INTERNATIONAL JOURNAL OF APPLIED METHODS IN ELECTRONICS AND COMPUTERS

www.ijamec.org

International Open Access

Volume 12
Issue 02

June, 2024

Research Article

Enhancing Travel Experience: Predicting Flight Delays for Informed Journey Planning

***Esma Ergün** [*,a], **Süha Tuna** [b]

[a]*Istanbul Technical University, Informatics Institute, Maslak 34469, Istanbul, Türkiye*
[b]*Istanbul Technical University, Informatics Institute, Maslak 34469, Istanbul, Türkiye*

ARTICLE INFO

Article history:
Received 03 April 2024
Accepted 29 June 2024

Keywords:
Classification
Flight Delay Prediction
Flight Recommendation
L-GBM
Machine Learning

ABSTRACT

Flight delays pose significant inconveniences for travelers, potentially causing missed connections, schedule adjustments, and time wastage. This study presents a machine-learning driven approach to mitigate these challenges by developing an application that predicts flight delays, empowering passengers with insights to minimize travel disruptions. Leveraging diverse machine learning algorithms and datasets from the United States Department of Transportation and the National Oceanic and Atmospheric Administration Service, our model aids travelers in making informed decisions by suggesting optimal flight times and carriers based on historical flight data and weather conditions. Addressing the issue of imbalanced data, we explore techniques such as Synthetic Minority Over-sampling Technique (SMOTE) and under-sampling. Our comparative analysis highlights the superior performance of Light Gradient Boosting Machine (L-GBM) in predicting flight delays. With an F1-score of 56% and an AUC value of 0.76, our study offers a promising solution to enhance passenger experiences through improved flight recommendations.

This is an open access article under the CC BY-SA 4.0 license.
(https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

Today one of the biggest problems of many airlines is flight delays and cancellations, which can be caused by weather conditions, air traffic, airport conditions and crew planning. This situation adversely affects the business or holiday plans of the passengers.

In this study, we aim to predict flight irregularities (delay and cancellation) through various classification techniques using flight data published publicly by the United States Department of Transportation. With the software to be developed in the present study, delays will be predicted and the most appropriate time interval and airline suggestions will be provided to the passengers. Therefore, it will be easier for the passengers using air transportation, to choose the correct flight time and airline.

In most of the studies on the prediction of flight delays, delay prediction is performed for the destination station by using the departure delay information obtained after the aircraft has taken off. Such predictions are generally used for airline or airport operations rather than passengers. The purpose of this study is to help passengers to choose a flight by making predictions with the information afforded days before the flight.

Using the results of this study, an application can be developed for passengers and utilized in daily life for business trips, vacation trips, etc. For air travel to be carried out for various reasons, one of the flights with less probability of delay is preferred. In this way, the possibility of negative effects on business trips or personal plans will be reduced.

However, airlines and airports can integrate the produced software in this study into their operational applications, detect these flights in advance, optimize personnel planning, and help operation teams easily track risky flights. By following the operational processes of the flights determined as has irregularity more closely and taking precautions, the probability of delays can be decreased. Airlines make compensation payments to passengers when there is more than a certain amount of

* Corresponding author. E-mail address: *kurbane@itu.edu.tr*
DOI: 10.58190/ijamec.2024.96

delay. Therefore, airlines can save significant compensation payments by estimating the delay rates with the proposed method.

2. Related Work

There are several attempts to predict the flight delays using various datasets and techniques throughout the years. To this end, Belcastro et al. conducted research for this task by employing the flight data from the US Department of Transportation Bureau of Statistics acquired between January 2009 and December 2013. They also benefited the weather data from the National Climatic Data Center. They performed several machine learning based methods such as C4.5 Tree, SVM, Random Forest, Stochastic Gradient Descent, Naive Bayes and Logistic Regression in order to determine the most suitable model for the problem. According to their observations the most accurate results were obtained with the Random Forest method [1]. In this work, an unbalanced dataset was assessed. With the help of the Random Forest method, an accuracy rate of 69.1% was obtained when weather information was not used, while an accuracy rate of 85.8% and a Recall rate of 84.7% were obtained when weather information was included.

