Simulation based Predictive analysis of Indian Airport transportation system using Computational intelligence techniques

Perepi Rajarajeswari[1] , P Lalitha Surya Kumari[2,*] , A. Sree Lakshmi[3] , S.R. Mugunthan[4] ,
Prathipati Ratna Kumar[5] , Pranayanath Reddy Ananthula[6]

1. Vellore Institute of Technology University ROR – School of Computer science and Engineering – Department of Software Systems – Vellore/Tamilnadu – India. 2. Koneru Lakshmaiah Education Foundation ROR – Department of Computer Science and Engineering – Hyderabad/Telangana – India. 3. Geethanjali College of Engineering and Technology – Department of Computer Science and Engineering– Hyderabad/Telangana – India. 4. Sri Indu College of Engineering and Technology – Department of Computer Science and Engineering – Hyderabad/Telangana – India. 5. Koneru Lakshmaiah Education Foundation ROR - Department of Computer Science and Engineering – Hyderabad/Telangana – India. 6. GITAM University ROR – School of Technology – Department of Computer Science and Engineering – Hyderabad/Telangana – India.

*Correspondence author: vlalithanagesh@gmail.com

ABSTRACT

Normally, flight delays and cancellations have significant impact on airlines operations and passenger's satisfaction. Flight delays reduce the performance of airline operations and make significant effect on airports on time performance. Previously statistical models have been used for flight delays analysis. This study was applied in Indian aviation industry and it has given statistical analysis of domestic airlines. In this research paper, we have applied Machine Learning models with the help of computational intelligence techniques for predicting airport transport management system. We have also applied computational intelligence techniques such as Particle Swarm Optimization (PSO) and Ant Colonization Optimization (ACO) to optimize the prediction model for delay period time and calculating the most optimal dependability. We have made comprehensive analysis of Data Efficiency Model for different airlines with various approaches as well as comparative analysis of accuracy for predicting airport model by using various machine learning models. In this study we have presented invaluable insights for the analysis of flight delay models.

Keywords: Particle Swarm Optimization; Ant Colonization Optimization; Delay Period; Airport delay; Air transport.

INTRODUCTION

India is on the verge of overtaking the UK and becoming the third largest aviation market in 2024, with growing traffic. In 2018, air passenger traffic amounted to 341.05 million. Travel and tourist contributions are expected to produce the Indian GDP (gross domestic product) by increasing US$247.3 billion in 2018. The expenditures in business travel climbed in 2018 from $201,71bn in 2017 to $234,44bn in 2018 and from $11,61bn in 2017 to $12,86bn in 2018. Pilots, flight attendants and aircrafts can also have extraordinary schedules to preserve plans for maintaining airplanes. Hence, any disruption in the device can have an impact on the subsequent flights of the identical airline (Rebollo and Balakrishnan 2014). Flight extends prediction difficulty can be handled by taking distinct factors of view: (i) lengthen propagation, (ii) root extend and cancellation. Reynolds-Feighan and Button (1999)

focused on assessment of the capacity and congestion levels at European airports. Wong *et al.* (2002) introduced an optimization model for assessing flight technical delay. Traffic may be especially affected when transportation networks are interrupted, which causes delays and makes passengers anxious about their arrival time. Delays cause additional travel time. While travelers, on the other hand, may modify their trip plans to meet unforeseen delays, delays in travel time may necessitate changes to itineraries. This decreases the overall usefulness of the device and therefore increases the cost.

Therefore, researchers targeted on cancellation evaluation attempt to determine which prerequisites lead to cancellations. Moreover, it explores the airlines' decision-making technique for deciding on the flights to be canceled. The estimation of flight delays can enhance the tactical and operational choices of airports and airways managers, in addition to warn passengers to rearrange their plans (D'Ariano *et al.* 2012). To recognize the whole flight ecosystems at higher level, great volumes of facts from business aviation are accrued each and every second and saved in databases.

This evaluation is conducted by the way of organizing a flight lengthening process called Lookup taxonomy, which organizes methods in accordance to the kind of problem scope, information issues and computational methods. The time spent on flight delays costs airlines extra money since it involves using people, fuel, and maintenance on planes. To lessen aircraft delays and increase capacity, several initiatives such as developing more runways, including better runway layouts, upgrading air traffic control infrastructure and more advanced air traffic control procedures are in place. Due to flight delays and cancellations, airlines and passengers are both affected. When passengers get disgruntled, it results in a loss of market share. In the event of a delay or cancellation, not only do your costs rise, but the time spent as a consequence of the delay or cancellation is also lost, reducing your productivity as well. Complexity and intractability are often due to many contributing factors. Anomalous weather and technical issues, such as facility capacity, poor scheduling, process changes and limited buffer time are often characterized as consequences. Difficulty finding the root cause and devising appropriate remedies because of the variety of reasons for flight delays (Schaefer and Millner 2001). Recent studies have shown that historical observations-based approaches that are completely data-driven are free of previous constraints and have the capacity to accurately reflect future dynamic features (the other approach to enhance system cognition and decision-making is to fully use information based on the available data. If confronted with poor weather, for example, it is feasible to use information from the past to examine what the air traffic controllers did on other days when dealing with comparable weather circumstances.

Several prior research has focused on identifying permanent patterns in air traffic management (Liu *et al.* 2014). Simulation, machine learning methods, probabilistic methods, statistical methods, regression methods, Correlation analysis, Prediction and classification methods for determining airport delay periods. In the commercial aviation industry, improving flight operations efficiency can be considered as one of the important challenges of flight analysts. Air traffic flow managers are facing more problems managing airline operations properly due to bad weather conditions. Actually, delay is considered as one of the important costs of airport stakeholders,because flight delays are having negative impact on the economic model of airlines and passenger's satisfaction. Air traffic transport management is a multidimensional process that can evaluate many contributing factors to specific different conditions. Effective strategies as data processing techniques are used for reducing delay periods and NASA development. In this research paper we have applied computational intelligence techniques for predicting flight delay periods at different times. The present study is focused on the investigation of flight delay periods with datasets by using Machine Learning models. Additionally, studies are focused on implementation of Machine Learning approach for predicting the Indian Airport transport management system. We have summarized all research techniques which are to be used to describe various methods of Machine Learning techniques for building prediction models. This analysis is performed by using computational methods.

The major contributions of the paper are: in this research paper we have analyzed delay periods of Aircraft system based on various Machine Learning algorithms, also, we have applied computational intelligence techniques for reducing the flight delays based on computational intelligence techniques and made comparative analysis for predicting the flight delay periods.

The organization of the paper is as follows. Section 2 describes the related work, section 3 analyzes the proposed system, section 4 provides results and discussion and section 5 concludes the paper and future work.

