] suggested a two-stage model: the prediction part using a probabilistic Bayesian network and the reliability part with a Markov chain approach.

Statistical methods are generally based on probabilities and approximate measurements, which might result in misleading outcomes. Machine learning has the advantage of resulting in increased accuracy and being able to address enormous quantities of data, automation, and working better with unstructured data, according to Alla *et al.* [11]. Qu *et al.* [12] used a deep learning technique for assessing and projecting aircraft delays. The prediction accuracy was 8.7 percentage points higher compared with the traditional machine learning technique. So as to predict delayed domestic flights operated by American Airlines, Chakrabarty [13] have deployed a gradient boosting classifier model with data sampling and hyperparameter tuning. The suggested method has accomplished an accuracy of 85.73%. For airport delay prediction, a long short term memory (LSTM) neural network framework using historical flight data from several airports in the U.S. from 2015 to 2018 has been proposed by researchers in [8]. According to the experimental results, the suggested technique outperforms existing methods regarding reliability and precision.

Bisandu *et al.* [14] have recommended a deep recurrent neural network (DRNN) model in order to analyze and solve flight delay prediction issues. The suggested method's efficiency and computing time were compared to existing benchmark approaches. In order to help in decision-making and predicting air traffic delays, Nibareke and Laassiri [15] performed analysis on a flight dataset using decision tree, naïve Bayes, and linear regression. The calculation and comparison of accuracy, error, and score metrics have generated decision trees as the best model and naïve Bayes as the weakest one. Huo *et al.* [16] have chosen five methods, which are naïve bayes, logistic regression, k-nearest neighbors, random forest, and decision trees to forecast aircraft delays at Hong Kong International Airport. For estimating aircraft departure delays, Khan *et al.* [17] have proposed a new model using various neural network algorithms combined with different sampling techniques. By examining variables that are in relation with delays such as weather data, operations in airports ground, capacity for demand and flow control qualities, Esmaeilzadeh and Mokhtarimousavi [18] developed a support vector machine (SVM) model to investigate the nonlinear connection of the air delays in the three biggest airports in New York city. Alla *et al.* [19] experimented with gradient boosting, linear regression, extreme gradient boosting, random forest, and decision trees algorithms in order to forecast the arrival time of a specific flight. random forest was the most successful model, with the biggest accuracy of 98.11% compared to the other ones.

The content of this paper is arranged as outlined: section 2 examines the studies and efforts conducted in the area of flight delay estimation. Section 3 highlights the method provided in this research, the algorithms used, as well as all the features analyzed and proposed so as to boost the method's performance. Section 4 provides the empirical results of the suggested technique as well as the time complexity computed in this work. Section 5 discusses the conclusion, viewpoints, and possible future developments.

2. METHOD

2.1. Problem statement

Passengers may experience severe difficulty as a result of flight delays, such as missed connections, missed appointments, and unplanned overnight stays. Airlines can proactively alert customers and provide alternate choices, such as re-booking on another aircraft, if delays are predicted in advance. Flight delays can also be costly for airlines, resulting in increased use of fuel, customer charges paid for the annoyed and dissatisfied passengers, stuffing overtime, and other operational costs. Airlines can take preventive actions to minimize these costs, such as adjusting flight schedules or optimizing ground operations, which can also improve customer satisfaction and loyalty. In order to avoid safety issues engendered by the stress and fatigue of pilots and the crew in general, after experiencing delays, airlines may take precautions to guarantee that their crew members are well-rested and that all safety protocols are followed if traffic delays are known in advance. For this reason, we decided in our study to develop a predictive model that allows passengers, airlines, and airport managers to be aware of delays so as to take proactive actions and measures. The objective is to provide travelers with high flexibility and peace of mind while also assisting airlines in more successfully directing their operations and improving customer satisfaction. Furthermore, it will help airport authorities and leaders in the decision-making process. We started by extracting historical flight information from the BTS database. The data underwent meticulous preparation, segmentation and normalization, ensuring its readiness for examination. From the dataset at hand, we derived different attributes that describe the performance of every flight. To boost the accuracy of our model and as far as we know, we suggested novel features that, according to air transportation organizations and associations, are extremely important and lead to air traffic delays. After that, we segmented the final data as follows: 70% for training and 30% for testing. We utilized the multilayer perceptron (MLP) to train our model. To ensure effective training of the proposed system, hyperparameter tuning was adopted using the Bayesian optimization approach. We decided to perform a data sampling using the synthetic minority oversampling technique (SMOTE) combined with Tomek links. We extensively examined and assessed the evaluation of the proposed model's performance against different metrics. We ended the study with a complex computation. Figure 2 offers a summary of the architecture of the general structure of our approach.

