PREDICTING RUNWAY CONFIGURATION IN THE CONTEXT OF THE CONGONHAS AIRPORT USING MACHINE LEARNING

Bruno M. Franzini[1,2], Tábada Domingues[1], Ivan M. Silva[1], Pollyanne E. da Silva[1],
André L. E. Melo[1], Carlos H. Q. Forster[2], Jean P. O. Lima[1,*] and Marcos R. O. A. Máximo[2]
[1]Airspace Control Institute (ICEA), Research Division
[2]Autonomous Computational Systems Lab (LAB-SCA), Computer Science Division, Aeronautics Institute of Technology (ITA)
*** Corresponding author e-mail address**: jeanphelipedeoliveira@gmail.com

PAPER ID: SIT904960

Abstract

The complexity and criticality of airspace control, coupled with the increasing demand for flights, highlight the need for advanced tools to optimize and secure operations. Additionally, many variables influence the air traffic scenario and, given the difficulty of analytically assessing the impact of each variable, a possible strategy is to use Machine Learning techniques to more accurately assign the weight of each factor, by processing historical data. In this context, considering that one of the variables that most influences air traffic is runway reconfiguration, this paper presents the development of a system to predict which runway will be in use at Congonhas Airport (SBSP) using a Gradient Boosting technique, called LightGBM. A classification model based on LightGBM was developed using data from the Meteorological Aerodrome Report (METAR), Terminal Aerodrome Forecast (TAF), and Weather Research and Forecasting (WRF). The model makes predictions about possible runway reconfigurations, deciding which side of the runway will be used in the coming hours. With this data, an accuracy of 98% in predictions was achieved and when testing the model on a period outside the training and test dataset, an accuracy of 88% was obtained. These results highlight the efficiency of the developed model compared to the currently used rule-based methods, which achieved approximately 81%, according to related works. This work contributes by providing a better quantification of the influence of meteorological variables on airspace operations, according to a feature importance analysis, which can support a more efficient and safer airspace activity planning.

Keywords: Air Traffic Management, LightGBM, Meteorological Data, Runway Configuration.

Acknowledgements

Marcos Maximo is partially funded by CNPq - National Research Council of Brazil through the grant 307525/2022-8.

PAPER ID: SIT904960

1. INTRODUCTION

Airspace control is a complex and critical challenge due to the involved risks and the continuous increase in flight demand, estimated at 4.3% per year over the next 20 years (ACI World, 2023a). This growth demands that regulatory agencies adapt to manage a higher aircraft volume (ACI World, 2023b). In this context, one of the greatest challenges is balancing safety and efficiency, as an extremely safe operation can be inefficient, while a flexible operation may pose risks.

The efficient use of airspace economically impacts passengers and industry workers. In 2023, more than 112 million passengers flew in Brazil (Poder360, 2023), and about 1 million people were employed in this sector (IATA, 2018). Inefficiencies, such as delays and vectoring, particularly in terminal areas, increase operational costs, negatively impacting both airlines and passengers. It is estimated that the average cost of in-flight delays is 100 euros per minute (EUROCONTROL, 2015), highlighting the need for optimization.

Multiple variables, such as weather conditions and runway reconfigurations, affect flight punctuality. Quantifying the influence of these variables is challenging for humans due to the complexity and the need for real-time massive data processing (Prandini et al., 2011). Machine Learning can assist in extracting patterns from historical data to predict issues more accurately.

A significant problem for efficiency in terminal areas is runway reconfiguration. This process involves changing landing and takeoff operations to the other end of the runway. One of the determining factors for this event is the change in weather conditions, such as wind direction and speed. Runway reconfiguration is critical to maintaining the safety of air operations, as ideally, airplanes should take off and land against the wind to achieve proper lift and ensure safe maneuvers. Thus, one of the main catalysts for runway reconfiguration is tailwind, that is, the wind component in the same direction as the aircraft. Congonhas Airport (SBSP) has two parallel runways, with the thresholds: 17, the main threshold, which is most used under normal conditions, and 35, the opposite threshold. Anticipating these reconfigurations can allow for adequate planning by responsible agencies to minimize negative impacts such as delays and congestion.