RELATED WORK

There has been rapid growth in passenger travel in the new century, particularly due to increasing household income and lower fares. Passenger traffic in India in FY 2018 was 341.05 million. CAGR growth for FY16-FY20 was 11.13%. It is anticipated that India would pass the United Kingdom as the third-largest air passenger market within a decade. Between 2016 and 2036, India is estimated to have around 480 million fliers, more than either Japan or Germany (both about 225 million) combined. Domestic passenger traffic had an average growth rate of 12.91% per year during the two-year period from FY16–FY20. International passenger traffic has grown by 5.01% each year, on average, during the previous five years. Passenger traffic in FY20 reached 274.5 million domestically, while foreign traffic was 66.55 million. In 2022, passenger traffic was 166.8 million at Indian airports and international passenger traffic was 22.1 million.

Over the next four fiscal years, domestic aircraft movement grew by 9.83% per year, while overseas movement rose by 3.57% each year. The role of the private sector in the industry has risen since liberalization. 'In principle', the Indian government has granted 'approval' to the construction of 19 airports, of which seven would be built via a public-private partnership with an investment of Rs 27,000 crore (US$ 41.89 billion). The media, academics, and even authorities all around the globe have been intrigued by flight delays during the past decade (Lawson and Castillo 2012). According to the Department of Infrastructure and Regional Development (DIRD), flight delays in arrivals and departures account for 17.3% and 16% of total arrivals and departures correspondingly (DIRD 2022). According to the data, flight delays have increased over the last several years (Abdel-Aty *et al.* 2007). Delays for departures rose from 16.2% in 2016 and 2017 to 18% in 2017 and 2018. As a consequence of population growth and increased air travel, delays will rise. Other factors such as the airline's operating model may also influence flight delays. The average flight delay trend in Australia for the country's four main domestic airlines from 2004 to 2015 is shown in Fig. 1 (Mohammadian *et al.* 2019).
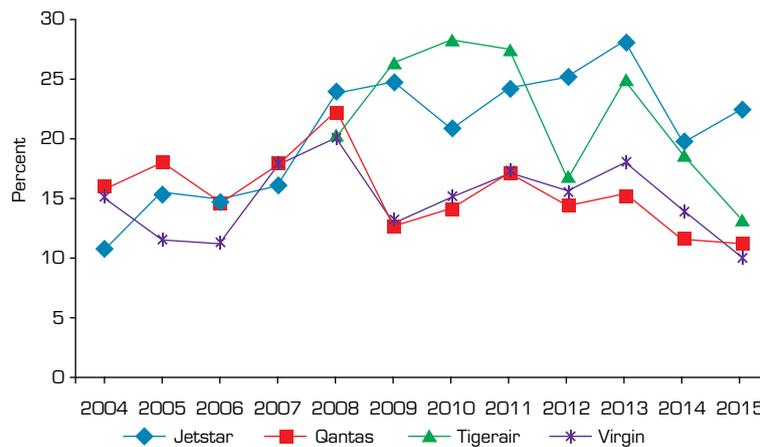

Source: Mohammadian *et al.* (2019).

Figure 1. Delay periods from 2004-2015.

During 2004-2015, the airlines had the largest and lowest average flight delays when measured against their rates of occurrence. Delays are able to create financial losses. According to a study released in April of this year, flight delays in the U.S. aviation industry would exceed $30 billion this year, with carriers owing $8.7 billion. A 10% decrease in air traffic delays would result in $17.6 billion in overall US welfare gains. According to the newest Value Penguin study, in America, 20.1% of flights were delayed by more than 15 minutes from January to May 2022. In the first quarter of 2022, the average airfare was $328.49 and inflation rates were adjusted in dollars in the first quarter of 2021. In 2020, 10.2 % of flights were delayed from January to May, 18.9% of flights delays happened in 2019. The time spent on flight delays costs airlines extra money since it involves using people, fuel, and maintenance on planes. Several prior researches focused on identifying permanent patterns in air traffic management. Rajarajeswari *et al.* (2022) focused on Machine Learning concepts and Mathematical modeling techniques in 2022. In their paper "Semi-Supervised Learning for Classifying

Groups of Similar Days", Liu *et al.* (2014) presented a semi-supervised learning technique that can differentiate between similar days groups. To identify the days with the lowest total distance, the first step was to compute the distance between hourly weather predictions.

According to empirical research, individuals have a significant propensity to avoid delays (Vieira *et al.* 2016). Two competing factors affect the amount of time passengers experience at a hub airport (i.e., airports with one or a few major carriers). Both on the one hand and on the other, hubs are strived to provide their clients with as many connection options as possible, but they are constrained by airport capacity. Since there is a trade-off between growing the number of connections and the increase in marginal costs (delays and connecting times), there is also a limit to the total number of connections that may be added. Airlines that have hub operations, on the other hand, are able to flexibly choose their departure and arrival dates, enabling them to partially offset the increase in congestion (Dorigo and Gambardella 1997). Graham (2019) has prepared a technical review on airport management techniques. Electric aircraft design was discussed by Nagy (2019).

Figure 2 shows the methods of flight delay prediction.
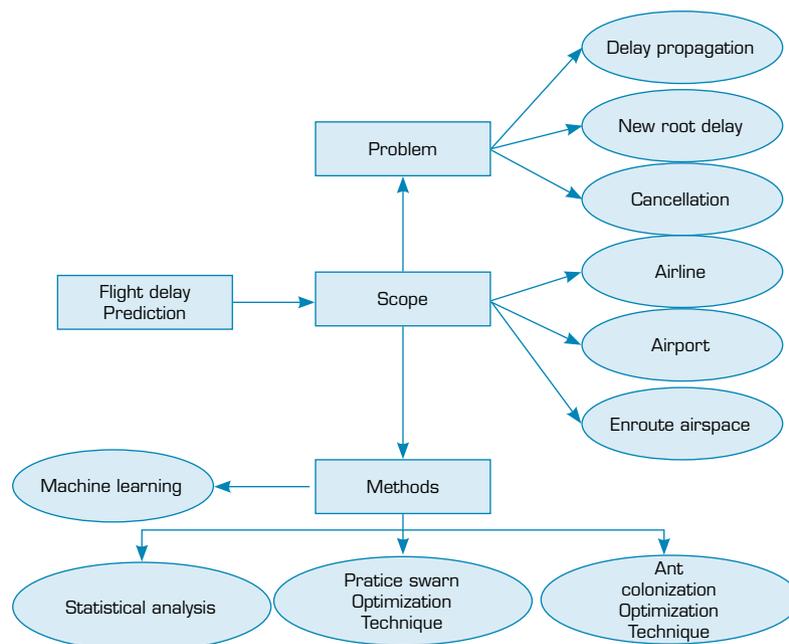
Source: Elaborated by the authors.

Figure 2. Flight delay Prediction Methods.

Analytical statistics

Also employed in statistical models, there are correlation analysis, parametric and non-parametric tests, multivariate analysis, and economic models. Government agencies have built econometric models to better understand the connection between delay, passenger demand, fare, aircraft size, and other variables.

Probabilistic models

Probabilistic models use analytical techniques that calculate the probability of an event by using facts from the past. A probability distribution is given for the estimated outcome. Every model includes randomness, and unpredictability has an impact on the result of the model.