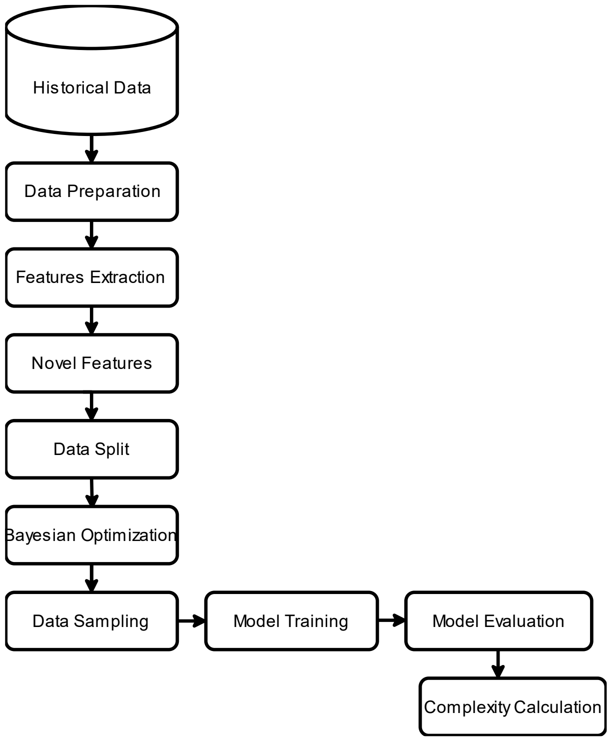

Figure 2. Sequential workflow of the proposed model

2.2. Data collection

Historical flight records for non-stop domestic flights within the United States for the year 2019 were obtained from the BTS [2]. Data on runway dimensions, junctions, and other pertinent details from about 106 airports across the U.S. were retrieved from the Federal Aviation Administration (FAA) [20] airport database. Moreover, information on navigational aids, equipment, and facilities at these airports was accessed from the air navigation database website [21].

2.3. Data preprocessing

Various data mining and machine learning methodologies can be utilized to uncover intriguing insights and patterns from extensive databases [22]. The data preprocessing operations prepare the input dataset for the following data mining actions. They also contribute to enhancing and boosting the accuracy and performance of machine learning systems, particularly in classification, according to [23]. In this paper, we adopted two preprocessing techniques: data cleaning and data normalization.

In data cleaning, the process of data cleaning consists of the elimination of duplicate elements, the handling of missing information, the correction of inconsistent values, and the proper formatting of data, so it can be ready and prepared to be analyzed and used, according to [11]. It is an important phase in the data analysis process. It guarantees that the conclusions obtained from the data are correct and dependable. In data normalization, normalization adjusts the range of attribute values to fit within a new scale. This kind of approach is crucial for classification methods because it enhances the learning process and ensures that attributes with higher values do not dominate those with lower values, as highlighted in [23].

2.4. Features selection

Database features used to predict arrival delays for domestic flights operated in Unites States Airports, we used data for the year 2019. We extracted the statistical information from the BTS [2]. The dataset comprising all the relevant information about the flight is summarized in Table 1. Proposed features are used by international organizations and associations are working to enhance air travel safety and efficiency, as well as to improve the passenger experience and fight against flight delays.

Table 1. Analysis of database features

Feature	Category	Description
Day of month	Numerical	The day of the month in which the flight was executed
Day of week	Numerical	The day of the week during which the trip was executed
Carrier code	Numerical	The airlines designation
Tail number	Numerical	The airlines registration/matriculation
Flight number	Numerical	The number of the flight
Origin	Categorical	The airport of origin
Destination	Categorical	The airport of destination
CRS_DEP	Numerical	The programmed departure time
Actual_DEP	Numerical	The true departure time
DEP delay	Binary	1 if the flight is delayed on departure 0 if not
CRS_ARR	Numerical	The programmed arrival time
ARR delay	Binary	1 if the flight is delayed on arrival, 0 if not (The dependent variable in our research)
Distance	Numerical	The distance in miles between the airport of origin and the airport of destination

The ICAO [1] has established standards and recommended practices (SARPs) for airlines and airports to guarantee safe and efficient operations. Air traffic management, airport operations, and airline safety are among the topics covered by these SARPs. The ICAO [1] also collaborates with member states to put these guidelines into action, improve the safety and efficiency of air transport, and manage flight delays. The European Union (EU) has proposed a "passenger rights" policy that establishes guidelines for airlines to follow in the case of aircraft delays, cancellations, or refused boarding. Passengers are entitled to compensation, assistance, and refunds in particular cases. Airports Council International (ACI) [24] has created the "airport service quality" (ASQ) program to assess consumer contentment with airport services. The initiative solicits passenger feedback on many areas of the airport experience, such as check-in, security, boarding, on-time arrival, and flight delays. Airports may employ this information to improve the passenger experience by identifying areas for improvement. Through its awards for the year 2022 in the US, the FAA [20] has established the airport improvement program (AIP), which funds a number of initiatives and projects such as the building of new and upgraded airport infrastructure, repairs to runways and taxiways, maintenance of airfield components such as lighting or signs, and the purchase of airport equipment.

In order to meet the excessive growth in Brazilian passenger traffic and the extra demand generated by the World Cup 2014, a lot of projects were established notably the construction of a new airport in the nearby town of São Gonçalo do Amarante, which was designated to serve the city of Natal [25]. According to Kennedy [26], intersections between runways are expected to experience additional delays in the National Airspace System as a result of increased wait periods. The need to queue while waiting for an intersecting runway to clear might cause considerable delays. Delays can also be caused by waiting to traverse one of the runways while taxiing. The International Air Transport Association (IATA) [27] has launched the "Airport Collaborative Decision-Making (ACDM)" initiative to optimize the information exchange between airports, airlines and air traffic control, with the objective of minimizing flight delays and enhancing the use of airport resources such as aircraft gates and parking. According to the international federation of Air Traffic Controllers' Associations (IFATCA) [28], a marked increase in the number of approach categories studied has occurred over the past few years. This is mostly caused by the application of cutting-edge technologies including the GPS and the RNAV. Such techniques allow pilots to land with better precision and confidence, lowering the possibility of missed approaches or go-around procedures by giving precise and reliable guidance and position information, which can reduce the occurrence of flight delays.