In another study, the impact of pre-prepared flight schedules on flight delays was examined, and a model was developed to predict flight delays and cancellations, which could be used for better schedule planning [2]. Classification was separately performed for departure delays, arrival delays, and flight cancellations. The study compared Random Forest, L-GBM, and Multilayer Perceptron models, identifying L-GBM as the best model based on both AUC values and processing times. According to results obtained with L-GBM, F1-scores were measured at 0.516 for departure delays, 0.560 for arrival delays, and 0.600 for cancellations. The AUC values were calculated as 0.786 for departure delays, 0.803 for arrival delays, and 0.929 for cancellations.

In the study conducted by Choi et al. in 2016, flight data from the US Department of Transportation Bureau of Statistics flight data from 2005 to 2015 and weather data from the National Climatic Data Center were benefited[3]. They exploited various models and, a comparison was made using the Area Under Curve (AUC) value since the unbalanced dataset was practiced. They determined as the model with the best Receiver Operating Characteristic (ROC) curve with an AUC value of 0.68. For the testing process, 3 different results were obtained by using the weather forecast 5 days ago and 1 day ago and the weather information on the same day. It was determined that the weather information on the same day gives much higher accuracy rates than the others. This is due to the inaccuracy of the weather forecast. If we compare the results with our proposed method, it is possible to observe that our method yields higher ROC-AUC values.

Ding conducted a research on predicting the arrival delays of flights with Multiple Linear Regression using the amount of departure delay and flight distance [4]. Using the 5-month flight data from 78 distinct airports of 175 different airlines, the accuracy value of the proposed model was observed to be approximately 80%. However, in this study predicting whether a flight will be delayed was not attempted, but to estimate the amount of delay in a delayed flight.

In another study [5], using the 8-month data of the United States Department of Transportation in 2015 and the aircraft information data of the Federal Aviation Administration (FAA), flight delays were predicted using Decision Trees, Random Forest and Multilayer Perceptron models. It was observed that the model using Multilayer Perceptron with 85% accuracy gives the best results.

On the other hand, using 2015 data from the United States Department of Transportation, the success of the model using Random Forest was achieved to be higher than 0.85 by Musaddi et al. [6]. Since the data volume is large, only the delayed flights were used, omitting the non-delayed flights. As a result of the study, it was stated that the delay probabilities according to the airline company could be determined and used for the airline selection for passengers.

In the study by Esmaeilzadeh and Mokhtarimousavi, departure delays for the primary airports of New York (EWR, JFK, and LGA) were predicted [7]. Regarding the model prediction performance, the results indicated that SVM could be a promising tool for predicting and analyzing flight delays. Departure delays were categorized into three levels: low (<15 minutes), medium (15-45 minutes), and high (45+ minutes). Relative probabilities were used to interpret the relationships among explanatory variables, considering percentage increases/decreases relative to other variables or the same variable's categories. Overall accuracy was found to be 0.855 with an AUC value of 0.959.

In our study, unlike many other studies, the Light Gradient Boosting Machine (L-GBM) algorithm was exploited to compare this algorithm with other methods. On the other hand, since the delay estimation is intended to be predicted before the flight, the departure delay information used in many studies was not employed. It was also examined how L-GBM works with SMOTE, a combination not commonly found in the literature.

