2025 Global Conference on Information Technology and Communication Networks (GITCON) | 979-8-3315-3348-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/GITCON65266.2025.11376823

Enhancing Flight Departure Delay Predictions Due to Weather Conditions Using LSTM and XGBoost Models

Dr. Krishna Annaboina,
Assistant Professor,
Department of CSE(IOT),
Guru Nanak Institutions Technical Campus,
Hyderabad, Telangana, India,
anneboina.krishna@gmail.com

Dr. Pradosh Kumar Sharma,
Head, Department of Physics, Chinmaya Degree College BHEL
Haridwar, Uttrakhand,India,
drpksharma25@gmail.com

Bandla Ramesh,
Department of Computer Science and Engineering,
Dhulapally, Telangana, India,
bandlaramesh.cse@gmail.com

P. Joel Josephson,
Associate Professor.
Malla Reddy (MR) Deemed to be University
Maisammaguda Dhulapally Secunderabad
Telangana India 500100
, joeljosephsonp@gmail.com

Huma Qamar Khan,
Department of Computer Science Engineering,
Joginpally B R Engineering College,
Moinabad, Hyderabad, Telangana, India,
humaqk@gmail.com

U D Prasan,
Department of Computer Science and Engineering, Aditya Institute of Technology and Management,
Tekkali, Andhra Pradesh, India,
udprasanna@gmail.com

***Abstract* –Regrettably, the current infrastructure and operating capabilities are insufficient to accommodate the escalating demand for air travel, which is gaining popularity. Aircraft delays have increased in frequency due to the erratic weather patterns caused by climate change. Extensive study has been conducted on short-term forecasting; nevertheless, due to the intricacies of international flight networks, there is an increasing demand for long-term flight delay prediction and analysis. This work proposes a comprehensive technique that encompasses data preparation, feature selection, and model training, thereby fulfilling the criteria. The preprocessing phase entails data cleansing and quality assurance, while the Boruta approach is employed to identify pertinent attributes. XGBoost enhances feature selection by retaining features that exceed a specified relevance threshold. A hybrid model, LSTM-XGBoost, is proposed to improve performance. It integrates the optimal threshold with a tailored loss function. The model attained a binary flight delay prediction accuracy of 95.57%, with a RMSE of 0.35, a MAPE of 3.71, and a MAE of 0.2, based on actual flight data evaluation. The results indicate that the LSTM-XGBoost model effectively enhances long-term flight delay predictions, hence optimising aircraft operations amid increasing air traffic and climate-related disturbances.**

Keywords—National Airspace System (NAS), Information Gain (IG), Naïve Bayes Classifier (NBC), Bidirectional Long-Shot Term Memory (BiLSTM), Random Forest Classifier (RFC).

I. INTRODUCTION

The flight delay prediction studies in the past most of those studies have focused on short timeframes like or up to and have mostly been applied to airline services. International flights traverse great distances across seas and continents and their durations range from 10 to 20 hours. This highlights the need for delay prediction over longer timeframes. The capacity to forecast aviation delays over lengthy time periods using input data opens up new possibilities for long-haul flights and different flight itineraries which is a practical use of such models. Flight resource management is only one area that stands to gain from this enhanced capability are many other uses [1]. Uncertainties abound in the aviation industry which is characterized by its dynamism, competitiveness and volatility. These uncertainties include flight delays caused by a variety of circumstances including bad weather technical difficulties operational delays and crew-related issues. When a flight's arrival time is more than fifteen minutes later than the intended arrival time in relation to the departure time the flight is viewed as delayed according to the Federal Aviation Administration. When flights and the economy are able to anticipate when flights will be delayed everyone wins [2]. By accurately predicting when delays will occur flights can take proactive measures to reduce or avoid them. Customer satisfaction and loyalty can raise as a result of improved on-time performance more effective use of resources and streamlined operations. Increased income and market share might result from passengers choosing an airline that reliably offers on-time flights and a positive experience.