The ICAO [1] has produced airport design standards that recommend minimum runway lengths depending on the size and type of aircraft expected to operate in the airport. These standards can contribute to guaranteeing that airports are equipped to accept a wide range of aircraft, minimizing the chance of delays due to capacity limits. Similarly, the IATA [27] has produced the ADRM, which offers recommendations on airport planning and development, including runway length and other infrastructure requirements. Airports can ensure that their facilities are enhanced for safety and efficiency by following these standards, which can help reduce flight delays and provide a better passenger experience. This explains why we opted to stay focused on achieving the objectives and policies established by the international organizations and associations previously explained by creating new features that fulfill the needs in terms of airport infrastructure and buildings, the maneuvering area of the aerodrome (runways, taxiways, intersections, and gates,), and the aids used to operate the flights safely and efficiently. Table 2 outlines the features that were proposed in this study.

Table 2. Description of proposed features

Feature	Category	Description
Runway Length	Numerical	The distance of the most used runway
Number of Runways	Numerical	The number of runways in a specific airport
Runways Intersections	Numerical	The number of runway intersection spots in a specific airport
Runways Precision Rate	Binary	1 if the runway is precision-aids equipped (ILS, RNAV, and GNSS) 0 if non-precision-aids equipped (VOR, DME, NDB, and LOCATOR)
Number of Gates	Numerical	The number of airplane parking in a specific airport
Number of Terminals	Categorical	The number of passenger terminals in a specific airport

2.4.1. Runway length

The length of a runway can affect flight delays in different ways. Although longer runways allow for larger and heavier aircraft to take off and land, which can increase the capacity and demand at an airport, it also means a longer time to vacate and clear the runway, which causes delays for following flights. However, if a runway is too short for a particular aircraft, the landing or take-off distance will also be short, and it can lead to a runway excursion. As a preventive action, that aircraft may need to be diverted to another airport, causing a delay. Additionally, inclement weather can also cause flight delays, particularly if the runway is not long enough for an aircraft to safely take off or land in poor visibility conditions, especially if the runway is wet or slippery, which results in a longer braking distance.

2.4.2. Number of runways

The number of runways at an airport can have an impact on flight delays. Having numerous runways allows an airport to serve more air traffic and reduce delays caused by congestion, enabling aircraft to take-off and arrive simultaneously, thereby increasing the airport’s overall capacity. Though an increased number of runways can also result in a complicated architecture for the airport. In general, the more runways an airport has and the more complex its layout, the greater the potential for delays due to runway intersections.

2.4.3. Runways intersections

Flight delays can be increased with runway intersections. In airports, runway intersections are spots where two or more runways cross or join. They can cause traffic delays if aircraft are not able to take off or land on the intersecting runways at the same time, due to safety. Let us take the example in Figure 3 of two intersecting runways, R10 and R09. To take off on R10, aircraft A must cross R09, but aircraft B is authorized to land on R09. For that, aircraft A is forced to wait and hold position to give way to aircraft B for safety reasons.

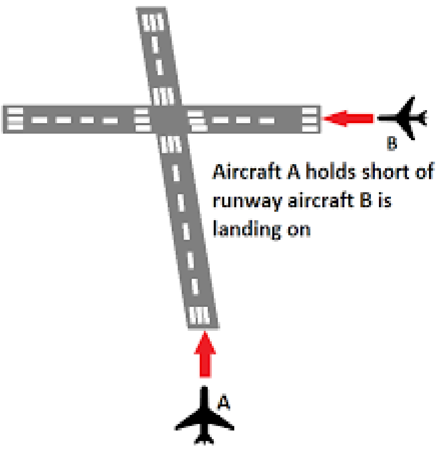

Figure 3. Traffic operations on two intersecting runways (Source: [29])

2.4.4. Runways precision rate

Runway precision aids are systems and equipment that assist pilots in navigating safely and correctly during take-off and landing. Inaccuracies in the information provided by the non-precision runway aid can lead to errors in approach and landing. The discrepancies cause go-arounds and miss approach operations that contribute to flight delays.

2.4.5. Number of gates

Flight delays can also be affected by the number of gates at an airport. It is critical and mandatory for an airport to have a sufficient number of parking spaces to manage the number of flights that are planned to take-off and land. If there are not enough gates, flights may have to wait for one to become available, causing delays.

2.4.6. Number of terminals

Airports with more terminals tend to welcome more flights and passengers. This can lead to more congestion, longer security lines, and more potential for delays. Nevertheless, airports with few terminals can sometimes experience flight delays, especially if one or more are unserviceable.

2.5. Multilayer perceptron

The MLP is a popular and basic neural network generally used for classification problems. Basically, it is a feed-forward neural network composed of many perceptron [11]. With one or more hidden layers, the MLP is generally employed for pattern recognition, classification, prediction, and function approximation [30]. The input layer nodes receive the input data, while the output layer nodes generate the network's predictions. The neurons in each layer use activation functions to calculate a weighted sum of their inputs and produce a non-linear output.

The MLP follows the process of multiplication, summation, and activation used in neural networks [11], as expressed by (1):

$$Y = F\left(\sum_{i=0}^{n} w_i * x_i + b\right) \quad (1)$$

where x_i refers to the i^{-th} input where i ranges from 0 to n inputs. w_i indicates the weight matrices for both the hidden and output layers, with i spanning from 0 to n inputs. b denotes the bias term. F represents the activation function. Y signifies the output value.