3. Method

As the aim of the study is to determine whether the flights will have an irregularity or not, some classification techniques have been employed. Additionally, regression methods have been utilized, where class labels are assigned to probabilities based on a predefined threshold

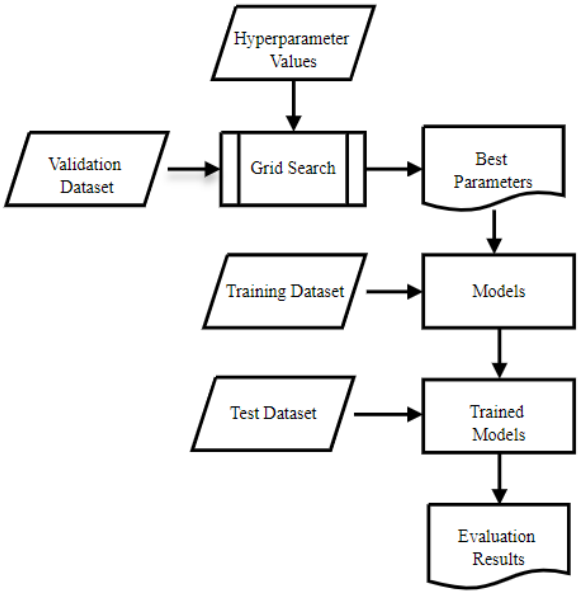

Figure 1. Working Structure of the Model

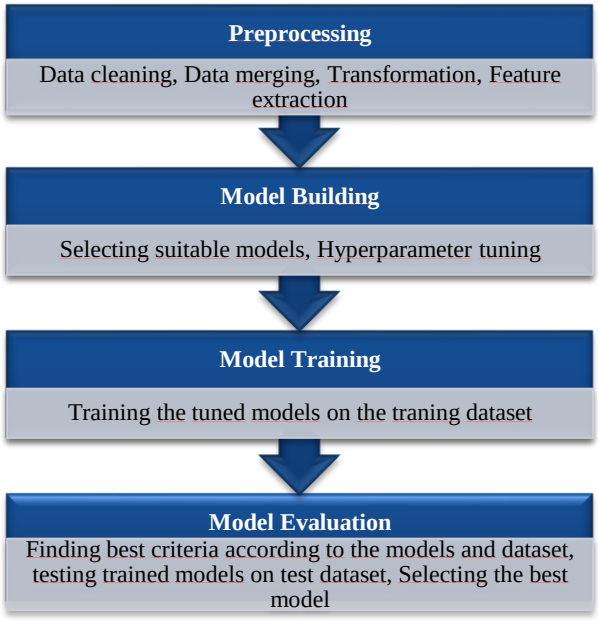

Figure *2*. Stages of the Study

value. The methods applied in this study encompass the L-GBM [8], a decision tree-based algorithm, Logistic Regression [1], a curve fitting method, Gaussian Naive Bayes [1] from the Bayes classifier family, K- Nearest Neighbor (KNN) [3], Multilayer Perceptron (MLP) [5] and Support Vector Machines (SVM) [1].

The study is split into stages that include data preprocessing, model building, training, and evaluation, respectively, as depicted in Figure 1. The model architecture is visually presented in Figure 2.

To address the imbalance sample amount with and without delay in the dataset, we employed two techniques: Synthetic Minority Over-sampling Technique (SMOTE) and Undersampling, respectively. These methods involve synthetic data duplication and reduce the number of majority class samples, respectively. The results obtained from both approaches are then compared.

Afterward, we establish a model using L-GBM and determine appropriate parameter values. To evaluate the performance of the model, we test it on samples that were not considered during training. Figure 3 illustrates the process of the algorithm. By providing the model with route information and time interval through our developed algorithm, we generate output for the user, suggesting the most suitable flight date and airline option for that specific route.

3.1. Synthetic Minority Oversampling Technique (SMOTE)

SMOTE involves generating synthetic data points for the minority class to balance the dataset, and since the positive class had fewer samples in the dataset used for this study, positive class samples were replicated while reducing the number of negative class samples.

Fernandez et al. described SMOTE as, rebalancing the original training set by using an oversampling approach. Rather than simply replicating instances of the minority class, SMOTE generates synthetic examples. These new data points are created by interpolating between various instances of the minority class that are in a specified neighborhood [9].