Weather the flight takeoff or landing management, airline management, air traffic, air traffic control passenger reasons and a host of other factors make it more difficult to pinpoint exactly what causes flight delays these days [3]. Flight security operations and resource scheduling will be further burdened as a consequence of aircraft delays which will upset the allocation plans for limited airport resources including runways, aprons and routes and so on. Flight operating, maintenance and human resource expenses may rise due to aircraft delays cutting into profitability. If a passenger's flight is delayed it can ruin their vacation plans or business trip plans. Insurers rely heavily on flight delay prediction data to set premiums and run their flight delay and travel insurance policies. These issues can be ameliorated with the use of this categorization prediction of flight delays. Flight schedules that are likely to experience delays by analyzing huge data using a variety of techniques to avoid these issues [4]. The private airline's enormous data set includes every flight that had place that year. There is flight data and then there is meteorological data for the time period. Flight decision-making processes and ground operations have been examined from numerous angles due to the topic's crucial role in air traffic control. In conducted the initial research [5]. Using weather data collected between domestic flights data is used to forecast flight delays caused by weather.

The remaining sections of the document are organised as follows: Part 2 provides an overview of the Litrature Survey.

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.

In Section 3, a concise summary of prior research on flight delay prediction using machine learning techniques is provided.The recommended approach to this research.Section 4 follows with the predictive model's experimental outcomes. Section 5 concludes the work and makes recommendations for further reading.

II. LITERATURE SURVEY

Probabilistic models and statistical analysis are the go-to tools for forecasting flight delays. Nevertheless found that these models had difficulty handling high-dimensional data and extracting non-linear relationships. In order to anticipate flight delays it suggested a CNN-LSTM and DL system [6]. There are primarily three parts to the suggested CNN-LSTM model: For the purpose of predicting flight delays a CNN an LSTM network and finally a RFC are utilized. To begin the spatial adjustments between regions are extracted using a CNN framework [7]. The LSTM network is then used to describe temporal dynamics based on the output of the CNN algorithm. Numerous ML models are investigated in this study such as LR, NB, NN, RF, XGBoost, CatBoost and LightGBM [8]. Adding meteorological data to further tests using the SMOTE helps fix the data imbalance problem. This comprehensive method guarantees accurate forecasting in different environments [9]. The majority of the studies that were reviewed used RF as their method of choice. Nearly as popular are GB, ANN, and DT.

The ML which DL is a subset allows computers to construct complicated ideas using simpler representations [10]. The DL is able to tackle intricate problems by describing each layer of the model as a nested simple mapping which simplifies the complicated mapping [11]. For complicated data patterns DL provides a wealth of advanced frameworks such as LSTM and CNN [12]. Images and data with a defined grid-like layout are ideal for the analysis and processing tasks that CNN are designed to handle. In contrast LSTM and NN excel in handling data sequences [13]. It appears that DL is a practical and effective method for extracting features from data and predicting future flight delays caused by weather given that weather has both spatial and temporal components. Utilizes LSTM and BiLSTM models to provide a categorization method for flights delays [14]. Flight delays are a big problem for airlines since they are inconvenient for customers and cost money. A classification test has demonstrated encouraging results for the powerful DL approaches of BiLSTM and LSTM models. For this research it trained and tested the BiLSTM and LSTM models using a dataset of flight on-time performance data obtained from the of Transportation Statistics and utilized in the research [15]. However, in order to analyze the performance comparative similarities with others it is highly significant to use a BiLSTM type of RNN model with structure and an offline dataset such that used for flight delays [16]. Finally improve the SVM model for flight delay prediction it adds the previous knowledge of flight delay as an inequality matrix [17]. Using the flight records for the month and distance employed the CatBoost method to estimate flight delays based on weather data for variables including wind speed and visibility. Flight delay prediction using several ML models such as KNN, SVM, NBC and RF.