Multilayer perceptron is widely used for a variety of applications. Alla *et al.* [11] used a MLP neural network with selective training for the prediction of delays on arrival. To estimate the coefficient of soft soil consolidation, Pham *et al.* [31] combined MLP and biogeography-based optimization (BBO).

The proposed method, MLP-BBO, had the biggest predictive performance with the lowest root mean square error (RMSE) of 0.397 compared with other models. Mubarek and Adali [32] adopted a MLP for fraud detection. The proposed model revealed that it was the most accurate, with the greatest degree of accuracy of 99.47% using nine selected features. For drought forecasting, Zulifqar *et al.* [33] applied and tested the MLP in several climatological stations situated in the northern area and Pakistan. The model was able to predict drought conditions with different time scales and higher accuracy. To determine the warming and calming requirements of energy-efficient buildings, Xu *et al.* [34] have utilized and optimized an MLP method using different optimization algorithms. Radhakrishnan *et al.* [35] have developed an MLP model for predicting mechanical ventilator settings by changing the hidden layers and comparing the results. The best model was the one with three hidden layers.

In the present study, the MLP was implemented for assessing and foreseeing the occurrence of flight delays using new features. We opted for the MLP for many reasons: i) Because MLPs have few parameters, they can be employed by individuals without previous experience, and their implementation techniques are easy to understand [36]; ii) MLPs have the capability to be utilized across diverse fields for solving a variety of problems [36]; iii) MLPs serve as tools for discrimination, recognition of patterns, empirical modeling, and many other applications [36]; iv) When applied to similar issues, MLPs often outperform standard statistical approaches [36]; v) While traditional linear models struggle to model data with nonlinear properties, MLPs can effectively capture both linear and nonlinear interactions [36]; and vi) MLPs are effective in extracting structural or pattern characteristics from both static and dynamic data [36].

To create a machine learning system, weight parameters are set up and adjusted using an optimization approach. According to the study [37], this action keeps occurring till the objective function attains a minimum or the accuracy reaches a maximum. In our investigation, we employed Bayesian optimization to tune the hyperparameters.

2.6. Bayesian optimization

The hyperparameter optimization paradigm involves the use of four pertinent elements: an estimator (which might be a regression or classification model with some sort of objective function), a defined search space, a technique for exploring or optimizing hyperparameter combinations, and an evaluation metric for comparing the effectiveness of various hyperparameter configurations. Bayesian optimization models compute the next hyperparameter by considering the previous outcomes of tested hyperparameter values avoiding numerous unnecessary evaluations. As a consequence, the Bayesian approach can find the best hyperparameter combination in fewer iterations than other optimization techniques such as random search and grid search. Bayesian optimization employs two main components to select the next hyperparameter configuration: a surrogate model and an acquisition function [37].

2.6.1. Surrogate model: Gaussian processes

The surrogate model is applied to direct the search for the target model's global optimum. The most common surrogate model for objective function modeling is the gaussian process (GP), which follows a normal distribution according to (2). It is an advanced probabilistic model that is widely used in machine learning for regression and classification problems [37].

$$p(y \mid \mathrm{x}, \mathrm{D}) = \mathrm{N}(\mathrm{y} \mid \hat{\mu}, \hat{\sigma}^2) \tag{2}$$

where D corresponds to the hyper-parameter configuration space, and $y = f(x)$ denotes the evaluation outcome for each hyper-parameter value x, σ^2 is the covariance and μ the mean.

After making predictions, the subsequent evaluation points are selected based on the confidence intervals generated by the BO-GP model. Every new data point is incorporated into the dataset, and the BO-GP model is updated accordingly. This process is reiterated several times until the set stopping criteria are met. For a dataset with size n, the BO-GP model has a time complexity of $O(n3)$ and a space complexity of $O(n2)$. A significant drawback of the BO-GP model is its cubic time complexity concerning the number of instances, which impacts its scalability and parallel processing capabilities. Additionally, the BO-GP model is mainly designed for optimizing continuous variables [37].

2.6.2. Acquisition function

To select the next candidate from the search space, the acquisition function can be described to mean the anticipated gain:

$$A(x) = E[G(x) \mid \mathrm{X,y}] \tag{3}$$

where

$$G(x) := min(y) - f(x). G: R^d \rightarrow R \tag{4}$$

is the gain for unknown solutions. In every loop, an additional candidate solution x' is selected via the maximization of the acquisition function [38]:

$$x' = arg_{x \in S}\, max\, A\,(x) \tag{5}$$

2.6.3. Bayesian algorithm

The key stages of the Bayesian optimization technique are demonstrated in Algorithm 1 [38]. The initial round produces the basic datasets x and y. A stochastic surrogate model of the objective function is consequently developed. Following that, a sample is chosen by maximizing the acquisition function. The sample is assessed via the objective function. The surrogate model is subsequently revised with the novel data. This technique will continue until the maximum number of iterations is met.

Algorithm 1. Bayesian optimization algorithm
Require: An acquisition function A
1. Construct the primary data set X, y using the objective function f.
2. Build the Gaussian process model utilizing the dataset X, y
3. While the stopping criteria have not been met, do
4. Maximize the acquisition function:
$$X' = \underset{X \in S}{arg\, max}\, A(X)$$

5. Evaluation: $y' \leftarrow f(X')$
6. Augment the data set by adding X', y' to X, y
7. Rebuild the Gaussian process model for f using the expanded dataset X, y
8. End while

3. RESULTS AND DISCUSSION

3.1. Experimental setup

The simulations were conducted in Python 3.9.15 using the scikit-learn library. The program was coded on an HP computer with Windows 10. In our research, delayed flights are represented as 1 and on-time flights as 0.