The Airspace Control Institute (ICEA) has a runway configuration predictor (ICEA, 2024) that uses a Rule-Based Reasoning approach (Frye et al., 1995) built from ICA 100-37 definitions (DECEA, 2020) on data derived from the Weather Research and Forecasting (WRF). The predictor suggests a runway reconfiguration when the forecasted tailwind component exceeds 6 kt for the runway in operation. With this approach, 81% accuracy was achieved using a method that does not perform correlation analysis of variables, historical behavior, etc. Therefore, it is expected that a machine learning-based predictor can improve this accuracy by better understanding the influence of other variables.

This work aims to develop a Machine Learning model to predict the runway in operation at Congonhas Airport, based on training with historical data. This paper contributes to research in this field by proposing a methodology to train and evaluate various machine learning algorithms for the proposed task. Additionally, it employs data preprocessing techniques that can be used to standardize the representation of various variables and allow for better comparison between studies in the literature.

2. LITERATURE REVIEW

Given the number of variables involved in air traffic management, many studies seek to apply AI techniques in this context, with different approaches. Broadly and initially, Gosling (1987) proposes various areas within the context of airspace control where AI could contribute and yield positive results. The author identifies seven possible control strategies ranging from visual and electronic collision prevention to proposed improvements in the U.S. air traffic control systems, strategies in which aircraft follow predefined, non-conflicting trajectories. AI applications are grouped into seven functional areas and discuss ways to incorporate these applications into different control strategies, also considering implementation issues that may arise in the course of applying AI techniques for air traffic control.

Since the publication of this work, much has been developed in the aviation industry, and currently, there is increasingly more data available related to flights, leading some studies to seek to extract information from this data to assist in air traffic management. In this sense, Murça and Hansman (2019) present a data-driven framework to identify, characterize, and predict traffic flow patterns in the terminal area of multiple airport systems, known as metroplexes. This framework uses machine learning methods applied to historical flight trajectories, weather forecasts, and airport operational data. Through a multi-layered clustering analysis, the study identifies recurring patterns of runway and airspace use, as well as relevant decisive factors, and develops descriptive models for metroplex configuration prediction and capacity estimation. Applied to the New York metroplex, this framework demonstrates high accuracy in predicting traffic flow patterns for planning horizons of up to eight hours, highlighting the importance of metroplex configuration prediction for better flow rate planning and, consequently, better traffic regulation.

Additionally, other studies highlight the significant influence that external factors have on flight planning, further complicating the task of managing air traffic. Oliveira et al. (2021) investigate the impacts of weather conditions on the punctuality of domestic flights in Brazil. Using historical flight schedules and weather data, the authors estimate a logit model to analyze the effects of variables such as visibility, ceiling height, wind gusts, and precipitation on the probability of delays. The study concludes that adverse weather conditions significantly increase the likelihood of delays, contributing to a better understanding and management of weather impacts on flight punctuality.

Still in the context of runway conditions, Midtfjord, De Bin, and Huseby (2021) developed a decision support system for safer aircraft landings, using XGBoost to predict these conditions. This system combines classification and regression models trained with weather and runway data, achieving high accuracy and providing explainable predictions through techniques such as SHAP (SHapley Additive exPlanations).

Another significant work is by Herrema et al. (2019), who also use Gradient Boosting, but to predict runway exit times in Vienna, i.e., how long it will take for the aircraft to leave the runway after landing. Applying the technique to 54,679 arrival flights and analyzing scenarios that impact runway use, this model achieved 79% accuracy, demonstrating the applicability of Gradient Boosting in airport contexts to predict operational changes and improve efficiency.

Finally, Lau (2021) specifically focuses on predicting transition times during runway reconfiguration using ensemble methods such as Random Forest Regressor, AdaBoost, and Gradient Tree Boosting. The study showed that these methods could achieve R2 scores of at least 0.8, providing accurate predictions for runway configuration changes based on dynamic conditions like wind and clouds.