3.2. Light Gradient Boosting Machine (L-GBM) Algorithm

The L-GBM algorithm, is a decision tree-based ensemble learning method that brings weak learners together and turns them into strong learners, by providing computation speed and high accuracy. According to the research conducted by Ke et al. in 2017, the algorithm demonstrated similar accuracy rates compared to the standard Gradient Boosting Decision Tree algorithm but with a significant speed improvement, being approximately 20 times faster [8].

In the L-GBM model, leaf-based growth is used instead of level-based growth. Due to its high speed and high accuracy, this algorithm can be easily applied to large datasets.

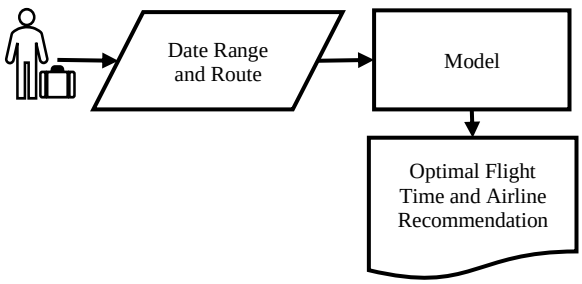

Figure 3. Process of the Software

4. Experimental Studies

In this work, flight and weather data were combined and adopted. The flight data [10] from the United States Department of Transportation for the year 2022 were used. The weather data [11] were acquired from the National Oceanic and Atmospheric Administration Service for the same period. John F. Kennedy International Airport (New York, JFK), which is one of the busiest airports in the United States, was selected and the flights departing from this airport were examined.

The features and their explanations in the data are provided in Table 1. The decision variable used in this case is Irregularity, which combines both the Cancelled and Delayed features. This variable signifies the presence of any irregularity, such as a flight delay exceeding 15 minutes or a cancellation.

4.1. Data Preprocessing

In the data preprocessing step, the acquired data were prepared to be employed in model training. This preparation process consists of the following stages.

- Feature Selection: The features that would not be useful in classification were determined and these features were removed from the dataset. For example, airline numeric code, airline name and city name. Similarly, in the weather data, the Source and Report Type features were omitted. These decisions were made based on judgment.
- Data Integration: The weather data and flight data were merged into a unified dataset. Initially, flight dates were converted to Coordinated Universal Time (UTC). Then, the datasets were combined based on date and time values that were the closest. For example, a flight departing at 01.01.2022 12:15 was matched with weather data from 01.01.2022 12:00
- Data Imputation: Missing weather features were filled using an interpolation method that calculates the average of the nearest above and below values. Some features contained values of 999 or 9999, which were also handled using the same interpolation process.
- Feature Extraction: IsHoliday feature extracted from Flight Date feature by using Python Holidays Library.
- Data cleaning: Within the weather data, there were daily summary records in addition to hourly records. Daily summaries with the Report Type SOM or SOD were filtered out.
- Feature Conversion: Categorical features were converted to binary variables so that they can be processed by the models. Additionally, numerical features were converted to take values between 0-1 by applying the min-max normalization procedure.
- Additionally, the WIND feature had direction and speed information in the same column. It's divided into two features Wind Direction and Wind Speed.

4.2. Parameter Optimization

Except for the Naive Bayes method, all models are parametric models. Therefore, it is necessary to determine the most appropriate parameters for the data. There are various methods for determining the most suitable parameters. The Grid Search Cross Validation (CV) method in the Python scikit-learn library was adopted to determine the best parameters. With this method, the parameters to be optimized for a model and the list of values that these parameters can take, are given to the Grid Search CV method, so that the algorithm creates a model for each value of each parameter in the given range. These models are evaluated according to the measurement criteria and given to the method, and the parameter set that gives the best measurement result is given as output. In this study, the F_1-Score was used as the measurement criterion.

The dataset is partitioned into three segments: a validation set, a training set, and a test set, accounting for 20%, 20%, and 60% of the data, respectively. The validation set is only used for parameter optimization.