3.2. Results and analysis

Choosing the most appropriate hyperparameters possesses an enormous effect on the performance model. A selection of different hyperparameter values has been considered as a search space to build and optimize the proposed MLP model: i) Hidden layer size: (100,), (50), (50,50), (100,100), (100,50), (50,100); ii) Activation function: logistic, Tanh, and ReLU, iii) Solver: SGD, Adam; iv) Alpha: 0.01, 1e-6, 1e-2; v) Learning rate: constant, invscaling, adaptive; and vi) Max iteration: 100, 500, 1000, 2000.

Actions can vary in duration and occur concurrently, possibly overlapping in time [39]. To validate the choice of the hyperparameter optimization method, we compared the Bayesian optimization with that of grid search [37] and random search [37] techniques regarding the accuracy and the elapsed time. According to the results in Table 3, we deduce that the Bayesian algorithm is the best optimization method to be proposed in this study. In Table 4, the combination of the best hyperparameters and the best accuracy generated by the Bayesian optimization is presented. Our proposed model was evaluated in terms of normalization, data sampling, and parameter tuning. Table 5 shows the evaluation results before and after normalizing the data regarding the recall, the F1 score, the accuracy, and the precision. It shows the importance of data normalization to enhance the efficiency and training consistency of the model being proposed. We performed the data balancing using the Smote-Tomek technique, which combines under sampling and oversampling for better sampling. Table 6 presents the evaluation findings with and without data sampling. Based on the results, we deduce that balancing data with the Smote-Tomek technique was very successful, which increased the accuracy to 90.13%. All the other metrics were improved compared with the previous findings.

Table 7 highlights the effect of optimizing the suggested MLP model by monitoring the results before and after the Bayesian optimization. We notice that the Bayesian optimization has resulted in a higher accuracy of 92.49%. All the other metrics were improved compared with the previous findings. Figure 4 monitors the receiver operating curve (ROC) for the final proposed model with normalization, sampling and optimization. The area under the curve is 0.9230.

Table 3. Comparison of hyper-parameter algorithms

	Bayesian	Grid search	Random search
Accuracy (%)	92.49	88.08	87.61
Elapsed time (seconds)	1063.79	3994.03	4382.37

Table 4. The best hyper-parameters and accuracy of the proposed model

Model	Hyper-parameter (HP)	The best HP value	The best accuracy
Multilayer perceptron	Activation	ReLU	92.49%
	Alpha	0.01	
	Hidden layer sizes	(50)	
	Learning rate	constant	
	Max Iter	1000	
	Solver	Adam	

Table 5. Evaluation metrics in terms of data normalization

	With normalization	Without normalization
Accuracy (%)	79.40	75.98
Precision	78.96	74.00
Recall	77.73	70.66
F1 score	78.13	73.22

Table 6. Evaluation metrics in terms of data balancing

	With sampling	Without sampling
Accuracy (%)	90.13	79.40
Precision	91.10	78.96
Recall	79.93	77.73
F1 score	90.48	78.13

Table 7. valuation metrics in terms of hyperparameter tuning

	With optimization	Without optimization
Accuracy (%)	92.49	90.13
Precision	92.15	91.10
Recall	80.27	79.93
F1 score	92.13	90.48

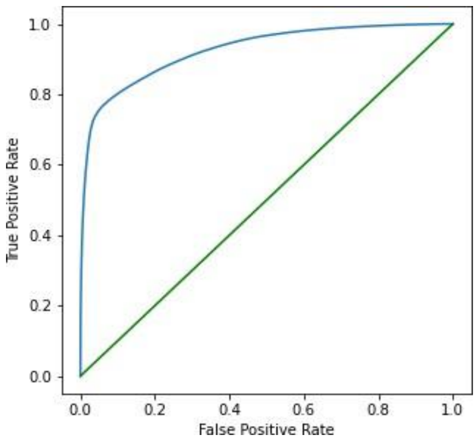

Figure 4. ROC curve for the proposed model. Source: own calculation

3.3. Benchmark findings

To demonstrate the efficacy of our suggested approach, we compared our findings in Table 8 to those from previous studies. The comparison was made based on the recall, the F1 score, the accuracy, and the precision. We remark that our suggested strategy is the most accurate when compared to the others.

Table 8. Related works and proposed method comparison

Model	Features used	Objective	Metric 1: accuracy (%)	Metric 2: F1 score (%)	Metric 3: precision (%)	Metric 4: recall (%)
Henriques *et al*. [40]	Flight information, weather data; aircraft data; delay propagation information	To predict flight arrival delays	85.63	79.00	-	-
Stefanovic *et al*. (arrival) [41]	Flying period, trip number, airline, destination, origin, climate, ceiling data, velocity of the wind, wind direction, visibility, planned time, classes	To predict the flight arrival time deviation for Lithuanian airports	47.43	50.77	47.43	56.73
Stefanovic *et al*. (departure) [41]	Flying period, trip number, airline, destination, origin, climate, ceiling data, velocity of the wind, wind direction, visibility, planned time, classes	To predict the flight departure time deviation for Lithuanian airports	85.65	87.86	85.65	90.90
Pamplona *et al*. [42]	Flight data, delay justification code	To predict air traffic delays	91.30	77.00	87.00	69.00
Vonitsanos *et al*. [43]	Flight data, same-origin-flights count, average airline delay, average origin delay, cancelled (classification), arrival delay (regression)	To forecast air flight delays	-	54.35	54.37	53.40
Our model	Flight data, distance, runway length, number of runways, runways intersections, runways precision rate, number of gates, number of terminals.	To predict flight arrival delays	92.49	92.13	92.15	80.27