3. MATERIALS AND METHODS

3.1. Experimental Data

The data used in the development of the predictive model for runway reconfiguration, from training to testing, were collected from different sources. It is worth noting that the period sought for training and testing was from 12/28/2023 to 04/14/2024, limited by the available WRF data set. A total of 20,000 time instances (discretized into 15-minute intervals) were observed, with runway 17 being in use for 14,942 of these instances and runway 35 for the remaining 5,058 instances. In this sample, there were a total of 62 reconfigurations from runway 17 to 35 and 62 reconfigurations from 35 to 17. A brief explanation of the data used is presented below:

- **Weather Research and Forecasting:** The WRF data set is the result of a numerical weather prediction model developed to provide detailed, high-resolution simulations of weather conditions. The variables used were Wind Speed, Wind Direction, and Tailwind Component, generated for moments spaced 15 minutes apart;

- **Meteorological Aerodrome Report:** METAR is a routine weather observation report used in aviation to provide information on weather conditions at aerodromes. Issued every hour, the METAR includes data on wind speed, visibility, clouds, temperature, and pressure;

- **Terminal Aerodrome Forecast:** TAF is a weather forecast for an aerodrome, issued every 6 hours. It provides detailed information about the expected weather conditions for the next 24 to 30 hours, including wind, visibility, and clouds;

- **Runway Reconfiguration History:** Originating from tower data, the runway configuration history provides the sequence of landings performed at the aerodrome, including the times and the respective runway used (17 or 35, in the case of Congonhas).

3.2. Data Preprocessing

After collection, the data were integrated using WRF as the reference due to its higher sampling frequency (data every 15 minutes). The METAR and TAF were aligned with the nearest preceding hour for each WRF data point to simulate a real-time environment, organizing the data as they would be available in a production setting. Finally, the runway reconfiguration history was indexed by the runway in use at the reference time of each WRF data point.

After integration, preprocessing was conducted to prepare the data for the machine learning models to be tested, which included:

- Transforming date and time data into the time difference between the reference dates of the integrated data sets, to allow the models to properly understand the relevance of each variable. For example, if a METAR data from 01:00 is associated with a WRF data from 01:45, a column is created with information on the temporal distance between these variables, i.e., 45 minutes. This information can be useful for the model to interpret the importance of the data, i.e., a METAR data from 45 minutes ago is less relevant than a data from 15 minutes ago. This feature was called $time\_diff\_prev\_and\_ref$;

- Applying the One-Hot Encoding technique to categorical variables;

- Normalizing numerical variables using the standard scaler method. This representation usually improves model performance (Pedregosa, 2011). The normalization scale was saved to ensure consistency in the transformation of new test data sets.

3.3. Machine Learning Pipeline

After the data were prepared for input into the model, the next step was to choose the classification technique to be employed. For this purpose, 27 classification models were compared (following the methodology of Pandala et al. (2022), which involves minimal model parameterization) according to their respective performances. The model that achieved the highest accuracy was LightGBM (Ke et al. 2017). Additionally, some tools developed in ICEA initiatives have shown good results for similar tasks using LightGBM, such as in Teles and Zaneti (2024). Therefore, these factors, combined with its other advantages described in previous sections, led to LightGBM being chosen for the classification task in this work. Following this decision, the model underwent refinement in hyperparameter selection to seek pseudo-optimal performance, as will be explained below.

To optimize hyperparameters and train the model, it was necessary to partition the data set using an 80/20 Holdout, where 80% of the data were used as the training set and 20% as the test set. The training set, containing 16,000 examples, was used to adjust the model's parameters to enable the learning of existing patterns. The test set, with 4,000 examples, was used for the final evaluation of the model. It is important to note that this set was not used during training or hyperparameter tuning, ensuring a fair and unbiased evaluation of the model's performance.

The optimization of the LightGBM model’s performance was carried out using the Randomized Search technique (Bergstra et al. 2012), aimed at finding the best hyperparameters. This method was chosen due to its good performance in optimization (Bergstra et al. 2012). This process involved defining a distribution of hyperparameters: Leaves Number, Learning Rate, Feature Fraction, Bagging Fraction, maximum tree depth, and the regularization parameters lambda_l1 and l2. The search was conducted using cross-validation with the K-Fold method (Kohavi, 1995) with K=5 to ensure the robustness of the results. After training and evaluating the accuracy of all hyperparameter combinations, the set that achieved the highest accuracy on the validation set was selected.

Based on the selected hyperparameters, the LightGBM model was retrained using the entire training set, not just 80% as in the 5-fold cross-validation. This retraining policy allows the optimized model to enhance its learning capacity with the complete training set. It is worth noting that the training was conducted with the goal of minimizing binary error, iteratively adjusting the model parameters to improve its performance.