Machine Learning

Supervised Machine Learning is the process of identifying the input and output data for a given dataset and then using a variety of Machine Learning algorithms to assign new data to those positions. In this case, it is important to anticipate a flight delay (Rebollo and Balakrishnan 2014).

Simulation based Predictive analysis of Indian Airport transportation system using Computational intelligence techniques

5

Computational Intelligence Techniques

Using Machine Learning models, computer-intelligent techniques play a significant role in time-domain prediction optimization. Computational Intelligence Techniques are used for multiple optimization and estimation methods for doing adaptive control mechanisms. Machine Learning techniques are used to generate unpredictable traffic condition issues (Graham 2019). This methodology is based on particle swarm optimization and ant colonization optimization methods.

PROPOSED METHODOLOGY FOR FLIGHT DELAY PREDICTION SYSTEM

During the first four months of this year, airline companies were forced to pay passengers compensation totaling over Rs. 25 crores for various difficulties. The findings here may serve as a basis for identification of operational factors that cause delays in any country scenario. The recognized and structured taxonomy of known problems associated with aircraft delay prediction may be found in this document. It addresses the root cause of the flight delay, as well as the effects that it has on the relevant institutions and methods to handle flight delay prediction problems. It takes into consideration flight domain (issue and scope) choices. Delays, including flight cancellation, are caused by many factors, including delay propagation, delay created at the departure point, and flight cancellation. Operators and administrators will be able to prepare for a seamless operation (delay prediction tool helps here). Flight delays affect all sectors, which are independent entities that operate in harmony. In order to create a system that predicts flight delays, Machine Learning, Probabilistic models, Statistical analysis and network representations may be used. Considering the traditional taxonomy of flight delays and related problems, the breadth of the predictions may be considered as a combination of different variables. For the prediction of delays at airports, the models are built as components of the system might be employed. Such forecasting abilities would simplify the planning of travel mitigation methods for traffic management and airline dispatcher. This challenge may be resolved by utilizing the following approaches to create a tool to forecast aircraft delays.

Computational Intelligence Techniques for prediction optimization

Using Machine Learning models, computer-intelligent techniques play a significant role in time-domain prediction optimization. This methodology is based on particle swarm optimization and ant colonization optimization methods.

Particle Swarm Optimization

The best-fit particle in the whole swarm affects the location of each particle in the PSO. A star social network architecture is used because all the social information that is known about the swarm's particles (Carlisle and Dozier 2001) is taken into consideration. In this approach, each individual particle has equal importance, a current position in search space, a current velocity and a personal best position in search space. The ideal posture for everyone refers to the search location of the particle in view of a reduction issue, the lowest value defined by the objective function F. Furthermore, the lowest value position among all the best of the personal is referred to as global best position. This example included our use of three distinct models: Decision Trees, Logistic Regression and Neural Network.

For accuracy and efficiency, we used a particle swarm optimization method to differentiate amongst airline service providers. Equations 1 and 2 are both equations that represent how personal and global best values are updated. Time domain iteration is described by N.

Kennedy and Eberhart (1995) introduced PSO techniques. Fielding and Zhang (2018) and Rana *et al.* (2011) have also described PSO with different optimization variants. Sedighizadeh and Masehian (2009) focused on minimal issues and calculated best position vector of Particle at time t+1 by using the following equations:

$$P_{best,i}^{t+1} = \{P_{best,i}^{t+1} \; if \; f(x_i^{t+1}) > P_{best,i}^{t+1} \; x_i^{t+1} \; if \; f(x_i^{t+1}) > P_{best,i}^{t+1}$$

$$\text{if } f(x_i^{(t+1)}) > P_{best,\, i}^{t+1} \tag{1}$$

Where is the fitness function? The global best position at time step is calculated as:

$$G_{best} = \{P_{best,i}^{t+1}\}, where\ i \in [1, ..., n]\ and\ n > 1 \tag{2}$$

Since personal best is defined as the best location for the individual, the particle is essential to keep in mind that the PB is the best position that is the individual particle. Since it was considered as the first-time step. When we look at the rest of the swarm, it was found by a single particle and was thus the best location in the whole swarm.

For PSO method, the velocity of particle i is calculated by (Eq. 3)

$$v_{ij}^{t+1} = v_{ij}^{t} + c_1 r_{1j}^{t} [P_{best,i}^{t} - x_{ij}^{t}] + c_2 r_{2j}^{t} [G_{best} - x_{ij}^{t}] \tag{3}$$

Figure 3 shows the proposed algorithm:
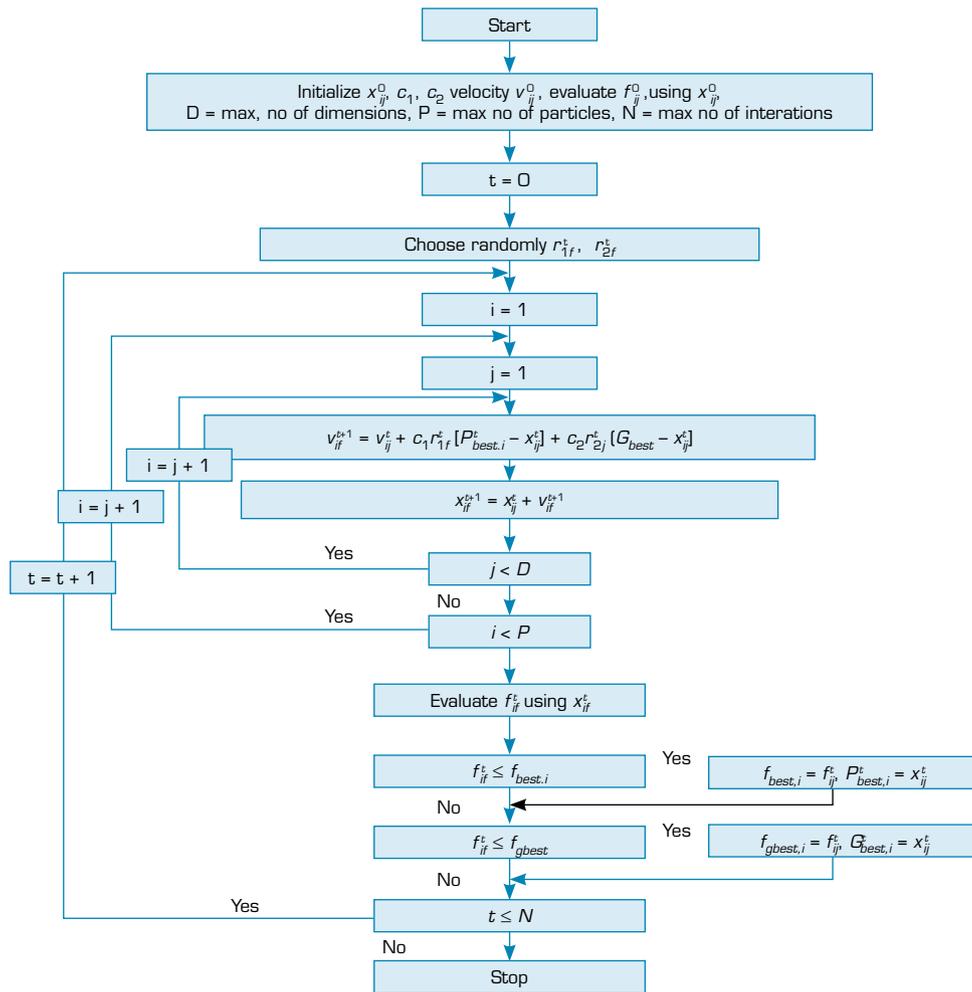
Source: Elaborated by the authors.

