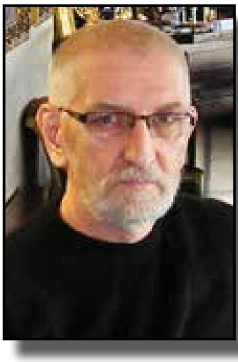
ORIGINAL ARTICLE
DOI: https://doi.org/10.30932/1992-3252-2024-22-1-6

World of Transport and Transportation, 2024,
Vol. 22, Iss. 1 (110), pp. 190–195

Studying the Quality of Airline Customer Service Using Machine Learning Methods



Vyacheslav A. DUKE

Igor G. MALYGIN

***Vyacheslav A. Duke*[1], *Igor G. Malygin*[2]**

[1, 2] Solomenko Institute of Transport Problems of the Russian Academy of Sciences (IPT RAS), St. Petersburg, Russia.
[1] Russian Science Citation Index SPIN-code: 3148-2066.
[2] ORCID: 0000-0002-0403-8095; Web of Science Researcher ID: E-2182–2018; Scopus Author ID: 57159964300; Russian Science Citation Index Author ID: 375896.
✉ *[1] v_duke@mail.ru.*

ABSTRACT

The article presents the results of using machine learning methods to study data from a special questionnaire that considers the general characteristics of flights, the characteristics of passengers and their opinions on various aspects of the flight. The objective of the study is to identify in experimental data factors that negatively affect passengers' attitudes towards airline services.

When conducting the study, well-known algorithms were used that are part of free WEKA (Waikato Environment for Knowledge Analysis) software for data analysis and machine learning by University of Waikato (New Zealand), distributed under the GNU GPL license: naïve Bayes classifier; multilayer perceptron using backpropagation algorithm; k-nearest neighbour method (KNN) with adaptive selection of parameters; decision trees – J48 is an open-source Java implementation of the C4.5 algorithm; random forest; logistic regression; adaptive boosting algorithm (AdaBoost); support vector machine – the SMO (Sequential Minimal Optimization) algorithm which is one of the possible implementations of the support vector machine algorithm.

It is shown that the best accurate models reflecting passenger satisfaction with airline services are obtained using random forest algorithms (error on the test sample is of 3,9 %) and a neural network approach (error on the test sample is of 3,7 %). At the same time, these algorithms do not allow us to explicitly identify factors characteristic of air passengers who are dissatisfied with the quality of service. This gap is filled by an algorithm based on the method of structural resonance in multidimensional data (SRMD), which made it possible to identify precise logical rules in the data with high completeness. The resulting logical rules are highly interpretable patterns of passengers who either negatively or neutrally evaluate the airline's services in general.

Keywords: *air transport, marketing, airline services, quality of passenger service, artificial intelligence, machine learning.*

Funding: *the work was carried out within the framework of implementation of the state assignment of IPT RAS No. 075–00602–24–0 for 2024 and for the planning period of 2025 and 2026 at the expense of funds provided by the Ministry of Education and Science of Russia (an agreement on provision of a subsidy from the federal budget for financial support for implementation of the state assignment dated January 18, 2024 No. 075–03–2024–666).*

For citation: *Duke, V. A., Malygin, I. G. Studying the Quality of Airline Customer Service Using Machine Learning Methods. World of Transport and Transportation, 2024, Vol. 22, Iss. 1 (110), pp. 190–195. DOI: https://doi.org/10.30932/1992-3252-2024-22-1-6.*

The original text of the article in Russian is published in the first part of the issue.
Текст статьи на русском языке публикуется в первой части данного выпуска.

INTRODUCTION

Air passenger service quality assessment plays an important role in the airline industry for several reasons:

1. Understanding what aspects of service are important to passengers allows airlines to identify weaknesses and focus on them to improve service.

2. Measuring service quality helps companies provide better service than their competitors.

3. Good service contributes to passenger satisfaction and, as a result, increases the likelihood that they will be returning clients.

4. Positive reviews about the quality-of-service help attracting new customers. Passenger recommendations often have a big influence on other people's decisions to choose a given airline.

5. Improving service quality helps reduce costs because increased efficiency and improved processes lead to fewer problems and corresponding compensation associated with poor service.

There are several known methods for assessing the quality of service for air passengers:

– Questionnaires and surveys: airlines often use questionnaires and surveys to obtain feedback from passengers. This can be either a paper post-flight questionnaire or an online survey sent by email or accessible through the company web portal.

– Monitoring of social networks: reviews and comments from passengers on various social networks can give an idea of how passengers perceive the airline's service.

– Focus groups: organising focus groups, during which a group of passengers discuss their experience of the service, can provide a deeper understanding of their needs and expectations.