4. RESULTS AND DISCUSSIONS

4.1. Model Performance

The hyperparameter search phase yielded the following results: 101 for Leaves Number, 0.11 for Learning Rate, 0.79 for Feature Fraction, 0.92 for Bagging Fraction, 18 for Maximum Tree Depth, 0.57 for lambda_l1, and 0.52 for lambda_l2. After the process of finding the best model and subsequent training, as explained in Section 3, the accuracy obtained with the best hyperparameters was 98%, indicating the model's ability to accurately predict runway reconfiguration at Congonhas Airport. The model's training curve can be seen in Figure 1.

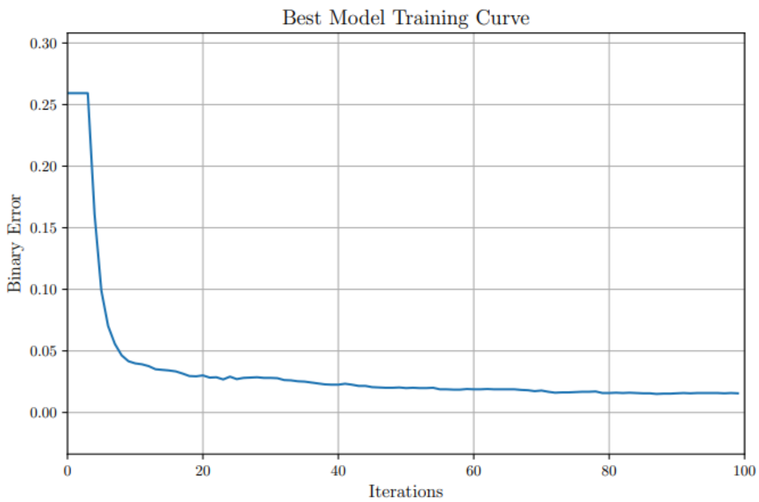

Figure 1: Training curve for the best model found by the Randomized Search.

Additionally, the intermediate results of the cross-validation can be seen in Figure 2, which shows the accuracy obtained by each of the 5 folds in their respective validation sets. These values result in an average accuracy of 82% with a standard deviation of 2%.

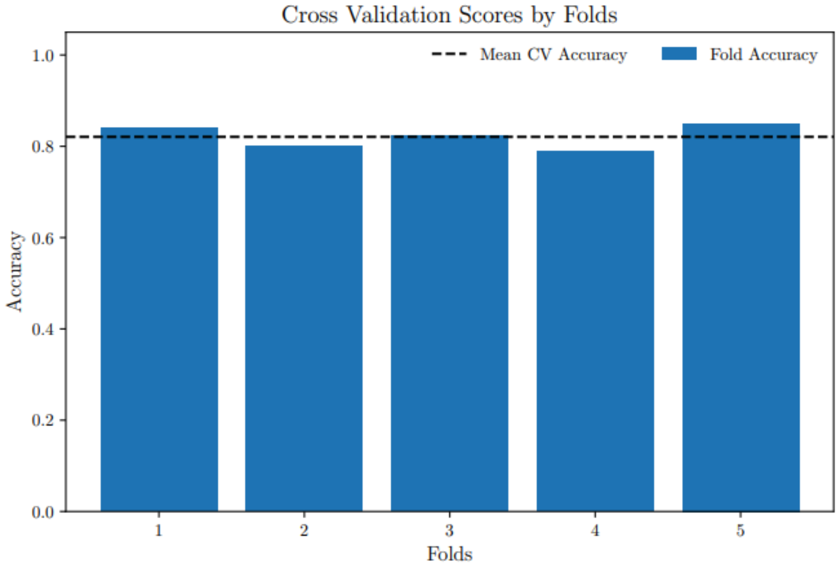

Figure 2: Accuracy of the different folds of the cross-validation.

Analyzing the training curve in Figure 1, it is noted that learning occurred as expected, with the binary error decreasing with each iteration and stabilizing. Moreover, the results obtained show that the LightGBM model is capable of predicting runway reconfiguration with good accuracy. However, to test the model in an even more unbiased manner, a final test was conducted with a dataset outside the period from 28/12/2022 to 14/04/2023, which was used for training and testing, to evaluate the model's generalization capability. Thus, the model was tested on the interval from 15/04/2023 to 15/05/2023, where it achieved an accuracy of 88% and an F1-Score of 86%, with the classification results presented in the confusion matrix of Figure 3.