Figure 3. Proposed Algorithm.

Ant Colonization Optimization

An artificial ant is a basic calculator that looks for an excellent solution to a certain optimizing issue in the ant colony optimization algorithms. The optimization problem must be transformed to the issue of finding the shortest route on a wet network in order to use an anti-colonial method. In each initial step, each ant builds a solution stochastically, i.e. the order to be followed by the edges of a graph. The routes of the several ants were compared in the second step (Robinson 1989). The last stage involves the update

Simulation based Predictive analysis of Indian Airport transportation system using Computational intelligence techniques

7

of each edge's pheromone levels. Here we are considering the solutions for predicting the flight delay where the following ACO procedure is optimized as follows:

Ant Colonization Optimization_Meta Heuristic is

while not completed do

Generate Solutions ( )

Daemon Actions ( )

Pheromone Update ( )

Repeat the process

End

Edge selection

In our case, the precision model is taken from Decision Trees, Logistic Regression and Neural Network were optimized using Ant Colonization Optimization technique to differentiate different airlines service providers for accuracy and efficiency model for flight delay prediction. Each ant needs to construct a solution to move through the graph of data of precision. When choosing the next edge of its tour, an ant takes into account the length and matching pheromone levels of each edge from its present position. Each ant travels from state x to state y at each step of the algorithm, which corresponds to a more complete intermediate solution. Thus, in each cycle, each ant k calculates a set of viable expansions to its present condition and will likely migrate to one of them. Ant k is subject to the combination of two values, the course attracting value of the move, which is calculated in some heuristic way. It is showing the priori desire of this movement and the trail level of the movement, indicating the competence in making that movement in the past. The level of the trail provides a retrospective indication of this change.

Generally, the ants travels with probability from state x to state y

$$p_{xy}^k = \frac{(\tau_{xy}^\alpha)(\eta_{xy}^\beta)}{\sum_{z \in allowed_x} (\tau_{xy}^\alpha)(\eta_{xy}^\beta)} \tag{4}$$

where, pheromone amount is deposited for transition from state is a parameter to control the influence of. It can be desired as the state transition and is a parameter which can be used for controlling the influence and representing the trail level and attractiveness for the other possible state transitions. The pheromone deposit function which gets updated for good and bad situations depends on similarity for delay impossibility or delay possibility. As the functions are inversed for ant colony algorithm, the pheromone update function is defined.

Pheromone update

Trails are updated after all ants have solved the problem, changing the trail difficulty to match the solution's class. An example of a global pheromone updating rule is

$$\tau_{xy} \leftarrow (1 - \rho)\tau_{xy} + \sum_k^m \Delta\tau_{xy}^k \tag{5}$$

where, the quantity of pheromone deposited for a state change is determined by the cost of transition. The pheromone evaporation coefficient, the number of ants and the amount of pheromone deposited by the ant is used to represent a TSP (movement of the graph correspond to arcs) issue.

$$\Delta\tau_{xy}^k = \{\frac{Q}{L_k} \qquad 0 \ if \ ant \ k \ uses \ curve \ xy \ in \ its \ tour \tag{6}$$

where it is the cost of the ant's tour (typically length) and is a constant. Each kth tour defines the analysis for different flight carrier prediction of flight delay.
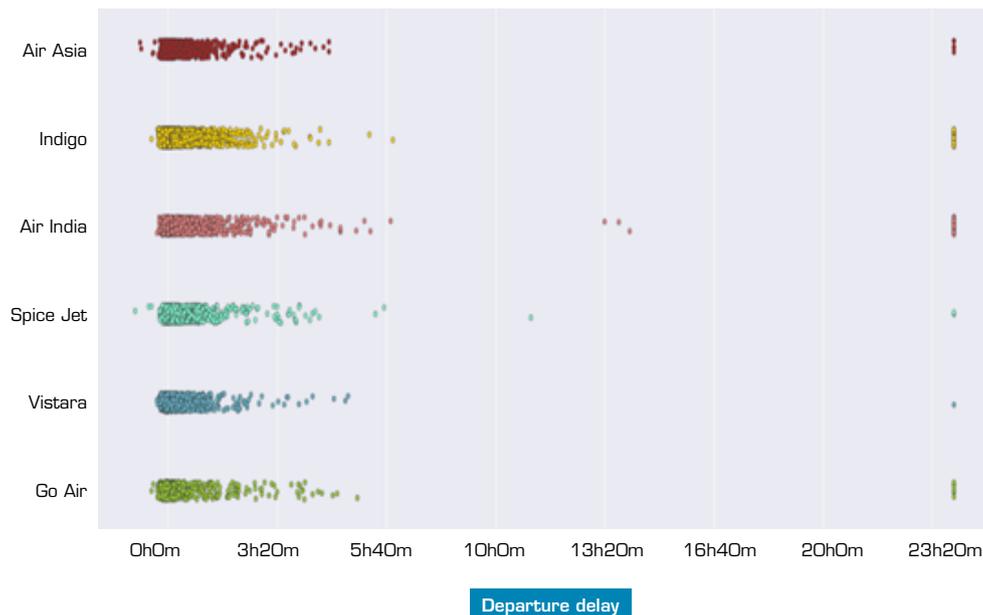
RESULTS AND DISCUSSION