3.4. Time complexity

According to [11], $O(a * b * c)$ is the time complexity of matrix multiplication for:

$$M' = M_{ab} * M_{bc} \tag{6}$$

In a neural network, to go from layer a to b, we have (7):

$$S_b = W_{ba} * X_a \tag{7}$$

where S_b denotes the calculated amount of the weights W_{ba} and the input data X_a. By implementing the activation function, we get:

$$Y_b = f(S_b) \tag{8}$$

where Y_b is the output.

If we have N layers, the operation will execute N-1 times (including input and output layers). In this scenario, we have three layers. Therefore, we will need two matrices for storing the weights: W_{ba} and W_{cb}. Here, a represents the number of nodes in the input layer, b signifies the number of nodes in the second layer (hidden layer), and c indicates the number of nodes in the output layer. W_{ba} is a matrix having b rows and a column which holds the weights connecting layer a to layer b, whereas W_{cb} has c rows and b columns that include the weights connecting layer b to layer c.

Given n training instances, we possess:

$$S_{bn} = W_{ba} * X_{an} \tag{9}$$

According to (6), time complexity of (9) is $O(b * a * n)$. Then we apply the activation function:

$$Y_{bn} = f(S_{bn}) \tag{10}$$

which has $O(b * n)$ as time complexity. Overall, we are dealing with the subsequent complexity: $O(b * a * n + b * n) = O(b * n * (a + 1)) = O(b * a * n)$.

Using the same calculation, for going from b to c in (11), we have $O(c * b * n)$.

$$S_{cn} = W_{cb} * X_{bn} \tag{11}$$

In total, the temporal complexity of feedforward propagation is going to be:

$$O(b * a * n + c * b * n) = O(n * (ab + bc) = O(n * a * b * c)$$

For one epoch (number of iterations), It is equivalent to $O(n * a * b * c)$. For i epochs, it is equivalent to $O(i * n * a * b * c)$. Also, the time complexity of the Bayesian optimization algorithm is $O(n^3)$. The total time complexity is: $O(i * n * a * b * c) + O(n^3)$. Since $i, a, b,$ and c are unchanged, the complexity can be simplified to: $O(i * n * a * b * c) + O(n^3) = O(n) + O(n^3) = O(n^3)$. The global temporal complexity of our proposed model is then: $O(n^3)$.

4. CONCLUSION

The aviation industry has recently experienced important and significant growth, leading to a massive increase in passenger and cargo demand. In fact, air travel is the safest, secure, and most rapid means of transport so far. The massive increase due to the high demand resulted in a density in the air and on the ground (airports), which has caused flight delays. Researchers and academics from all over the world are constantly working to find new approaches and strategies to address the challenge of aircraft delays. In this study, we proposed an optimized MLP using data for domestic flights operated in U.S. Airports so as to foresee flight delays. Not only the traditional features extracted from the database were used, but we also proposed new features that, according to air transportation organizations and associations, are very relevant and contributors to air traffic delays. To enhance the accuracy of the suggested model, we used data normalization, sampling methods, and hyper-parameter optimization. After a comparison of the elapsed time and accuracy for the Bayesian algorithm, grid search and random search techniques, we chose the Bayesian optimization to adjust the hyper-parameters for the MLP. The experimental results proved that data normalization, sampling techniques, and Bayesian optimization generated the most accurate model with the highest accuracy of 92.49%. Recall, precision and F1 score were also improved. To demonstrate the robustness of our suggested technique, we conducted a high-level comparison of our results with those from previous studies regarding the recall, the F1 score, the accuracy, and the precision. The time complexity for the proposed model was computed and presented at the end of the study. In the future, we would like to explore more variables that are responsible for flight delays, such as weather information, the busyness of the airports analyzed, and the type of aircraft studied.

REFERENCES

[1] ICAO, *Procedures for air navigation services (PANS) - aircraft operations - volume I - flight procedures (Doc 8168)*, 6th ed. The Internationa Civil Aviation Organization (ICAO), 2018.

[2] U.S. Department of Transportation: Beruau of Transportation Statistics, "On-time: reporting carrier on-time performance (1987-present)," *U.S. Department of Transportation: Federal Aviation AdministrationDepartment of Transportation: Beruau of Transportation Statistics*. https://www.transtats.bts.gov/DL_SelectFields.aspx?gnoyr_VQ=FGJ&QO_fu146_anzr=b0-gvzr (accessed Feb. 15, 2023).

[3] T. Lombaerts, S. Schuet, D. M. Acosta, J. Kaneshige, and K. H. Shish, "Piloted simulator evaluation of maneuvering envelope information for flight crew awareness," in *AIAA Guidance, Navigation, and Control Conference*, Jan. 2015, pp. 1–24, doi: 10.2514/6.2015-1546.

[4] K. Liu, Z. Zheng, B. Zou, and M. Hansen, "Airborne flight time: a comparative analysis between the U.S. and China," *Journal of Air Transport Management*, vol. 107, Mar. 2023, doi: 10.1016/j.jairtraman.2022.102341.

