11th International Conference on Electrical, Electronics and Computer Engineering (IcETRAN), Nis, 3-6 June 2024

Comparison of different machine learning approaches, for predicting aircraft departure delays, due to the circumstances of the defrosting process

ME Sanja Milošević, PhD student
Faculty of Transport and Traffic Engineering,
University of Belgrade,
Belgrade, Serbia
sanjav990@gmail.com

PhD čedomir Vasić, associate professor
Information-Technology-FB&L University MB,
Belgrade, Serbia
cedav1961@gmail.com

https://orcid.org/0009-0008-6634-6465

***Abstract*—One of the standard procedures carried out at airports during winter months is the de-icing and anti-icing procedure for aircraft. Circumstances accompanying this procedure such as aircraft type, amount of fluid used, external temperature, wind speed and dew point, directly affect the duration of the procedure and lead to delays as they are conducted immediately before takeoff. A useful tool that could efficiently predict the length of departure delays based on current circumstances, alert airport staff, and indicate the need for additional measures to reduce undesirable delays would be valuable. In this paper, specific data collected at "Konstantin Veliki" Airport in Niš, for the period from November 2020 to March 2024, were processed using two of the most common machine learning methods: regression analysis and classification. The results were compared, and the method that exhibited superior statistical parameters and more accurate prediction was proposed for practical use.**

***Keywords*— defrosting, regression, classification**

I. INTRODUCTION

Flight delays in air traffic are an unwanted occurrence with a significant impact on the finances and reputation of both airports and airlines. Weather conditions, traffic congestion, airport service issues, connecting flights, and other factors contribute to delays. M. Alfarhood [1] examined the usability of several machine learning models using data collected from multiple sources over a period of 6 years, with the CatBoost [2] model achieving the best results in delay prediction. Y. Zhao [2] published a comprehensive study targeting airport service operations and connecting flights as causes of airplane delays. In this study, airport services were modeled using a graphical model (Bayesian network), and then four convolutional neural network models were simultaneously evaluated.

In our work, we started from the fact that the delay we are investigating is solely due to de-icing and anti-icing procedures and the given circumstances under which these procedures are carried out. The concept of a “Clean aircraft” is crucial for smooth operations in low external temperatures. Removing ice and snow from wings, tail, engines, and other critical areas is necessary for flight safety. The de-icing and anti-icing procedure is conducted by ground handling staff, adhering to prescribed regulations. Safety takes priority over regularity and economic aspects that this procedure disrupts. However, if, based on delay predictions, ground handling engages additional available resources and conducts the entire procedure within the turnaround time frame, it would positively impact the airport's reputation and reduce additional costs. In his case-study work, Norin [4] also emphasizes other factors influencing the turnaround time. A good example of delay prediction based on the analysis of meteorological conditions was provided by J. Qu [5].

The author's intention in this work is to explore which modeling approach yields the best results and to propose a practical method for predicting delays in cases where de-icing and anti-icing procedures need to be applied. A good starting point is a highly usable set of input data, enabling the application and comparison of two different machine learning approaches. For each individual delay case, the following data are provided: aircraft type, amount of fluid used, procedure execution time, wind speed, external air temperature, dew point, and delay in minutes. With minimal adjustment of input data necessary to make the learning dataset acceptable for both tested models, an input comprising 108 instances for learning and 20 instances for testing was formed, which is an acceptable ratio confirmed in practice. During the adjustment of input data, the content of information crucial for the result was not compromised, nor was the content essential for the correlation between input data disrupted, only the format in which the data would be input into the model was changed.

The great usability of "multilevel input layer artificial neural network" in predicting airplane delays was demonstrated by Khanmohammadi [6], while Gardner [7] showed with a network based on "multilayer perceptron" that neural networks based on these elements model atmospheric influences very well.

II. OBJECTIVE

In this study, we attempted to determine which of the two basic approaches to machine learning yields better results and more accurate predictions if both approaches are applied to an identical dataset. In other words, we tried to determine whether the prediction problem we are addressing is a regression analysis problem or a classic classification problem .