Through extensive investigation into feature selection, pre-processing, and model design, they attempted to tackle the multiple obstacles associated with flight delay prediction networks. Here are some important points that our research brings to light:

- I propose extracting features from complicated flight delay data using the XGBoost model and then assigning relevance scores to those features. Set different feature priority levels to obtain different amounts of features, and then use numerous metrics to evaluate the best feature set.
- The MIX_LSTM model was designed with the goal of achieving higher performance in flight delay prediction and detection. It is based on the classic DL LSTM model.

III. PROPOSED SYSTEM

Accurately predicting when planes will take off allows for better use of airport support, apron, and runway resources, which in turn facilitates better team decision-making.To meet the increasing demand for transportation in civil aviation, more flights will need to be operated. Flight operations are required to make use of the apron, the airport's assistance, and the runway resources. Flight operations may experience delays due to the disparity between increasing flight demand and available resources.

The Kaggle dataset was the first source for the dataset. The data came from the United States Department of Transportation's Bureau of Transportation Statistics [18]. A total of 1,685,597 flights from January 2019, February 2019, and January 2020 make up the training dataset. A dataset consisting of 569,133 flights in February 2020 is used for testing purposes.

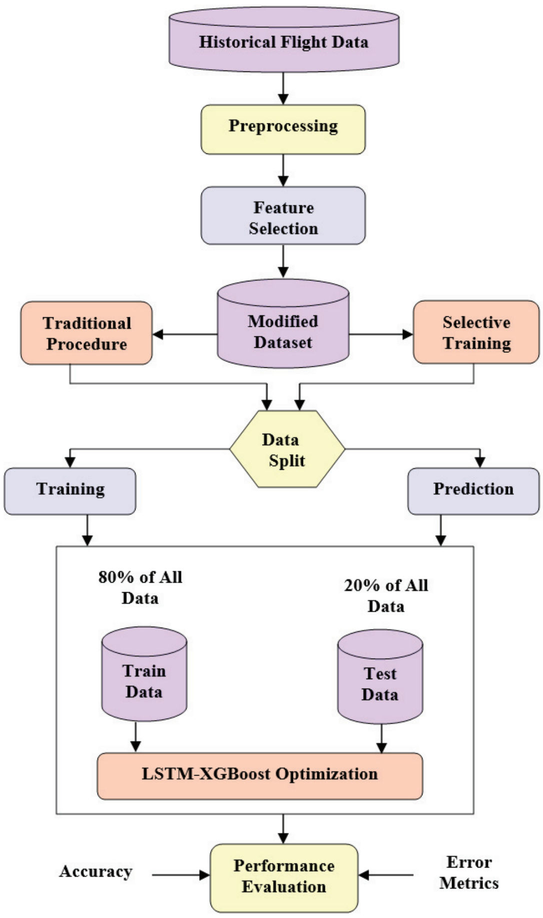

Fig. 1. The Proposed Model Flowchart

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.

It is common for data collected to be noisy, lacking certain details, and repeated. There must be cleaning and preprocessing procedures that include removing, replacing, transforming, and correcting data at this stage because of the inclusion of unnecessary information.After that, they separate the dataset into two parts: training and testing. Predictions are made using the test set.Grid search is employed to identify the ideal hyperparameters and model. They evaluate and contrast error metrics, computational expense, and accuracy for both standard and selective training to determine the efficacy of the proposed model. Figure 1 illustrates the flowchart of the proposed methodology.