[5] S. Borsky and C. Unterberger, "Bad weather and flight delays: the impact of sudden and slow onset weather events," *Economics of Transportation*, vol. 18, pp. 10–26, Jun. 2019, doi: 10.1016/j.ecotra.2019.02.002.

[6] C. Chen, C. Li, J. Chen, and C. Wang, "VFDP: visual analysis of flight delay and propagation on a geographical map," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 3510–3521, Apr. 2022, doi: 10.1109/TITS.2020.3037191.

[7] I. Ostroumov *et al.*, "A probability estimation of aircraft departures and arrivals delays," in *Computational Science and Its Applications–ICCSA 2021–Lecture Notes in Computer Science*, vol. 12950, 2021, pp. 363–377, doi: 10.1007/978-3-030-86960-1_26.

[8] L. Zeng, B. Wang, T. Wang, and Z. Wang, "Research on delay propagation mechanism of air traffic control system based on causal inference," *Transportation Research Part C: Emerging Technologies*, vol. 138, May 2022, doi: 10.1016/j.trc.2022.103622.

[9] Z. Chen and Y. Wang, "Impacts of severe weather events on high-speed rail and aviation delays," *Transportation Research Part D: Transport and Environment*, vol. 69, pp. 168–183, Apr. 2019, doi: 10.1016/j.trd.2019.01.030.

[10] Á. Rodríguez-Sanz, F. G. Comendador, R. A. Valdés, J. Pérez-Castán, R. B. Montes, and S. C. Serrano, "Assessment of airport arrival congestion and delay: prediction and reliability," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 255–283, Jan. 2019, doi: 10.1016/j.trc.2018.11.015.

[11] H. Alla, L. Moumoun, and Y. Balouki, "A Multilayer perceptron neural network with selective-data training for flight arrival delay prediction," *Scientific Programming*, vol. 2021, pp. 1–12, Jun. 2021, doi: 10.1155/2021/5558918.

[12] J. Qu, S. Wu, and J. Zhang, "Flight delay propagation prediction based on deep learning," *Mathematics*, vol. 11, no. 3, pp. 1–24, Jan. 2023, doi: 10.3390/math11030494.

[13] N. Chakrabarty, "A Data mining approach to flight arrival delay prediction for American airlines," in *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*, Mar. 2019, pp. 102–107, doi: 10.1109/IEMECONX.2019.8876970.

[14] D. B. Bisandu, I. Moulitsas, and S. Filippone, "Social ski driver conditional autoregressive-based deep learning classifier for flight delay prediction," *Neural Computing and Applications*, vol. 34, no. 11, pp. 8777–8802, Jun. 2022, doi: 10.1007/s00521-022-06898-y.

[15] T. Nibareke and J. Laassiri, "Using big data-machine learning models for diabetes prediction and flight delays analytics," *Journal of Big Data*, vol. 7, no. 1, pp. 1–18, Dec. 2020, doi: 10.1186/s40537-020-00355-0.

[16] J. Huo, K. L. Keung, C. K. M. Lee, K. K. H. Ng, and K. C. Li, "The prediction of flight delay: big data-driven machine learning approach," in *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Dec. 2020, pp. 190–194, doi: 10.1109/IEEM45057.2020.9309919.

[17] W. A. Khan, H.-L. Ma, S.-H. Chung, and X. Wen, "Hierarchical integrated machine learning model for predicting flight departure delays and duration in series," *Transportation Research Part C: Emerging Technologies*, vol. 129, Aug. 2021, doi: 10.1016/j.trc.2021.103225.

[18] E. Esmaeilzadeh and S. Mokhtarimousavi, "Machine learning approach for flight departure delay prediction and analysis," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 8, pp. 145–159, Aug. 2020, doi: 10.1177/0361198120930014.

[19] H. Alla, L. Moumoun, and Y. Balouki, "Flight arrival delay prediction using supervised machine learning algorithms," in *Intelligent Systems in Big Data, Semantic Web and Machine Learning, Vol 1344*, Springer, Cham, 2021, pp. 231–246.

[20] U.S. Department of Transportation: Federal Aviation Administration, "Aviation data and statistics," *U.S. Department of Transportation: Federal Aviation Administration*, 2024. https://www.faa.gov/data_research/aviation_data_statistics (accessed Feb. 15, 2023).

[21] Department of Transportation-Federal Aviation Administration, "Aeronautical Information Publication, United States of America, 29 edition," *Department of Transportation, Federal Aviation Administration*, 2022. Accessed: Feb. 15, 2023. [Online], Available: https://www.faa.gov/air_traffic/publications/media/aip_bsc_w_amd_1_dtd_11-3-22.pdf

[22] A. Dahbi, Y. Balouki, and T. Gadi, "Using multiple minimum support to auto-adjust the threshold of support in apriori algorithm," in *International Conference on Soft Computing and Pattern Recognition*, 2017, pp. 111–119, doi: 10.1007/978-3-319-76357-6_11.

[23] L. Moreira, C. Dantas, L. Oliveira, J. Soares, and E. Ogasawara, "On evaluating data preprocessing methods for machine learning models for flight delays," in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–8, doi: 10.1109/IJCNN.2018.8489294.

[24] ACI, "Airport customer experience and ASQ," *Airports Council International*, https://aci.aero/programs-and-services/asq// (accessed Mar. 04, 2023).