ISBN 978-86-6200-001-9 ©2024 AM & ETRAN

https://doi.org/10.69994/11Ic24006

The result of regression analysis is a model that determines the functional dependence of one or more dependent variables on one or more independent variables. In our case, the goal is to find a model that best "fits" one variable, which is the delay time, in relation to changes in several independent variables. Since time is a continuous quantity, and our delay time can theoretically have any value from 0 to infinity, it follows that our problem is typically a regression problem. In our specific case, we will seek a solution using multiple linear regression.

If we carefully examine the initial dataset, we notice that the delay time, as the dependent variable, is expressed in minutes. At the same time, the delay time is never longer than 30 minutes, so it follows that our output data can actually be viewed as one element of a finite set of discrete integer values from 0 to 30. The set of integers from 1 to 30 can also be viewed as 30 categories, each of which can be a solution in the classical problem of classifying input data. Classification problems are most elegantly solved today using neural networks.

Based on the above, it is possible to solve the prediction problem we have posed as a task using both regression and classification. For this purpose, we will use a tool that has implemented linear regression and a neural network named: Weka ("Waikato Environment for Knowledge Analysis") ver. 3.8.6. [10]. The software tool Weka has demonstrated great usability in modeling similar prediction problems. To illustrate this point, let's mention just two papers here: Alla [8] and Abubakar [9].

III. Material and Methods

The dataset, consisting of 108 instances of input data, has been uploaded to the Zenodo repository [11], for practical purposes. Since access to this data is in an "open-access" status, the dataset is available for download and inspection (doi: 10.5281/zenodo.11154142).

The input data are packaged in a .arff file, which is prescribed format for inputting data into the Weka software tool. In the header of the .arff file, it is necessary to specify the characteristics of variables using the "@attribute" parameter, and for classification purposes, this header is prepared in the following way:

- @attribute Defrost_lit_total numeric
- @attribute Plain_type {B738, A319, A321, IL62M, C25B, EMB505, A320, ATR72}
- @attribute Month {NOV,DEC,JAN,FEB,MAR,APR}
- @attribute Temperature numeric
- @attribute Wind numeric
- @attribute Dew_point numeric
- @attribute Delay {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30}

First, we will use the prepared .arff file as input for classification using the "MultilayerPerceptron" function. It is evident that the variables "Plain_type" and "Month" are of string type. Similarly to these two independent variables, the dependent variable "Delay," which determines the delay in minutes, also takes one of the values listed in the set of possibilities. After processing using the neural network, the following result was obtained:

=== Run information ===

Instances: 108

Attributes: 7

Defrost_lit_total, Plain_type, Month,Temperature

Wind, Dew_point, Delay

=== Summary ===

Correctly Classified Instances	*101*	*93.5185 %*
Incorrectly Classified Instances	*7*	*6.4815 %*
Kappa statistic	*0.9269*	
Mean absolute error	*0.0235*	
Root mean squared error	*0.086*	
Relative absolute error	*30.013 %*	
Root relative squared error	*43.7175 %*	
Total Number of Instances	*108*	

The next step is modeling regression analysis. The input data had to be adjusted to the condition that regression analysis does not allow a string of characters as an independent variable. For this purpose, it was necessary to create a "look-up" table and replace the aircraft type with an appropriate numerical value. Since the primary surface treated during de-icing and protection is actually the wing surface, and since the wingspan directly corresponds to the size of the aircraft in the data preparation process, we decided to take the wingspan in meters instead of the aircraft type name. Table I shows the parameter values taken during the substitution of input data type regarding the type of aircraft. The replacement of the "Month" variable with an integer value was done by simply entering the ordinal number of the month in the year.