The values Grid Search CV Method provided, were utilized in the training of the models. The method was applied to both the unbalanced original dataset and the datasets with SMOTE and undersampling, leading to variations in the optimal parameter values.

4.3. Evaluation Metrics

As the evaluation metrics, Accuracy, Precision, Recall and $F1$ -Score which are the criteria employed for classification techniques, were exploited. Furthermore, we also analyzed confusion matrices for a deeper understanding of the results.

Since there is a data imbalance in this dataset, the Accuracy criterion is not sufficient to measure the efficacy of the model. For example, when the model run on a dataset containing 90% negative class samples assigns all samples to the negative class, the Accuracy value will be found as 90%. Although 90% seems like a good rate, it is a very weak model since the model makes no distinction between samples. Therefore, for unbalanced datasets, the $F1$ Score, which takes the harmonic average of the Precision and Recall values, is more suitable for measuring the performance of the model.

Additionally, in this context, misclassifying delayed flights as not delayed is a more significant issue than misclassifying not delayed flights as delayed. This is because travelers tend to select only those flights labeled as not delayed. Consequently, prioritizing Recall over Accuracy is of greater importance in addressing this problem.

ROC Curves and the area under the ROC Curve (ROC-AUC) criterion were used to compare the models with

Table 1. Features

Dataset	Feature	Description
Flight	Day of Week	Day (Mon, Tuesday etc.)
	Month	
	IATA Code Operating Airline	Airline Code
	Destination	Arrival Airport Code
	Departure Time Bulk	Departure Time Interval (ex: 12:00- 13:00)
	CRS Departure Time	Local Scheduled Departure Datetime
	Cancelled	Yes/No
	Delayed	Yes/No
Weather	Date	Date and Time of Observation
	Source	Source of Observation
	Report Type	Type of observation
	WND	Wind Observation
	CIG	Sky Condition
	VIS	Visibility
	TMP	Air Temperature
	DEW	Dew Point Air-Temperature
	SLP	Air-Pressure

each other. Generally, the model that gives the closest ROC-AUC value to 1 is considered as the model that yields the most successful classification.

4.4. Data Balancing

When the dataset is examined according to the decision variable Irregularity, it is observed in Table 2 that the samples are not evenly distributed, leading to an unbalanced dataset. To address the data imbalance, there are multiple approaches available, such as undersampling, oversampling, and SMOTE. For our study, we adopted the SMOTE technique.

Given that SMOTE is primarily an oversampling method, we also conducted undersampling to facilitate a comparison.

4.5. Evaluation Results

The classification models were implemented by using the Python scikit-learn library and by giving the optimized parameter values obtained from the Grid Search CV Method. These models were trained on the training set. The classification results for unbalanced data are presented in Table 3. According to the results, L-GBM has the highest accuracy and F_1-Score values. Also, the Recall value is the second highest among other models. In Figure 4 one

Table 2. Distribution of Decision Variable

Irregularity	Number of Samples
Yes	38748
No	97833

Table 3. Evaluation Results with Unbalanced Dataset

Model	Accuracy	F1-Criteria	Precision	Recall
Light GBM	0.776	0.506	0.667	0.407
SVC	0.758	0.390	0.673	0.274
Polynomial SVC	0.764	0.448	0.660	0.339
MLP	0.754	0.480	0.595	0.402
Naive Bayes	0.647	0.481	0.410	0.581
KNN	0.752	0.444	0.602	0.352
Logistic Regression	0.746	0.376	0.615	0.270