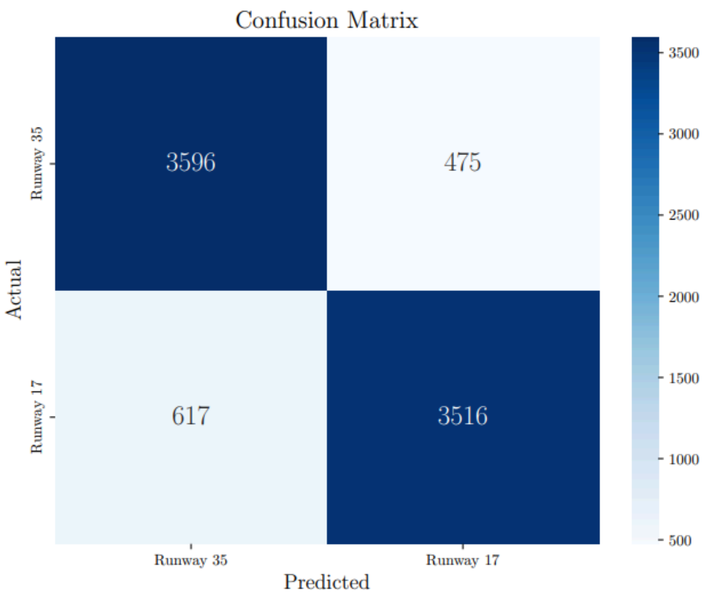

Figure 3: Confusion Matrix for the classification performed on the new dataset.

One of the problems when attempting to perform classification is the potential imbalance of the model, which can bias its decision and end up favoring one class over another. However, analyzing Figure 3, we can see that the model is making balanced classifications and, when an error occurs, it does not do so disproportionately for a specific class.

When analyzing the drop in accuracy from 98% to 88%, it is noted that the effect of distribution shift may be occurring, which refers to the difference in behavior between the test data and the new data exposed to the model. This may indicate that the problem has characteristics that vary over time, similar to a seasonality effect, meaning the data from April and May may behave differently from the data from December to March. However, analyzing Figure 2, it is noted that the cross-validation occurred as expected, since the accuracies of each fold are relatively similar, not indicating any anomalous behavior. Additionally, it is expected that the accuracy of the best fold would be better than that of the final model, since the accuracies of the folds are measured on their own validation sets.

4.2. Feature Importance

The calculation of feature importance is based on a sensitivity analysis of the trained model's loss function concerning each input variable. This was done to better understand the decisions made by the model. The importance of each variable is shown in Figure 4.