Table I: LOOK-UP TABLE

Airplain Type	Wingspam
A319	34.10
A320	35.80
A321	35.80
B738	34.00
ATR72	27.06
IL62M	43.20
C25B	11.00
EMB505	16.00

The .arff file modified like this was loaded into the Weka tool, and regression analysis was activated using the "LinearRegression" function. After applying this function and fitting the model, the following results were obtained:

=== Run information ===

Attributes: 7

Defrost_lit_total, Plain_type, Month, Temperature

Wind, Dew_poin, Delay

=== Classifier model (full training set) ===

Linear Regression Model

Delay =

*0.0211 * Defrost_lit_total +*

*0.3795 * Plain_type +*

*-0.334 * Month +*

*0.2104 * Wind +*

*0.6289 * Dew_point +*

-8.4838

=== Summary ===

Correlation coefficient	*0.718*
Mean absolute error	*4.2653*
Root mean squared error	*5.0179*
Relative absolute error	*71.3208 %*
Root relative squared error	*69.6038 %*
Total Number of Instances	*108*

Although the quality of classification models is assessed by the relationships between correct and incorrect classifications, which can also be true or false, positive or negative, Weka provides statistically processed results, allowing comparison by observing identical parameters for both observed models.

If we adopt:

x_a - independent variable defined by attribute a,

y_a - dependent variable influenced by same attribute,

$\bar{y}_{as}$ - the mean value of the dependent variable y_a,

n - the number of instances in the input dataset,

we can compare the obtained results by observing the following parameters:

Mean Absolute Error (MAE) as the average value of the absolute differences between the dependent and independent variables (average prediction error).

MAE = – (1)

Root Mean Square Error (RMSE) as the average value of the squared differences between the dependent and independent variables.

RMSE = – (2)

Relative Absolute Error (RAE) is expressed in percentages and represents the ratio of the MAE absolute error to the total absolute deviation of the dependent variable from the average value for all instances.

RAE = (3)

Root Relative Squared Error (RRSE) is expressed in percentages and represents the root of the ratio of the total squared error to the sum of the squares of the deviations from the average values.

RRSE = (4)

The Correlation coefficient and Kappa statistic are parameters present as a result of model analysis. However, they define relationships between independent variables and evaluate the interdependence of input data. Therefore, these parameters are not relevant for comparing the prediction efficiency achieved by the observed models, which is our goal.

From the summarized results after applying both functions, it is immediately noticeable that classification using the “MultilayerPerceptron" function is superior. We will have a better insight into the results if we represent the data in Table II.

IV. Results

Table II shows that according to all parameters for assessing statistical error defined by mathematical formulas (1), (2), (3), and (4), the classification method yields better results.

TABLE II: ERROR ANALYSIS

Error type	**Classification Model**	**Regression Model**
MAE (1)	**0.02**	**4.27**
RMSE (2)	**0.09**	**5.02**
RAE (3)	**30.01%**	**71.32%**
RRSE (4)	**43.72%**	**69.60%**

We will validate this conclusion by testing both models on a test dataset consisting of 20 new instances. These instances were not used during the model learning phase.

Table III presents the parallel prediction results of both models on identical instances, where only the aircraft type and month are listed in the first two columns, and the actual delay in minutes is listed in the third column. The subsequent columns display the results obtained through prediction. Upon calculating the Mean Absolute Error (MAE) for each method, it becomes evident that classification has demonstrated significantly superior results.

TABLE III: COMPARATIVE ERROR ANALYSIS

Aircraft Type	Month	**Real delay**	Classification Model Predicted Delay	Classification Model Absolute Error	Regression Model Predicted Delay	Regression Model Absolute Error
B738	FEB	**10**	10	0	4.9	5.1
A319	NOV	**0**	1	1	4.6	4.6
A320	NOV	**8**	11	3	9.6	1.6
A320	JAN	**12**	12	0	17.4	5.4
B738	JAN	**8**	8	0	10.4	2.4
A319	JAN	**12**	13	1	18.0	6.0
B738	FEB	**17**	12	5	12.7	4.3
ATR72	FEB	**0**	5	5	4.9	4.9
EMB505	MAR	**3**	3	0	-0.5	3.5
A319	FEB	**0**	0	0	7.3	7.3
ATR72	DEC	**0**	0	0	1.5	1.5
A320	JAN	**7**	15	8	7.7	0.7
A319	JAN	**7**	7	0	11.0	4.0
A319	FEB	**24**	19	5	24.2	0.2
ATR72	FEB	**0**	1	1	4.5	4.5
C25B	DEC	**5**	9	4	1.9	3.1
A319	DEC	**14**	10	4	7.4	6.6
A320	JAN	**11**	15	4	10.6	0.4
C25B	JAN	**8**	5	3	1.4	6.6
A319	MAR	**2**	1	1	10.4	8.4
				MAE =2.25		**MAE =4.05**

The graphic representation comparing the results of prediction on test data is shown in Fig. 1 and visually confirms the conclusion made.