[25] A. F. Braga, E. P. Guedes, and R. P. Filho, "FIFA World Cup 2014: analysis of the accessibility of airports," *Unpublished Paper, ResearchGate*, Jul. 2010.

[26] R. J. Kennedy, "Four runway configuration types and their relation to arrival delays," M.S. thesis, Graduate School, Purdue University, West Lafayette, India, 2015.
[27] IATA, "Airport Development Reference Manual (ADRM)," *The International Air Transport Association (IATA)*. https://www.iata.org/en/publications/manuals/airport-development-reference-manual/ (accessed Mar. 04, 2023).
[28] IFATCA, "Knowledge base category: TPM," *International Federation of Air Traffic Controllers' Associations (IFATCA)*, 2018. https://ifatca.wiki/kb-category/tpm/ (accessed Jan. 09, 2023).
[29] Australian Government: Bureau of Meteorology, "YMML air traffic operations," *Australian Government: Bureau of Meteorology*, 2023. Accessed: Jan. 09, 2023, [Online], Available: http://www.bom.gov.au/aviation/data/education/reference-card-ymml.pdf
[30] J. S. Sonawane and D. R. Patil, "Prediction of heart disease using multilayer perceptron neural network," in *International Conference on Information Communication and Embedded Systems (ICICES2014)*, Feb. 2014, pp. 1–6, doi: 10.1109/ICICES.2014.7033860.
[31] B. T. Pham, M. D. Nguyen, K.-T. T. Bui, I. Prakash, K. Chapi, and D. T. Bui, "A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography-based optimization for predicting coefficient of consolidation of soil," *CATENA*, vol. 173, pp. 302–311, Feb. 2019, doi: 10.1016/j.catena.2018.10.004.
[32] A. M. Mubarek and E. Adali, "Multilayer perceptron neural network technique for fraud detection," in *2017 International Conference on Computer Science and Engineering (UBMK)*, Oct. 2017, pp. 383–387, doi: 10.1109/UBMK.2017.8093417.
[33] Z. Ali *et al.*, "Forecasting drought using multilayer perceptron artificial neural network model," *Advances in Meteorology*, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/5681308.
[34] Y. Xu, F. Li, and A. Asgari, "Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms," *Energy*, vol. 240, Feb. 2022, doi: 10.1016/j.energy.2021.122692.
[35] S. Radhakrishnan, S. G. Nair, and J. Isaac, "Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning," *Biomedical Signal Processing and Control*, vol. 71, Jan. 2022, doi: 10.1016/j.bspc.2021.103170.
[36] Y.-S. Park and S. Lek, "Artificial neural networks: multilayer perceptron for ecological modeling," in *Developments in Environmental Modelling*, vol. 28, Elsevier, 2016, pp. 123–140.
[37] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/j.neucom.2020.07.061.
[38] H. Wang, B. van Stein, M. Emmerich, and T. Back, "A new acquisition function for Bayesian optimization based on the moment-generating function," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2017, pp. 507–512, doi: 10.1109/SMC.2017.8122656.
[39] M. Bouhdadi, Y. Balouki, and E. maati Chabbar, "Meta-modelling syntax and semantics of structural concepts for open networked enterprises," in *Computational Science and Its Applications – ICCSA 2007*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 45–54.
[40] R. Henriques and I. Feiteira, "Predictive modelling: flight delays and associated factors, hartsfield–Jackson Atlanta international airport," *Procedia Computer Science*, vol. 138, pp. 638–645, 2018, doi: 10.1016/j.procs.2018.10.085.
[41] P. Stefanovič, R. Štrimaitis, and O. Kurasova, "Prediction of flight time deviation for Lithuanian Airports using supervised machine learning model," *Computational Intelligence and Neuroscience*, vol. 2020, pp. 1–10, Oct. 2020, doi: 10.1155/2020/8878681.
[42] D. A. Pamplona, L. Weigang, A. G. de Barros, E. H. Shiguemori, and C. J. P. Alves, "Supervised neural network with multilevel input layers for predicting of air traffic delays," in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–6, doi: 10.1109/IJCNN.2018.8489511.
[43] G. Vonitsanos, T. Panagiotakopoulos, A. Kanavos, and A. Tsakalidis, "Forecasting air flight delays and enabling smart airport services in Apache Spark," in *Artificial Intelligence Applications and Innovations - AIAI 2021 IFIP WG 12.5 International Workshops - IFIP Advances in Information and Communication Technology, vol 628*, 2021, pp. 407–417.

BIOGRAPHIES OF AUTHORS

Hajar Alla is an air traffic controller at Marrakech Menara Airport. She is presently a Phd student at Hassan First University of Settat. She holds a master's degree in air traffic management and a bachelor's in computer science engineering. Her research areas are artificial intelligence, machine learning, deep learning, neural networks, air transportation and aviation. She has recently published research papers and conferences on flight time prediction based on different delay factors. She can be contacted at her institutional email: h.alla@uhp.ac.ma or her personal mail: allahaajar@gmail.com.

Youssef Balouki is presently working as professor in the Faculty of Science and Techniques in Hassan First University of Settat. He holds a PhD in computer science from University of Mohammed V, Agdal, Rabat., Morocco. He obtained a bachelor's in computer science from the University of Cadi Ayyad, Marrakech, Morocco. His research areas and expertise are advanced machine learning, data mining, aviation, formal verification, software development, software testing, software analysis, java programming and object-oriented programming. He can be contacted at his institutional email: y.balouki@uhp.ac.ma or his personal mail: baloukiy@yahoo.fr.