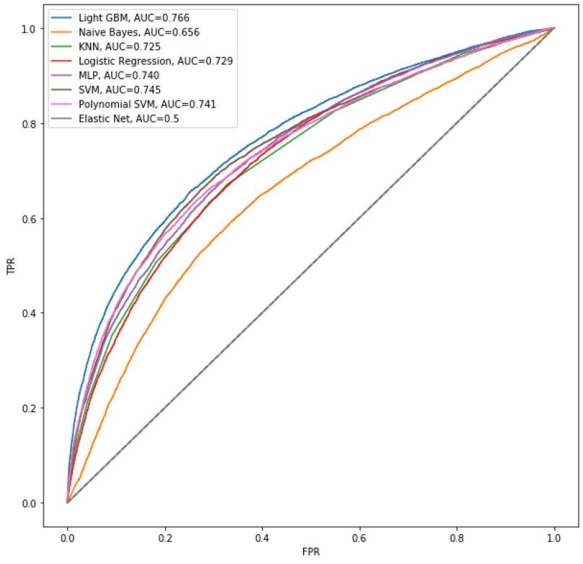

Figure 4. ROC Curves and ROC-AUC Values of Models with Unbalanced Dataset

can observe that the ROC-AUC value of L-GBM is the highest with 0.766.

The evaluation results of the models trained with the SMOTE applied data are shown in Table 4. While the highest $F1$-score was 0.506 for the L-GBM Model before SMOTE was applied, it increased to 0.523 after SMOTE was applied.

When the ROC Curves and ROC-AUC values provided in Figure 5 are examined, it is seen that the ROC-AUC value decreased from 0.766 to 0.759.

When undersampling is employed on the dataset, the top-performing model shifted the SVM Method, achieving a higher F_1-Score of 0.567 and a Recall of 0.681. The L-GBM approach closely follows with an F_1-Score of 0.564 and an identical recall value (Table 5).

Comparing the ROC-AUC values from Figure 6, the same characteristics are encountered. SVM with 0.765 and L-GBM with 0.763 have the best scores.

On the other hand, the undersampling outperforms SMOTE in addressing this problem. These findings suggest that the optimal data balancing technique may vary depending on the dataset and model in use.

The results indicate that the Light-GBM with unbalanced data yields the highest ROC-AUC value.

Table 4. Evaluation Results with SMOTE Applied Dataset

	Accuracy	F1-Criteria	Precision	Recall
Light GBM	0.761	0.523	0.598	0.466
SVC	0.761	0.466	0.629	0.37
Polynomial SVC	0.759	0.502	0.602	0.43
MLP	0.733	0.491	0.532	0.456
Naive Bayes	0.554	0.468	0.352	0.697
KNN	0.584	0.506	0.380	0.757
Logistic Regression	0.732	0.482	0.530	0.442

However, SVM with the undersampling has the highest F_1-Score with 0.567. Light-GBM consistently demonstrates excellent performance for both unbalanced and balanced datasets, with the added advantage of significantly shorter processing times compared to SVM.

This efficacy is attributed to the Light-GBM's ensemble learning approach, which leverages weak learners to minimize errors and its compatibility with categorical features.

Naive Bayes and K-Nearest Neighbor algorithms were the models with the worst results. The Naive Bayes algorithm assumes that there is no relationship between the features but there are related features like isHoliday and dayOfWeek, sky condition, visibility, etc. Since the K-Nearest Neighbor algorithm is distance-based, it is insufficient to detect complex relationships in the data.

When we analyze the complexity matrices of the Light-GBM models, Table 6 reveals that Light-GBM incorrectly classified 1560 negative samples and 4556 positive samples. Thus, it is evident that the performance of Light-GBM in detecting negative samples was poor when dealing with unbalanced data.