A. Data Preprocessing:

1) Cleaning and Preprocessing:

During preprocessing, they remove extraneous information and retain just pertinent data to ensure coherence [19]. This procedure involves the preparation and formatting of data prior to training through the following steps:

- ✓ Eliminating data entries with absent values
- ✓ Removing missing data and mistakes
- ✓ Data duplication elimination
- ✓ Transforming numerical data from a category format

B. Feature Extraction:

The objective of FS in ML is to reduce operational time, enhance model accuracy, and remove redundant features.The most common approaches to FS include filter, embedded, and wrapper methods [20]. The Boruta algorithm is used to select features in this study.Using RF as its foundation, Boruta is an encapsulated method for FS. Whether or not to keep each characteristic is decided by calculating its importance to the variable that is dependent.The Boruta algorithm 1 mostly involves the following steps:

Algorithm 1: Boruta Algorithm	
1.	Shadow feature establish:To generate a new feature matrix, they first randomly arrange the original features into a feature of shadow matrix, and then then combine the two matrices to get the new one.
2.	The new feature matrix is incorporated into a RF classifier for training, yielding the importances of features u.
3.	Equation (1) represents the formula for calculating the Y score of both the original feature and the feature of shadow. $$y_{score} = \frac{B_u}{T_u} \quad (1)$$ In this context, B_u denotes the mean feature importance value and T_u denotes the SD.
4.	In the feature of shadow, which is represented as y_{max}, the maximum $y_{score}$is searched.
5.	With a y_{score} larger than y_{max}, features are considered important. In contrast, the feature will be eliminated and labelled as unimportant if its original y_{score} is lower than y_{max} prime.
6.	When all characteristics have been marked, the process starts over with steps (1) to (5).

C. Model Training:

1) XGBoost:

In ML, XGBoost finds extensive use, particularly in DM and recommendation systems, because to its efficiency, flexibility, and lightweight nature.Fundamentally, it relies on training tree models using data derived from the derivatives second-order of the function target using a Taylor second-order expansion. Furthermore, it integrates the intricacy of tree models as a regularisation component in the optimisation target, hence improving the generalisation capacity of the acquired model [21]. Consequently, it is often referred to as gradient enhancing decision trees. Let ι be the goal function, e_s the feature parameters of the s tree, and $\Phi(e_s)$ the model's complexity. The Taylor second-order series of the function target ι at the s iteration is expressed as Equation (2):

$$\iota^{(s)} \simeq \sum l_{j=1} \left[k\left(z_j, z_j^{s-1}\right) + h_j e_s\left(w_j\right) + \frac{1}{2} g_j e_s^2(w) \right] + \Phi(e_s) \quad (2)$$

In which h_j stands for first-order gradients and g_j for second-order ones.

In order to find the optimal splitting node in equation (3), gain is utilised during model training, much as the Gini index and IG in DT:

$$gain = \frac{1}{2} \left[\frac{\left(\sum j\epsilon J_K{}^{hj}\right)^2}{\sum j\epsilon J_K^{gj} + \mu} + \frac{\left(\sum j\epsilon J_Q{}^{hj}\right)^2}{\sum j\epsilon J_Q^{gj} + \mu} - \frac{\left(\sum j\epsilon J h_j\right)^2}{\sum j\epsilon J_j^{g} + \mu} \right] - \beta \quad (3)$$

This is where the samples in the left and right nodes after splitting are represented by J_K and J_Q, respectively, and $J = J_K \cup J_Q$; The information scores for the left subtree and right subtree are represented by $\left(\sum j\epsilon J_K{}^{hj}\right)^2 / \sum j\epsilon J_K^{gj} + \mu$ and $\left(\sum j\epsilon J_Q{}^{hj}\right)^2 / \sum j\epsilon J_Q^{gj} + \mu$, respectively. The present information indivisible score is $\left(\sum j\epsilon J h_j\right)^2 / \sum j\epsilon J_j^{g} + \mu$. Parameters μ and β are penalty terms. The significance score of features is determined by averaging the gains of all features, with gain standing for the gain score of each tree split.By dividing the sum of all trees' gains by the sum of all features' splits, they can find the average gain. Features are kept in XGBoost if their relevance ratings are higher than a certain threshold, while features with lower scores are ignored.