Flight Dataset description

    The dataset is taken for six top airlines operational in India. There are a total of 3 lakh points. The dataset has 29 rows like Used Date, From, To, Airline, Scheduled Departure, SDEP, Departure DEP, Scheduled Arrival, SARR, Arrival ARR, Departure Delay, Arrival Delay, Status Distance, Passenger Load Factor, Airline Rating, Airport Rating, Market Share, OTP Index, Weather Hourly Wind Speed Kmph, Weather Hourly Weather, Desc Value, Weather Hourly Precip (MM), Weather Hourly Humidity, Weather Hourly Visibility, Weather Hourly Pressure, Weather Hourly Cloud, Cover, Category (Abdelmonaem *et al.* 2019). The dataset was between Jan 2019 to Jan 2020. Figures 2–6 shows different statistical analyses plot. As per the Fig. 3 and Fig. 5 which are departure, delay and mean delay respectively, we have observed that Air India tops the delay in every prospect, which is followed by SpiceJet, Go-Air, Indigo, Vistara and AirAsia. As the flight capacity is less for Vistara and AirAsia the pattern of less delay is acceptable but at the cost of machine learning analysis. Figure 7 shows the mean delay period which is shown below.
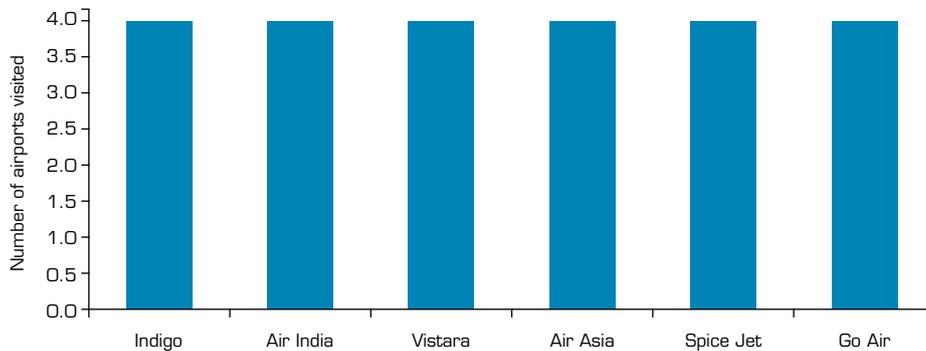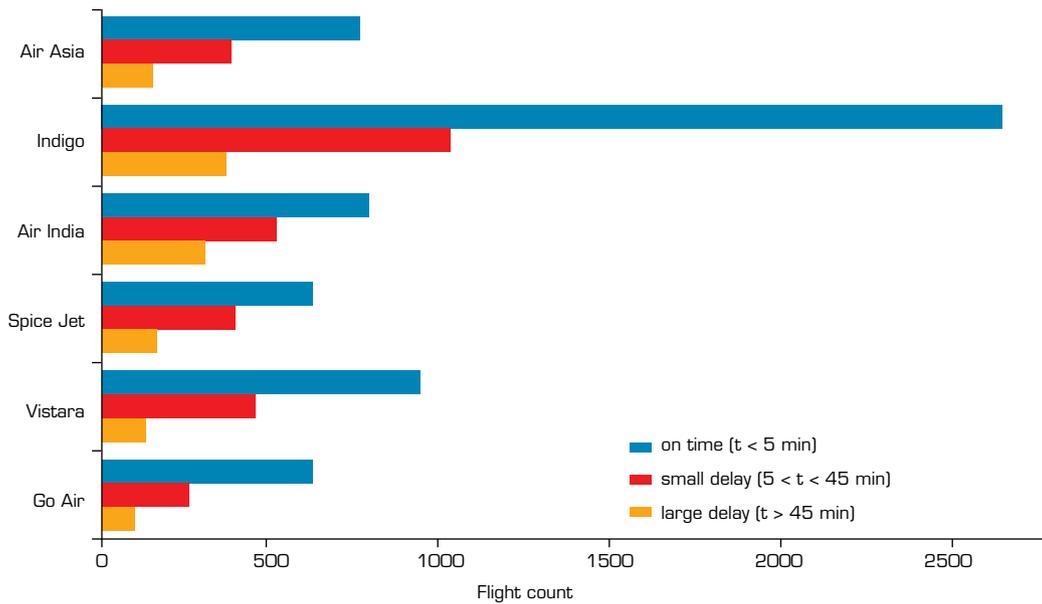

Source: Elaborated by the authors.
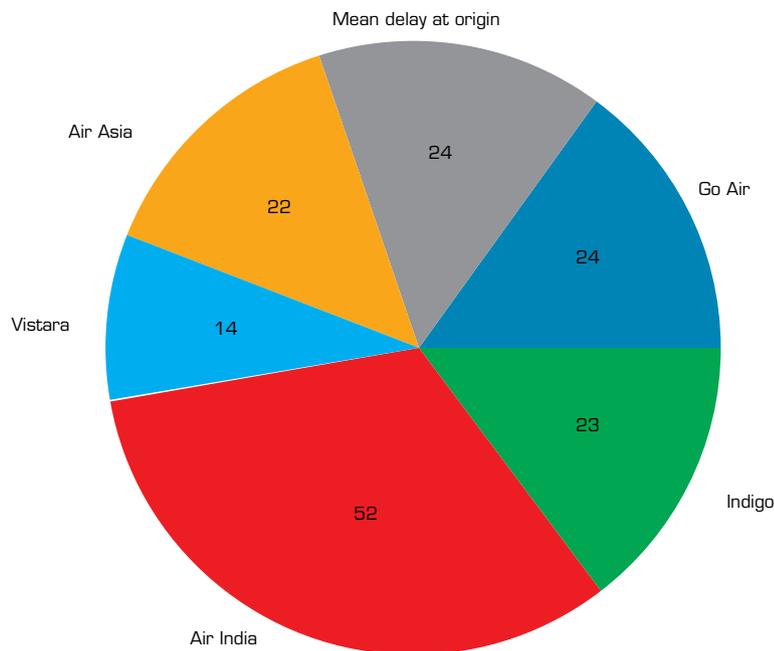
Figure 4. Visited Airports.

Source: Elaborated by the authors.

Figure 5. Departure Delay.

Simulation based Predictive analysis of Indian Airport transportation system using Computational intelligence techniques

9

Source: Elaborated by the authors.

Figure 6. Flight Count.

Source: Elaborated by the authors.

Figure 7. Mean Delay.

Naïve Bayes Classification

The classifications of Naive Bayes fall under the family of direct "likely" classifiers in the field of Machine Learning, which are supposed tobe independent of one another in the Naïve Bayes Classification. In the 1950s Naïve Bayes were more studied than ever, and since that time, in the 1960s the community had a special name in the problem of the retrieval and the resolution of texts, and it still is considered as a root method for categorization, judgment documents as they become one class or the other

(such as spam or legitimate, sporting or political), etc. After the threshold and regulations have been given in advance, this part is competitive in many sophisticated ways and supports vector machines. It also discovers automated diagnostics application.

Naïve classification systems from Bayes are extremely scalable. The evaluations of a closed-form expression that takes linear time rather than a repetitive cost approximation, as used for several alternative forms of classification, will make it necessary for a variety of maximum-liking training programs. Naïve Byeas models, along with easy Bayes and independence Bayes, are most well known in statistics and technical literature under the range of titles to use a theorem in the ruling classifier, although ingenuous Bayes are not (necessarily) a theorem technique (D'Ariano *et al.*2012).

Bayesian Network (BN) Algorithm

Bayesian Network is a mathematical learning method which shows the links between variables. A directed acyclic graph consisting of nodules and edges may easily be called network. Applied math is supported. Once a best BN is built, it conducts the probabilistic classification job. (Fenton and Neil 2007; Ketha and Imambi 2019).

The classification using Naive Bayes was unexpectedly successful in attaining error rates between 15 and 18 percent. However, deeper research by the classifiers revealed also that they predicted mainly that the aircraft would not be delayed. Roughly 75% of the remainder was precise, being 37.32% from foreign flights to 7.6% on a whole dataset (Fenton and Neil 2007; Ketha and Imambi 2019). Included in the cloud computing and smartphone analysis, weather forecasting and forecasting may be utilized for delayed forecasts on flights (Hopfield 1982; Quinlan 1987).