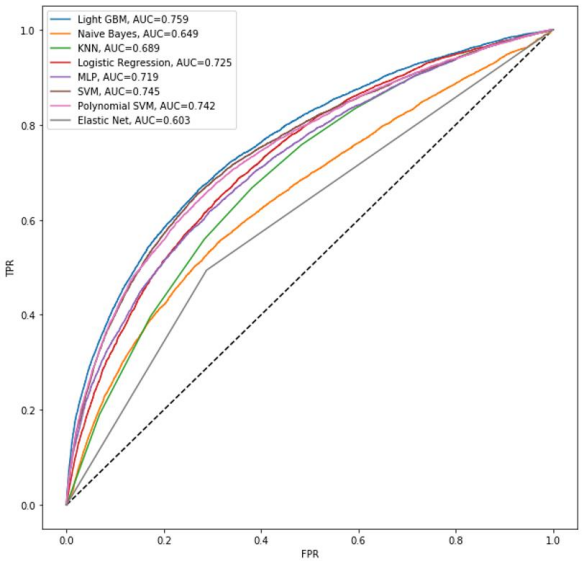

Figure 5. ROC Curves and ROC-AUC Values of Models with SMOTE

Table 5. Evaluation Results with Undersampling Applied Dataset

Model	Accuracy	F1-Criteria	Precision	Recall
Light GBM	0.704	0.564	0.482	0.681
SVC	0.707	0.567	0.485	0.681
Polynomial SVC	0.698	0.559	0.475	0.678
MLP	0.689	0.558	0.465	0.698
Naive Bayes	0.636	0.482	0.402	0.600
KNN	0.675	0.540	0.449	0.676
Logistic Regression	0.672	0.531	0.444	0.660

This conclusion is further supported by the Recall score, which stands at 0.407. On the other hand, utilizing Undersampling with Light-GBM yields a higher Recall value of 0.681, making it more suitable to address this issue. However, it is worth noting that this approach results in a slightly lower ROC-AUC score than using Light-GBM with an unbalanced dataset, with a difference of only 0.003.

5. Conclusion

In this study, various classification techniques were used. The Light-GBM Model was observed as the most efficient in computation speed while SVM is the most successful method in overall accuracy. To improve the results, parameter optimization was performed with the Grid Search and 8-fold Cross Validation methods. Among SMOTE and undersampling techniques, undersampling had better results and was employed in the proposed model.

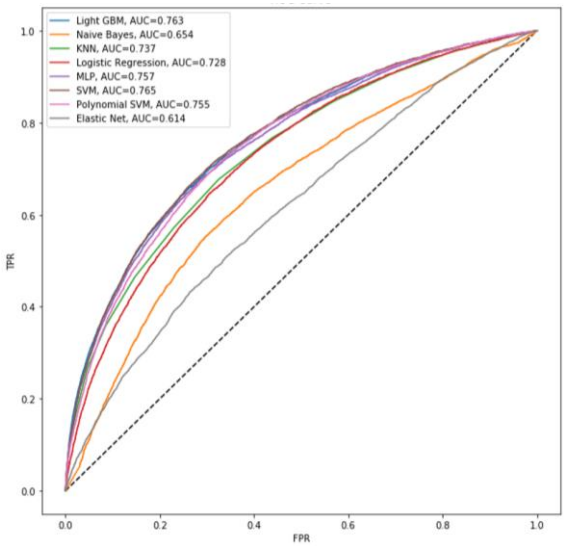

Figure 6. ROC Curves and ROC-AUC Values of Models with Undersampling

Table 6. L-GBM with Unbalanced Dataset Confusion Matrix

	Actually Positive	Actually Negative
Positive Classified	3138	1560
Negative Classified	4556	18062

The comparison of our results with other studies are provided in Table 7. According to the table, our AUC value and F1-criteria scores are higher than those of corresponding studies.

Using the outputs of this study, the results of the software for a user who wanted to travel from New York to Chicago on January 24th are shown in Table 8. According to these results, it is expected that there will be an irregularity in X Airline flights between January 24,15:00-16:00 and 17:00- 18:00. By considering the results, the user is guided to choose a flight other than these flights. Original Airline names were replaced with the values X and Y.

In future studies, features such as crew, planning, number of passengers and baggage and ground operation times from airline companies can be acquired and added to the dataset. In this way, since the dataset will contain more information about the problem, more efficient models can be developed. Also, in this study, the MLP yields similar results to the L-GBM Model. It seems possible to improve the results by using appropriate deep-learning methods.