2) The Model Mix_LSTM:

This study presents the Mix_LSTM model, which is based on LSTM and used for analysis and prediction of flight delays.The model draws on a special kind of LSTM called Bidirectional LSTM. This kind of LSTM can think about both the past and the future, unlike the more common standard LSTM.To create a completer and more detailed picture of the sequence, forward LSTM records data from the past and backward LSTM records data from the future.The input data, following XGBoost FS, is the leftmost component. The central module illustrates a multi-layer Bidirectional LSTM, while the subsequent three layers represent the configurations of a double, single, and triple layer of Bidirectional LSTM, respectively. By stacking many Bidirectional layers of LSTM sequentially, the network can learn to extract characteristics that are higher-order and more abstract. This makes it possible for it to represent the supplied data more accurately. The right-most FC layer performs the last classification step after merging the outputs of all Bidirectional LSTM modules. In order to perform precise data categorisation and analysis, the FC layer must link deep features to preset output categories.

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.

A typical LSTM network has a hidden state g_s, a candidate cell state D_s, the current cell state D_s, and three gates output, input, and forget.In this case, $g_s$represents the output of the network, and the user decides on the size of $\widetilde{D}_s$. Writing, reading, and resetting of data sequential are all handled by the gates of LSTM. It is worth mentioning that the final output of each gate is processed using a function sigmoid, which limits the values to 0 to 1^2. The specifications for these modules are given by equations (4) and (5).

$$\widetilde{D}_s = \tanh\left(X_{wgd} Ws + W_{ggd} g_{s-1} + a_d\right) \tag{4}$$

where the bias is denoted by a_d and the weight matrices are denoted by X_{wgd} and W_{ggd}.

$$D_s = E_s \odot D_{s-1} + J_s \odot \widetilde{D}_s \tag{5}$$

D_s denotes the cell state at the current time step; E_s represents the output of the forget gate, regulating the degree to which information from the preceding time step is discarded; $\odot$ signifies element-wise multiplication. The J_s output of the input gate regulates the extent of current time step information incorporation; This is essential for alleviating the issue of gradient disappearing.

The forget gate is a crucial element of LSTM. The function of sigmoid activation regulates the forget gate, which decides if information from the preceding time step's memory state should be retained or discarded by constraining the gate's output to a range of 0 to 1. This process enables the LSTM network to preserve and modify significant memories over an extended duration while selectively eliminating extraneous information, as demonstrated in the subsequent formula (6):

$$E_s = sigmoid(X_{we} W_s + X_{ge} g_{s-1} + a_e \tag{6}$$

In this context, W_s represents the input at the current time step, g_{s-1} denotes the hidden state from the preceding time step, while X_{we} and X_{ge} signify the weights, and a_e indicates the bias.

The input gate, second in command within the LSTM architecture, governs the extent to which the data supplied at the current time step modifies D_s. By controlling the influx of incoming data, the network can selectively update D_s, enhancing its ability to retain and disregard specific elements of the input data. The formula (7) is as follows:

$$J_s = sigmoid\left(X_{wj} W_s + X_{gj} G_{s-1} + a_j\right) \tag{7}$$

When a_j is the bias and X_{wj} and X_{gj} are the input gate's weight parameters.

IV. Result and Discussion

An earlier study found that the National Airspace System (NAS) incurs substantial expenses as a result of flight delays.Air traffic delay analysis and prediction has been the subject of numerous studies with the goal of cutting down on unnecessary expenditure.More effective and less disruptive methods of air traffic management might be put in place according to the results of the study.Modelling and simulation have been used in a lot of the earlier studies.Computer simulations are used to recreate circumstances that require analysis by creating a model that mimics the real behaviours of the system's components.