Neural Network Integration

To build the system that anticipates flight delay, several approaches are used. Only a few of these methods are covered in Fig. 8.
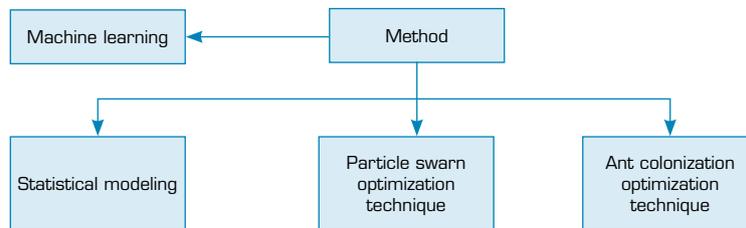
Source: Elaborated by the authors.

Figure 8. Neural Network Integration.

Decision Tree

The algorithm's fundamental premise is to use a tree-like structure to generate responses that are either true or incorrect. To end with an option in the model, start with a root node. Yes is assigned to each node. The information is passed on to the next node without inquiry and answer exchanged. All of the training dataset's data is sent down to the root node. Deciding what questions to ask and when to ask them is a tough task when planning to place a tall tree. Algorithmic software that makes decisions based on metrics like entropy or Gini-impurity uses well-known metrics to measure a particular node's level of uncertainty or impurity. To understand the relationship between the Gini impurity and entropy, see Eq. 4 and 5, respectively. The idea for flight delay analysis using decision tree was carried out for defining the uncertainty or impurity associated for dependability of airlines dataset.

$$H(s) = -\sum_{c \in C} \quad p(c)logp(c) \qquad\qquad H(s) = -\sum_{c \in C} \quad p(c)^2 \qquad\qquad (7)$$

Within the equations, C represents the number of different classes which represent the dependability of each variable in rows parameter for time domain prediction.

Logistic Regression

Logistic Regression is an algorithm that performs classification using $h_\theta(x) = g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}}$, and this, in turn, indicates the highest probability of estimate and gradient ascent possible once the variable has been split, the logistic regression method

is the most appropriate multivariate analysis to use (binary). The logistical regression analysis, like other regression studies, is a predictive analysis of some kind. Logical regression is used to interpret data and to explain the connection between a single dependent binary variable and one or more independent variables with nominal, ordinal, interval, or ratio levels.

Neural Networks

A neural network is constructed by stacking together numerous neurons in layers in order to get a final result. It is the input layer that is the first layer, and the output layer is the final layer that is created. The term "hidden layers" refers to all of the layers in between. Every nerve cell has the ability to activate other nerve cells. Sigmoid, ReLU, Tanh, and a variety of other activation functions are among the most common. The weights and biases of each layer serve as the network's parameters, which are described below. The objective of the neural network is to find the network parameters that have been given, and the anticipated result is that the network parameters are equal to the ground truth. For the purpose of determining network parameters, back-propagation on a loss-function is used (Eberhart and Shi 2000), (Carlisle and Dozier 2001). Table 1 shows the precision values for all three algorithms.

Table 1. Classification Reports for Classifiers Using Decision Trees, Logistic Regression, and Neural Network (Values in %).

Algorithm	Decision Tree	Logistic Regression	Neural Network
Precision	91	90	92

Source: Elaborated by the authors.

Networks reports are generated using 3 lakh test samples, which is a large number (Awasthi and Seth 2018).

Comparative analysis of accuracy for Airport model by using various Machine Learning models: we have given comparison of accuracy with respective all models which are given in Fig. 9. We have used Python programming for getting predictive analysis of the Airport model system. We have used Particle Swarm optimized Machine Learning model and Ant Colonization Optimized Machine Learning modeling techniques for predicting the system.
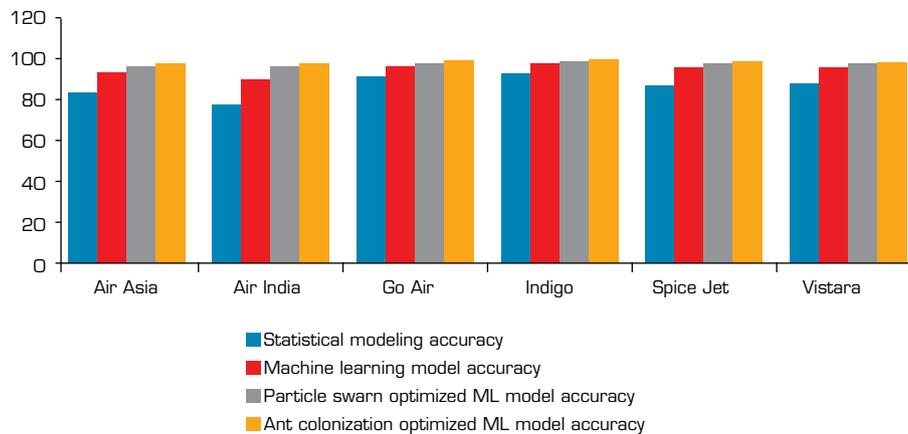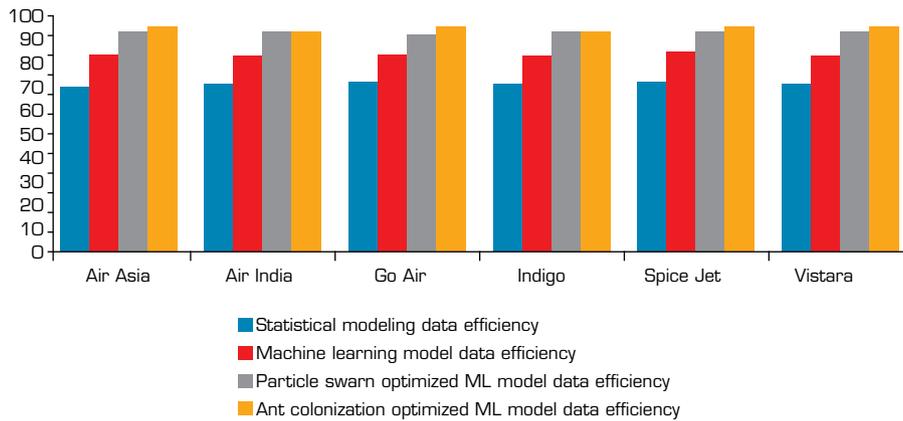

Source: Elaborated by the authors.

Figure 9. Comparison analysis on Accuracy for Flight Delay Prediction.

Figure 9 shows the accuracy chart for flight delay prediction. It consists of Statistical Modeling, Machine Learning, Particle Swarm optimized ML model and Ant colonization optimized. We have made comparisons between all types of models.