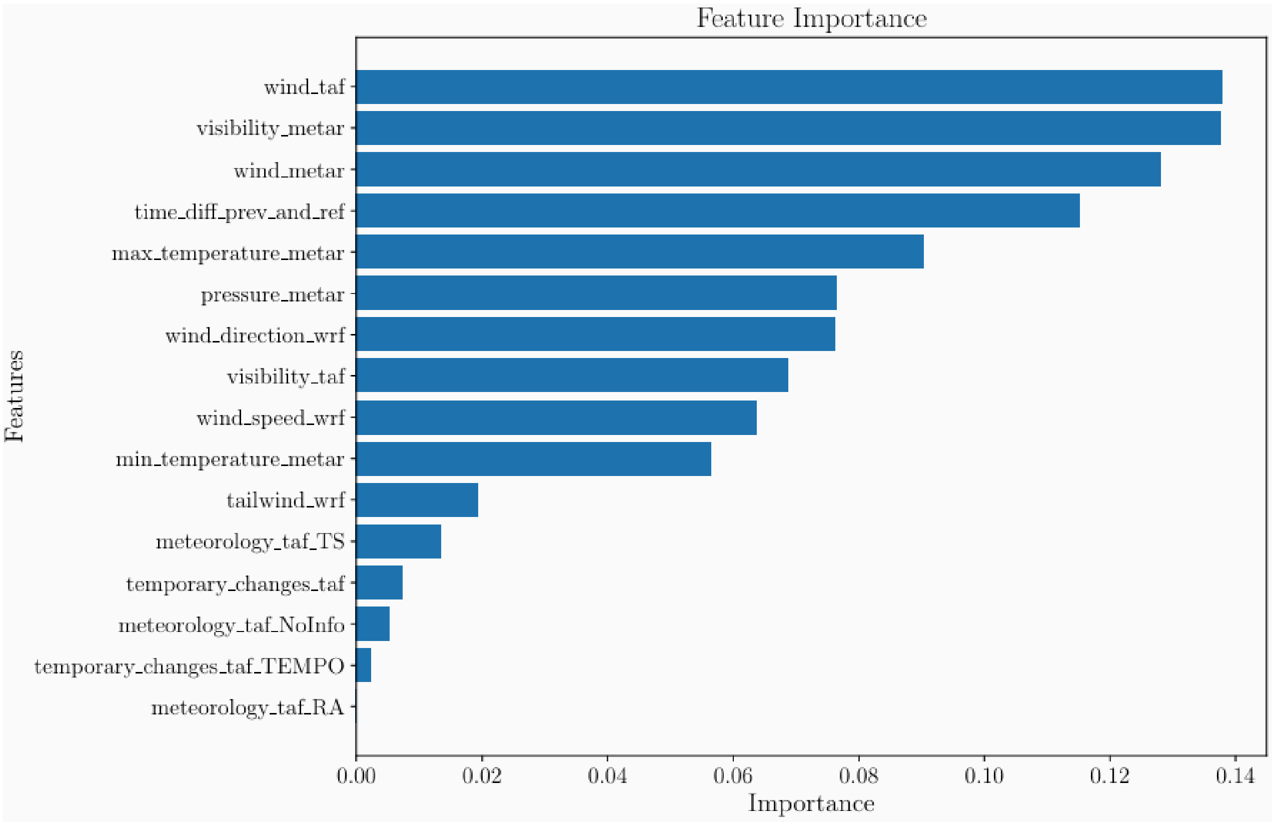

Figure 4: Importance ranking of each feature for the model.

From Figure 4, greater explainability of the developed model can be obtained. The analysis shows that the standout features among those used as input data are the wind from TAF, visibility from METAR, wind from METAR, and $time\_diff\_prev\_and\_ref$. The fact that the METAR and TAF variables proved to be more important justifies the model's better performance compared to related works, given that the model presented in ICEA (2024) uses only WRF data. The importance of the feature $time\_diff\_prev\_and\_ref$, validates the proposed data processing methodology, since this variable is calculated to suggest the relevance of the predicted variables based on the reference time, as explained in Section 2. Furthermore, the fact that the most important feature for the model was the wind from TAF aligns with the ICA 100-37/2020, where wind is one of the main characteristics to be considered for defining the runway in use (DECEA, 2020).

5. CONCLUSION

This study developed and evaluated a predictive model based on LightGBM to forecast runway configuration at Congonhas Airport, using meteorological data from sources such as METAR, TAF, and WRF. The model achieved an accuracy of 98% during training and 88% on an external test dataset, demonstrating its unbiased prediction capability of the runway to be used based on the input data. These results highlight a significant improvement over conventional rule-based methods, which achieved about 81% accuracy, as presented in related works.

The variable importance analysis highlighted wind as a critical factor in determining runway selection, aligning with the ICA 100-37 guidelines and contributing to more efficient and safer airport operations planning. Moreover, the use of advanced machine learning techniques enabled a more detailed and precise analysis of the complex interaction between meteorological variables. Future research could further explore the adaptation and expansion of these models to other airports and operational scenarios, considering different variables and specific conditions of each location.

6. REFERENCES

ACI World. (2023). ACI World Airport Traffic Forecasts 2023–2052. Montreal: Airports Council International. Disponível em: https://store.aci.aero/wp-content/uploads/2024/02/WATF-Executive-Summary.pdf. Acesso em: 3 jun. 2024.

ACI World. (2023). What to expect: Latest air travel outlook reveals short- and long-term demand. Montreal: Airports Council International. Disponível em: https://aci.aero. Acesso em: 3 jun. 2024.

BERGSTRA, J.; BARDENET, R.; BENGIO, Y.; KÉGIL, B. Random Search for hyper-parameter optimization. Journal of Machine Learning Research, JMLR.org, v. 13, n. Feb, p. 281-305, 2012.