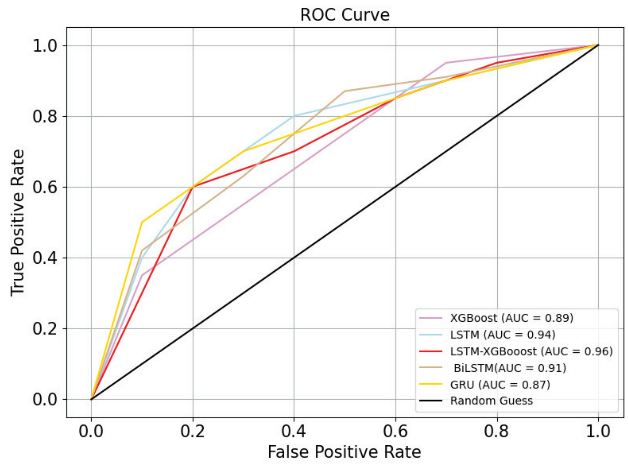

Fig. 2. ROC Curve

Figure 2 displays the model's ROC curve, which compares the different classifiers. K-Fold Cross Validation LSTM-XGBoost was a successful extra model evaluation tool.

TABLE I. Performance Comparison

Models	Accuracy	Precision		Recall	
		0	1	0	1
XGBoost	88.76	0.85	0.86	0.82	0.84
LSTM	93.84	0.90	0.92	0.87	0.89
LSTM-XGBoost	**95.57**	**0.92**	**0.94**	**0.89**	**0.91**
BiLSTM	90.61	0.87	0.89	0.84	0.86
GRU	86.42	0.83	0.85	0.80	0.82

The results of the model's classification phase are shown in Table 1. Among the classifiers tested, LSTM-XGBoost achieved the highest accuracy rate of 95.57%.

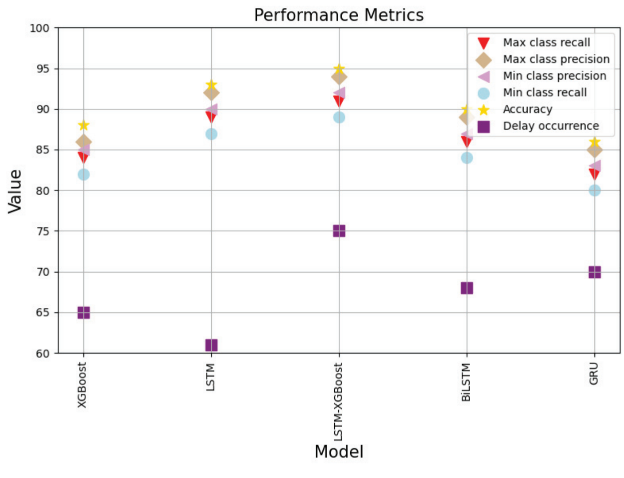

Fig. 3. Various Approach Comparison

Figure 3 shows the results for Minimum and Maximum Precision, Accuracy, Recall, and Delay Occurrence. Older planes were more likely to be delayed, according to the LSTM results.

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.

TABLE II. COMPARING THE ERROR RESULT

Models	RMSE	MAPE	MAE
XGBoost	3.14	4.28	1.74
LSTM	2.64	4.76	1.26
LSTM-XGBoost	**0.35**	**3.71**	**0.2**
BiLSTM	3.56	3.96	1.56
GRU	2.08	5.18	0.69

Table 2 shows the results of comparing the model established in this paper with the four classic models mentioned earlier. All models were tested using the same dataset and hyperparameters.

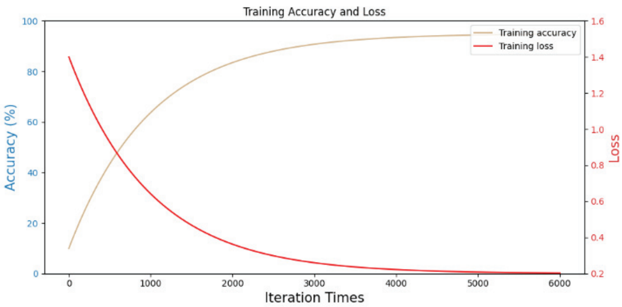

Fig. 4. Training Accuracy and Loss

The training procedure of the standard architecture is illustrated in Fig. 4. Moreover, as shown in Figure 4 shows that when the number of iterations grows, both the accuracy of the training dataset and the loss decreases.