As shown in Fig. 8, the accuracy chart for Airlines was defined for four statistical models: Modeling Accuracy, Machine Learning Model Accuracy, Particle Swarm Optimized ML Model Accuracy and Ant Colonization Optimized ML Model Accuracy. It was observed that Indigo accuracy was achieved maximum up-to 99.5% considering the huge data validation and lowest was Air Asia due to less data validation (Shahriari and Aydin 2018). The model was heavily predicting the huge delay for future prediction in Air India followed by SpiceJet, Air Asia, Vistara, Go Air and Indigo (Fig. 10).

Source: Elaborated by the authors.

Figure 10. Data Efficiency Model for Different Airlines.

Considering Fig. 9 the efficiency or dependability of data was considered for the validation analysis and it was found out that 98% above was achieved for all the airline services.
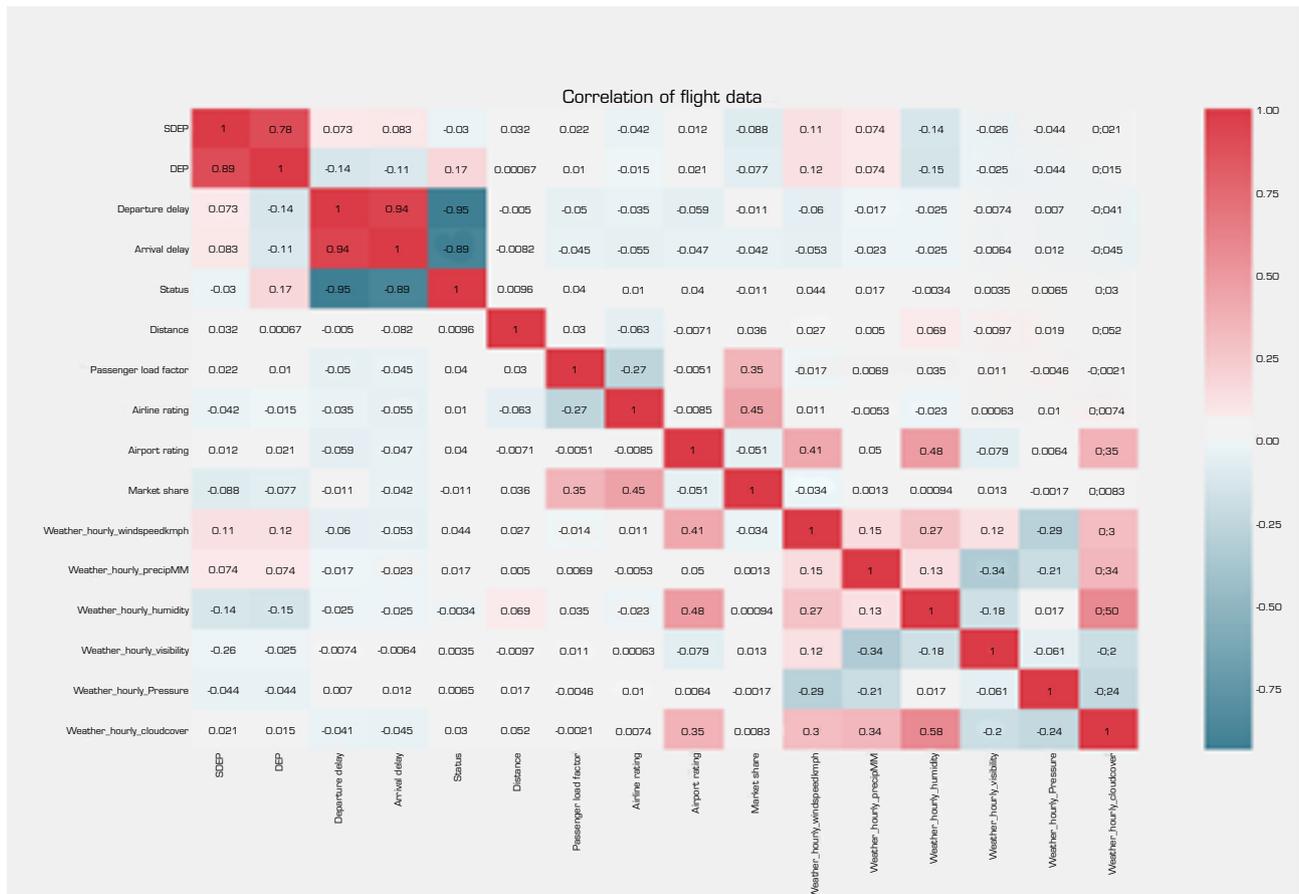
Source: Elaborated by the authors.

Figure 11. Correlation matrix for flight data.

As observed in Fig. 11 the correlation matrix was developed for various parameters like Scheduled Departure, SDEP, Departure DEP, Scheduled Arrival, Departure Delay, Arrival Delay, Status Distance, Passenger Load Factor, Airline Rating, Airport Rating, Market Share and weather conditions dataset. It was also discovered that the rainy season has the highest weather dependability,

with a 75 percent chance of flight delays. As the correlation matrix displays higher accuracy, the operational parameter demonstrates direct dependability of up to 87 percent.

CONCLUSION AND FUTURE WORK

The analysis also shows the dependability of various parameters like Scheduled Departure, SDEP, Departure DEP, Scheduled Arrival, Departure Delay, Arrival Delay, Status Distance, Passenger Load Factor, Airline Rating, Airport Rating, Market Share and weather conditions. It was also observed that the rainy season has dependability of weather with 75% for flight delay. Operational parameter shows direct dependability up to 87% as correlation matrix shows greater precision. The accuracy, precision and efficiency model for the airlines has been achieved up to 90% and above which shows great confidence in our prediction model for flight delay characteristics. The model was strongly forecasting a significant delay in Air India's future predictions, which was followed by SpiceJet, Air Asia, Vistara, Go Air, and Indigo, among others. For the validation study, the efficiency or dependability of data was taken into consideration in accordance with the result and it was discovered that a success rate of 98 percent or above was obtained for all air-line services. There is a huge correlation between the weather parameters which directly affects the delay in flights. Weather parameter as a function of time can be undertaken for creating the validate Machine Learning model as the weather data characteristics on the way path between source and destination of the flight. Such modeling can give us accurate result of prediction touching the accuracy and precision model more than 99%. In this research paper, we have introduced a new proposed methodology for the airport system model. We have described different computational intelligence techniques for modeling the aircraft system. We have applied Practical Swarm intelligence and Ant Colonization Optimization methods for predicting the aircraft traveling process. We have introduced Simulation and Machine Learning methods for determining airport delay periods. This system is implemented using a Python tool. In this paper, we have presented literature on flight delay prediction based on a systematic mapping study and comparative analysis of accuracy to predict airport travel system models by using various Machine Learning models. In addition, we will focus on a sensitivity analysis to investigate the impact on reduction of runway capacity and not including passenger inconvenience costs. In the future, we can apply formal methods for optimizing the flight routes to reduce the delay periods. We will be able to get more analysis of flight prediction system based on formal methods. In these formal methods, we will use semantic rules for calculating the delay periods in between the source to destination places.

CONFLICT OF INTEREST

No conflict of interest.