DECEA. ICA 100-37/2020 - Serviços de Tráfego Aéreo. Rio de Janeiro, RJ: DECEA, 2020. 147 p. Seção 6.5.

EUROCONTROL. (2015). European Airline Delay Cost Reference Values: Final Report. Dezembro. EUROCONTROL. Disponível em: https://www.eurocontrol.int/sites/default/files/publication/files/european-airline-delay-cost-reference-values-final-report-4-1.pdf. Acesso em: 3 jul. 2024.

FRYE, D.; ZELAZO, P. D.; PALFAI, T. Theory of mind and rule-based reasoning. Cognitive Development. Volume 10, Issue 4, Pages 483-527, October–December 1995.

GOSLING, Geoffrey D. Identification of Artificial Intelligence Applications in Air Traffic Control. Transportation Research Part A: General, v. 21, n. 1, p. 27-38, 1987.

HERREMA, F.; CURRAN, R.; HARTJES, S.; ELLEJMI, M.; BANCROFT, S.; SCHULTZ, M. A Machine Learning Model to Predict Runway Exit at Vienna Airport. Transportation Research Part E: Logistics and Transportation Review, 2019. DOI: 10.1016/j.tre.2019.10.002.

IATA. (2018). Brazil: Value of Aviation. Montreal, QC. Disponível em: https://www.iata.org/en/iata-repository/publications/economic-reports/brazil--value-of-aviation/. Acesso em: 3 jul. 2024.

ICEA. Sistema de Previsão de Reconfiguração de Pista. 2024. Disponível em: https://pesquisa.icea.decea.mil.br/cabeceira/. Acesso em: 12 jun. 2024.

KE, G.; MENG, Q.; FINLEY, T.; WANG, T.; CHEN, W.; MA, W.; YE, Q.; LIU, T. Y. LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, Volume 30, Pages 3146-3154, 2017.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence, Volume 2, Pages 1137-1143, 1995.

LAU, Max En Cheng. Prediction of Runway Configuration Change Transition Timings Using Machine Learning Approach. 2021. Nanyang Technological University. Disponível em: https://hdl.handle.net/10356/150292. Acesso em: 3 jul. 2024.

MIDTFJORD, A. D.; DE BIN, R.; HUSEBY, A. B. A Decision Support System for Safer Airplane Landings: Predicting Runway Conditions Using XGBoost and Explainable AI. Cold Regions Science and Technology, 2021. DOI: 10.1016/j.coldregions.2022.103556.

MURÇA, Mayara Condé Rocha; HANSMAN, Robert John. Identification, Characterization, and Prediction of Traffic Flow Patterns in Multi-Airport Systems. IEEE Transactions on Intelligent Transportation Systems, v. 20, n. 5, p. 1683-1696, 2019.

OLIVEIRA, M. d.; EUFRASIO, A. B. R.; GUTERRES, M. X.; MURÇA, M. C.; ´ GOMES, R. d. A. Analysis of airport weather impact on on-time performance of arrival flights for the brazilian domestic air transportation system. Journal of Air Transport Management, Elsevier, v. 91, p. 101974, 2021

PANDALA, S. Lazypredict: A Tool to Simplify the Training and Comparison of Several Models with a Single Line of Code. v.0.2.10. 2010. Disponível em: https://github.com/shankarpandala/lazypredict. Acessado em: 12-junho-2024.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, Volume 12, Pages 2825-2830, 2011.

PODER360. (2023). Setor aéreo movimentou 112,6 mi de passageiros no país em 2023. Disponível em: https://www.poder360.com.br/infraestrutura/setor-aereo-movimentou-1126-mi-de-passageiros-no-pais-em-2023/. Acesso em: 3 jul. 2024.

PRANDINI, Maria; PIRODDI, Luigi; PUECHMOREL, Stephane; BRÁZDILOVÁ, Silvie Luisa. Toward air traffic complexity assessment in new generation air traffic management systems. IEEE Transactions on Intelligent Transportation Systems, v. 12, n. 3, p. 809 – 818, 2011. DOI: 10.1109/TITS.2011.2113175.

TELES, P.; ZANETTI, M. Predicting Estimated Landing Time. 2024. Disponível em: https://github.com/pedroteles17/predicting-estimated-landing-time