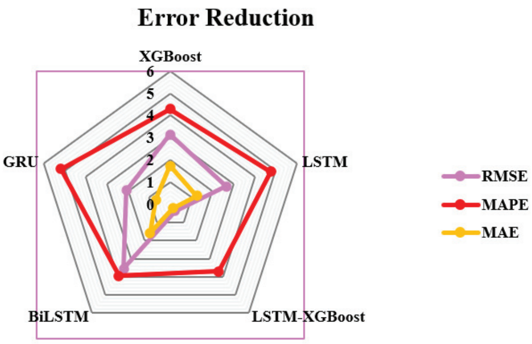

Fig. 5. Reduction of Error Performance

Figure 5 illustrates the percentage of errors reduced by LSTM-XGBoost. The suggested LSTM-XGBoost can greatly improve accuracy, as seen in the orange line for MAE, and the improvement in accuracy amounts to 0.2% for the remaining models.

V. CONCLUSION

Flight delay prediction is a crucial part of intelligent aviation systems; it has the potential to ripple across the entire aviation network, costing airlines and airports billions of dollars in lost revenue. As a result, this is rapidly becoming a pressing area of research for those looking to enhance the efficiency of airports and airlines.Punctuality of flights is a crucial indicator of the service excellence of an airport and airline. By predicting the duration of aircraft delays within a specified timeframe, airlines can promptly execute contingency measures, thereby averting revenue loss and incurring penalty costs.During preprocessing, the data is cleansed to ensure quality. The Boruta algorithm is employed for feature selection. Initially, by establishing various criteria for XGBoost in feature selection, characteristics with significance exceeding the specified threshold are preserved, but those falling below are disregarded. Ultimately, by integrating the optimal threshold and loss function, a model termed LSTM-XGBoost is proposed for the prediction of flight delays. The suggested LSTM-XGBoost model attains an accuracy of 95.57%, with RMSE of 0.35, MAPE of 3.71, and MAE of 0.2 in the binary flight delay prediction experiment conducted on the dataset.

REFERENCES

[1] S. Kim and E. Park, "Prediction of flight departure delays caused by weather conditions adopting data-driven approaches," *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-023-00867-5.

[2] Z. Mtimkulu and M. Maphosa, "Flight Delay Prediction Using Machine Learning: A Comparative Study of Ensemble Techniques," *Int. Conf. Artif. Intell. its Appl.*, vol. 2023, pp. 212–218, Nov. 2023, doi: 10.59200/ICARTI.2023.030.

[3] et al Ijemaru, Gerald K., "Image processing system using MATLAB-based analytics.," *Bull. Electr. Eng. Informatics*, 2023.

[4] M. Lu, P. Wei, M. He, and Y. Teng, "Flight Delay Prediction Using Gradient Boosting Machine Learning Classifiers," *J. Quantum Comput.*, vol. 3, no. 1, pp. 1–12, 2021, doi: 10.32604/jqc.2021.016315.

[5] I. Hatıpoğlu, Ö. Tosun, and N. Tosun, "Flight delay prediction based with machine learning," *Logforum*, vol. 18, no. 1, pp. 97–107, Mar. 2022, doi: 10.17270/J.LOG.2022.655.

[6] Q. Cai, S. Alam, and V. N. Duong, "A Spatial–Temporal Network Perspective for the Propagation Dynamics of Air Traffic Delays," *Engineering*, vol. 7, no. 4, pp. 452–464, Apr. 2021, doi: 10.1016/j.eng.2020.05.027.

[7] H. Khaksar and A. Sheikholeslami, "Airline delay prediction by machine learning algorithms," *Sci. Iran.*, vol. 26, no. 5 A, pp. 0–0, Dec. 2017, doi: 10.24200/sci.2017.20020.