Table 7. Comparison of the results

	Method	Accuracy	Sensitivity	F1-Criteria	AUC
R. Musaddi, A. Jaiswal, and M. Girdonia [6]	Random Forest	0.85			
R. Henriques and I. Feiteira [5]	Multilayer Perceptron, SMOTE	0.83		0.79	0.56
L. Belcastro, F. Marozzo, D. Talia, and P. Trunfio [1]	Random Forest	0.858	0.869		
S. Choi, Y. J. Kim, S. Briceno, and D. Mavris [3]	Random Forest	0.803			0.68
Esmaeilzadeh, E., & Mokhtarimousavi, S. [7]	SVM	0.855	0.853	0.850	0.95
Lambelho, M., Mitici, M., Pickup, S., & Marsden, A. [2]	L-GBM	0.794	0.516	0.516	0.786
Ergün, E. and Tuna, S.	L-GBM, Undersampling	0.704	0.681	0.564	0.763

Table 8. *Flight and Airline Recommendations on Jan 24th for JFK to ORD*

Departure Time Range	Airline	Number of Flights	Number of Flights Having Irregularity
$0600-0659$	Y	2	0
$0700-0759$	Y	2	0
$0800-0859$	X	2	0
$1100-1159$	Y	2	0
$1300-1359$	Y	2	0
$1500-1559$	Y	2	2
$1600-1659$	X	2	0
$1700-1759$	Y	2	2
$1900-1959$	X	2	0

Acknowledgment

This study has been presented in 7th International Conference on Engineering Technologies (ICENTE 2023), 23-25 November 2023, Konya/Turkey.

References

[1] L. Belcastro, F. Marozzo, D. Talia, and P. Trunfio, "Using Scalable Data Mining for Predicting Flight Delays," *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. 1, pp. 1–20, Oct. 2016, doi: https://doi.org/10.1145/2888402.

[2] Lambelho, M., Mitici, M., Pickup, S., & Marsden, A. (2020). Assessing strategic flight schedules at an airport using machine learning-based flight delay and cancellation predictions. *Journal of air transport management*, *82*, 101737.

[3] S. Choi, Y. J. Kim, S. Briceno, and D. Mavris, "Prediction of weather-induced airline delays based on machine learning algorithms," *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, Sep. 2016, doi: https://doi.org/10.1109/dasc.2016.7777956.

[4] Y. Ding, "Predicting flight delay based on multiple linear regression," *IOP Conference Series: Earth and Environmental Science*, vol. 81, p. 012198, Aug. 2017, doi: https://doi.org/10.1088/1755-1315/81/1/012198.

[5] 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: https://doi.org/10.1016/j.procs.2018.10.085.

[6] R. Musaddi, A. Jaiswal, and M. Girdonia, "Flight Delay Prediction using Binary Classification," *International Journal of Emerging Technologies in Engineering Research (IJETER)*, vol. 6, 2018, Accessed: Oct. 28, 2023. [Online]. Available: https://www.ijeter.everscience.org/Manuscripts/Volume-6/Issue-10/Vol-6-issue-10-M-09.pdf

[7] Esmaeilzadeh, E., & Mokhtarimousavi, S. (2020). Machine learning approach for flight departure delay prediction and analysis. *Transportation Research Record*, *2674*(8), 145-159.

[8] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," *Advances in Neural Information Processing Systems*, vol. 30, 2017, Available: https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

[9] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, Apr. 2018, doi: https://doi.org/10.1613/jair.1.11192.

[10] Bureau of Transportation Statistics. "On-Time : Marketing Carrier On-Time Performance (Beginning January 2018)" Available: https://www.transtats.bts.gov/DL_SelectFields.aspx?gnoyr_VQ=FGK&QO_fu146_anzr=b0-gvzr (accessed Oct. 28, 2023).

[11] National Oceanic and Atmospheric Administration. "Global Hourly Weather Data - 2022." Available: https://www.ncei.noaa.gov/data/global-hourly/access/2022 (accessed Oct. 28, 2023)