AUTHORS' CONTRIBUTION

Conceptualization: Kumari PLS; **Methodology:** Rajarajeswari P; **Software:** Rajarajeswari P; **Validation:** Mugunthan SR; **Formal analysis:** Kumar PR; **Investigation:** Rajarajeswari P and Kumari PLS; **Resources:** Lakshmi AS; **Data Curation:** Ananthula PR; **Writing – Original Draft:** Rajarajeswari P; **Writing – Review & Editing:** Rajarajeswari P; Kumari PLS; Lakshmi AS; Mugunthan SR; Kumar PR and Ananthula PR; **Visualization:** Rajarajeswari P; Kumari PLS; Lakshmi AS; Mugunthan SR; Kumar PR and Ananthula PR; **Supervision:** Rajarajeswari P.

DATA AVAILABILITY STATEMENT

The data will be available upon request.

FUNDING

ACKNOWLEDGEMENTS

REFERENCES

Abdel-Aty M, Lee C, Bai Y, Li X, Michalak M (2007) Detecting periodic patterns of arrival delay. J Air Transp Manag 13(6):355-361. https://doi.org/10.1016/j.jairtraman.2007.06.002

Abdelmonaem AMA, Gawad AFA, Nassief MM (2019) Aerodynamic CFD Simulation Study of a Commercial Aircraft Model. Paper presented at: 4th IUGRC International Undergraduate Research Conference, Military Technical College; Cairo, Egypt.

Awasthi A, Seth K (2018) Tourism and Hospitality Innovations, Opportunities and Challenges. New Delhi: Bharti Publications. Chapter 8, Aviation Industry: Issues & Challenges; p. 91-100.

Carlisle A, Dozier G (2001) An Off-The-Shelf PSO. Paper presented at: Proceedings of the Particle Swarm Optimization Workshop; Indianapolis, United States.

D'Ariano A, Pistelli M, Pacciarelli D (2012) Aircraft retiming and rerouting in vicinity of airports. IET Intell Transp Syst 6(4):433-443. https://doi.org/10.1049/iet-its.2011.0182

[DIRD] Department of Infrastructure and Regional Development (2022) Wikipedia [Internet]. [accessed 2022 Apr 20]. https://en.wikipedia.org/wiki/Department_of_Infrastructure_and_Regional_Development

Dorigo M, Gambardella LM (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans Evol Comput 1(1):53-66. https://doi.org/10.1109/4235.585892

Eberhart RC, Shi Y (2000) Comparing inertia weights and constriction factors in particle swarm optimization. Paper presented at: Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512); La Jolla, United States. https://doi.org/10.1109/CEC.2000.870279

Fenton N, Neil M (2007) Managing Risk in the Modern World: Applications of Bayesian Networks. London: London Mathematical Society.

Fielding B, Zhang L (2018) Evolving Image Classification Architectures With Enhanced Particle Swarm Optimisation. IEEE Access 6:68560-68575. https://doi.org/10.1109/ACCESS.2018.2880416

Graham A (2019) Airport management: a perspective article. Tour Rev 75(1):102-108. https://doi.org/10.1108/TR-05-2019-0200

Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. PNAS 79(8):2554-2558. https://doi.org/10.1073/pnas.79.8.2554

Kennedy J, Eberhart R (1995) Particle swarm optimization. Paper presented at: Proceedings of ICNN'95 - International Conference on Neural Networks; Perth, Australia. https://doi.org/10.1109/ICNN.1995.488968

Simulation based Predictive analysis of Indian Airport transportation system using Computational intelligence techniques

15

Ketha T, Imambi SS (2019) Analysis of Road Accidents to Indentify Major Causes and Influencing Factors of Accidents – A Machine Learning Approach. IJATCSE 8(6):3492-3497. https://doi.org/10.30534/ijatcse/2019/127862019

Lawson D, Castillo W (2012) Predicting flight delays. In: Technical report, Computer Science Department, CS 229. Stanford: Stanford University.

Liu Y, Seelhorst M, Pozdnukhov A, Hansen M, Ball MO (2014) Assessing terminal weather forecast similarity for strategic air traffic management. Paper presented at: International Conference on Research in Air Transportation; Istanbul, Turkey.

Nagy A (2019) Electric aircraft - present and future. Prod Eng Arch 23(23):36-40. https://doi.org/10.30657/pea.2019.23.06

Mohammadian I, Abbasi B, Abareshi A, Goh M (2019) Antecedents of flight delays in the Australian domestic aviation market. TRIP 1:100007.

Quinlan JR (1987). Simplifying decision trees. Int J Man-Mach 27(3):221-234. https://doi.org/10.1016/S0020-7373(87)80053-6

Rajarajeswari Perepi, Santhi K, Saraswathi R, Bég OA (2022) Predictive analysis of COVID 19 disease based on mathematical modelling and machine learning techniques. Computer Methods in Biomechanical and Biomedical Engineering Imagingand Visualization. https://doi.org/10.1080/21681163.2022.2120829

Rana S, Jasola S, Kumar R (2011) A review on particle swarm optimization algorithms and their applications to data clustering. Artif Intell Rev 35(3):211-222. https://doi.org/10.1007/s10462-010-9191-9

Rebollo JJ, Balakrishnan H (2014) Characterization and prediction of air traffic delays. Transp Res Part C Emerg Technol 44:231-241. https://doi.org/10.1016/j.trc.2014.04.007

Reynolds-Feighan AJ, Button KJ (1999). An assessment of the capacity and congestion levels at European airports. J Air Transp Manag 5(3):113-134. https://doi.org/10.1016/S0969-6997(99)00006-X

Robinson PJ (1989) The Influence of Weather on Flight Operations at the Atlanta Hartsfield International Airport. WAF 4(4):461-468. https://doi.org/10.1175/1520-0434(1989)004%3C0461:TIOWOF%3E2.0.CO;2

Schaefer L, Millner D (2001) Flight delay propagation analysis with the Detailed Policy Assessment Tool. Paper presented at: 2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat.No.01CH37236); Tucson, United States. https://doi.org/10.1109/ICSMC.2001.973100

Sedighizadeh D, Masehian E (2009) Particle Swarm Optimization Methods, Taxonomy and Applications. IJCTE 1(5):486-502. https://doi.org/10.7763/IJCTE.2009.V1.80

Shahriari M, Aydin ME (2018) Aviation Accident Analysis: A Case Study. In: Arezes P, editor. Advances in Safety Management and Human Factors. AHFE 2017. Advances in Intelligent Systems and Computing. Cham: Springer. https://doi.org/10.1007/978-3-319-60525-8_9

Vieira DR, Rebaiaia M-L, Chain MC (2016) The Application of Reliability Methods for Aircraft Design Project Management. Am J Ind Bus Manag 6(9):967-992. https://doi.org/10.4236/ajibm.2016.69094

Wong J-T, Li S-L, Gillingwater D (2002) An Optimization Model for Assessing Flight Technical Delay. Transp Plan Technol 25(2):121-153. https://doi.org/10.1080/03081060290033212