[8] I. B. Mustapha, S. M. Shamsuddin, and S. Hasan, "A Preliminary Study on Learning Challenges in Machine Learning-based Flight Delay Prediction," *Int. J. Innov. Comput.*, vol. 9, no. 1, pp. 1–5, May 2019, doi: 10.11113/ijic.v9n1.204.

[9] Z. Guo, B. Yu, M. Hao, W. Wang, Y. Jiang, and F. Zong, "A novel hybrid method for flight departure delay prediction using Random Forest Regression and Maximal Information Coefficient," *Aerosp. Sci. Technol.*, vol. 116, p. 106822, Sep. 2021, doi: 10.1016/j.ast.2021.106822.

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

[11] B. Yu, Z. Guo, S. Asian, H. Wang, and G. Chen, "Flight delay prediction for commercial air transport: A deep learning approach," *Transp. Res. Part E Logist. Transp. Rev.*, vol. 125, no. March, pp. 203–221, May 2019, doi: 10.1016/j.tre.2019.03.013.

[12] M. Lambelho, M. Mitici, S. Pickup, and A. Marsden, "Assessing strategic flight schedules at an airport using machine learning-based flight delay and cancellation predictions," *J. Air Transp. Manag.*, vol. 82, no. June, p. 101737, Jan. 2020, doi: 10.1016/j.jairtraman.2019.101737.

[13] M. F. Yazdi, S. R. Kamel, S. J. M. Chabok, and M. Kheirabadi, "Flight delay prediction based on deep learning and Levenberg-Marquart algorithm," *J. Big Data*, vol. 7, no. 1, p. 106, Dec. 2020, doi: 10.1186/s40537-020-00380-z.

[14] Q. Li and R. Jing, "Generation and prediction of flight delays in air transport," *IET Intell. Transp. Syst.*, vol. 15, no. 6, pp. 740–753, Jun. 2021, doi: 10.1049/itr2.12057.

[15] W. Shao, A. Prabowo, S. Zhao, P. Koniusz, and F. D. Salim, "Predicting flight delay with spatio-temporal trajectory convolutional network and airport situational awareness map," *Neurocomputing*, vol. 472, pp. 280–293, 2022, doi: 10.1016/j.neucom.2021.04.136.

[16] A. Aljubairy, W. E. Zhang, A. Shemshadi, A. Mahmood, and Q. Z. Sheng, "A system for effectively predicting flight delays based on IoT data," *Computing*, vol. 102, no. 9, pp. 2025–2048, Sep. 2020, doi: 10.1007/s00607-020-00794-w.

[17] J. Chen and M. Li, "Chained Predictions of Flight Delay Using Machine Learning," in *AIAA Scitech 2019 Forum*, Reston, Virginia: American Institute of Aeronautics and Astronautics, Jan. 2019. doi:

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.

10.2514/6.2019-1661.
[18] T. Matsui, "Flight Delay Prediction | Kaggle," *kaggle*. https://www.kaggle.com/c/1056lab-flight-delay-prediction (accessed Apr. 09, 2025).
[19] H. Alla, L. Moumoun, and Y. Balouki, "A Multilayer Perceptron Neural Network with Selective-Data Training for Flight Arrival Delay Prediction," *Sci. Program.*, vol. 2021, pp. 1–12, Jun. 2021, doi: 10.1155/2021/5558918.
[20] J. Yi, H. Zhang, H. Liu, G. Zhong, and G. Li, "Flight Delay Classification Prediction Based on Stacking Algorithm," *J. Adv. Transp.*, vol. 2021, pp. 1–10, Aug. 2021, doi: 10.1155/2021/4292778.
[21] Z. Chen, Z. Li, J. Huang, S. Liu, and H. Long, "An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM," *Sci. Rep.*, vol. 14, no. 1, p. 23969, Oct. 2024, doi: 10.1038/s41598-024-74822-6.

Authorized licensed use limited to: Alliance University. Downloaded on February 13,2026 at 06:15:09 UTC from IEEE Xplore. Restrictions apply.