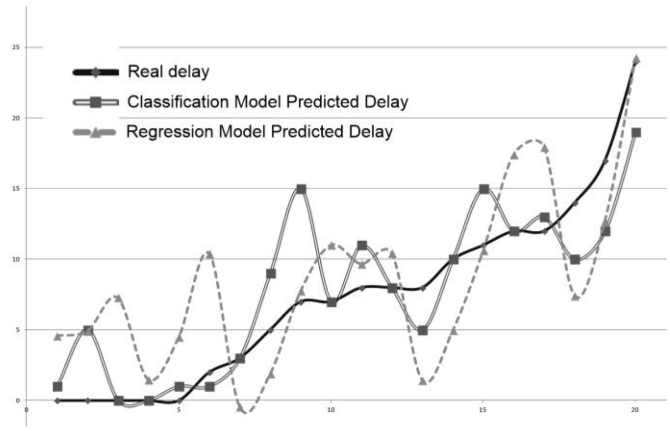

Fig. 1. Comparison of prediction results on test data.

V. CONCLUSION AND DISCUSSION

From the presented results of the analysis, the conclusion arises that for predicting delays in the circumstances outlined in this study, classification using a neural network is an approach that promises better results. The method developed in this study will be recommended for practical implementation. It remains a task to expand the dataset over time and verify the usability of the predictions.

REFERENCES

[1] Meshal Alfarhood, Rakan Alotaibi, Bassam Abdulrahim, Ahmad Einieh, Mohammed Almousa, Abdulrhman Alkhanifer, "Predicting Flight Delays with Machine Learning: A Case Study from Saudi Arabian Airlines", International Journal of Aerospace Engineering, vol. 2024, Article ID 3385463, 12 pages, 2024. https://doi.org/10.1155/2024/3385463.

[2] A. V. Dorogush, V. Ershov, and A. Gulin, “CatBoost: gradient boosting with categorical features support,” 2018, https://arxiv.org/abs/1810.11363. (View at: Google Scholar).

[3] Yang Zhao, Chen Yixin, Hu Jinbiao, Song Yilu, Mao Yi "Departure delay prediction and analysis based on node sequence data of ground support services for transit flights", Transportation Research Part C: Emerging Technologies, Volume 153, August 2023, 104217 DOI: 10.1016/j.trc.2023.104217

[4] Norin, A., et al., (2012). Airport logistics – A Case Study of The Turn-around Process. Journal of Air Transport Management, 20, 31-34

[5] J. Qu, et al., (2020). Flight Delay Prediction Using Deep Convolutional Neural Network Based on Fusion of Meteorological Data. Neural Processing Letters, 52, 1461-1484

[6] Khanmohammadi, S., et al., (2016). A New Multilevel Input Layer Artificial Neural Network for Predicting Flight Delays at JFK Airport. Procedia Computer Science, 95, 237-244.

[7] Gardner, M.W., Dorling, S.R., (1998). Artificial Neural Networks (the Multilayer Perceptron)—a Review of Applications in The Atmospheric Sciences. Atmospheric Environment, 32(14-15), 2627-2636

[8] Alla, H., et al., (2021). A Multilayer Perceptron Neutral Network with Selective-Data Training for Flight Arrival Delay Prediction. Scientific Programming, 2021.

[9] Abubakar, A.U., Tabra, M.S., (2020). Predictive Capabilities of Multilayer Perceptron (MLP) in WEKA Algorithm for High Strength Concrete with Steel Fiber Addition. Journal of Cement Based Composites, 2, 13-18.

[10] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

[11] Vasić, “Data set used for comparison of different machine learning approaches, for predicting aircraft departure delays, due to the circumstances of the defrosting process.”. Zenodo, May 08, 2024. doi: 10.5281/zenodo.11154142.