– Mystery passenger: some airlines use «mystery passengers» who travel under the guise of regular passengers, but evaluate the quality of service by testing the service without prior warning.

– Study of data and statistics: analysis of data on flight delays, level of passenger satisfaction, level of complaints and claims can be a useful tool for assessing the quality of service.

– Comparison with competitors: comparing their performance with that of competitors helps airlines understand where they stand in the market in terms of service quality.

The combination of these methods allows airlines to gain a better understanding of how their services are perceived by passengers and where they can improve. There are a few international and domestic studies of quality of passenger service using various methods [1–10]. In our study, we will focus on the survey method using a special questionnaire that considers both the general characteristics of flights and the characteristics of passengers and their opinions on various aspects of the flight.

INITIAL DATA

The data set used contains the results of a questionnaire, which reflects passengers' attitudes towards various aspects of the flight. The data were published by Timothy J. Klein on the popular portal Kaggle[1]. The target variable is «satisfaction», which takes two values: «neutral or dissatisfied» (neutral or negative assessment) and «satisfied» (positive assessment). The original names of other variables (questionnaire items) and interpretation of the values are given in Table 1.

The entire data sample is divided into two parts – a training set (train) and a test data set (test). The training set included the results of a survey of 103904 passengers, and the test set included 25976 passengers. Judging by the description of the data, the airline that provided it wished to remain anonymous. At the same time, these data have attracted the attention of quite many researchers. Reports related to construction of a model explaining passenger satisfaction and dissatisfaction are given on the corresponding Kaggle pages. The most complete and detailed analysis was carried out in Airline Passenger Satisfaction (Part 1) [2]. We will discuss the results of this analysis below and supplement the results obtained with data from our own research.

METHODS USED AND ANALYSIS RESULTS

In Airline Passenger Satisfaction (Part 1)[3] the Scikit-learn package, one of the most widely used Python packages for Data Science and Machine Learning, was used for analysis. Also, a set of standard procedures for statistical univariate and correlation analysis, and several machine learning algorithms were used: k-nearest neighbour method (kNN), support vector

---

[1] Airline Passenger Satisfaction. [Electronic resource]: https://www.kaggle.com/datasets/teejmahal20/airline-passenger-satisfaction. Last accessed 07.01.2024.

[2] Airline Passenger Satisfaction (Part 1) [Electronic resource]: https://www.kaggle.com/code/frixinglife/airline-passenger-satisfaction-part-1/notebook. Last accessed 07.01.2024.

Table 1

Questionnaire items on passenger satisfaction with various aspects of the flight [compiled by the authors based on Airline Passenger Satisfaction[1]]

№	Original item name	Decoding of values*
1.	Gender	Gender: male or female
2.	Customer Type	Customer type: regular or non-regular airline customer
3.	Age	Age: actual age of a passenger
4.	Type of Travel	Type of travel: personal or business trip
5.	Class	Class: business, economy, economy plus
6.	Flight Distance	Flight distance
7.	Inflight Wi-Fi service	Assessment of Wi-Fi connection on board (0: not ranked; 1–5)
8.	Departure/ Arrival time convenient	Departure/arrival time assessment (0: not ranked; 1–5)
9.	Ease of Online booking	Online booking assessment (0: not ranked; 1–5)
10.	Gate location	Assessment of gate location for boarding (0: not ranked; 1–5)
11.	Food and drink	Assessment of food and drink on board (0: not ranked; 1–5)
12.	Online boarding	Assessment of the service of online boarding pass (0: not ranked; 1–5)
13.	Seat comfort	Assessment of seat comfort in the passenger cabin (0: not ranked; 1–5)
14.	Inflight entertainment	Assessment of entertainment on board (0: not ranked; 1–5)
15.	On-board service	Assessment of servicing on board (0: not ranked; 1–5)
16.	Leg room service	Additional comfort for legs (0: not ranked; 1–5)
17.	Baggage handling	Assessment of baggage handling (0: not ranked; 1–5)
18.	Checking service	Assessment of checking service (0: not ranked; 1–5)
19.	Inflight service	Assessment of servicing on board during flight (0: not ranked; 1–5)
20.	Cleanliness	Assessment of cleanliness on board (0: not ranked; 1–5)
21.	Departure Delay in Minutes	Departure delay in minutes
22.	Arrival Delay in Minutes	Arrival delay in minutes

* The values were also decoded in Russian as shown for the use in the survey held by the authors.

machine (SVM), AdaBoost algorithm, decision trees (DT) and random forest. The following conclusions have been drawn:

– For kNN method (with k = 10), the error on the training set was 5,2 %, and on the test set – 6,5 %.

– For the «support vector machine» the error on both the training and test sets was 5 %.

– For the «random forest» classifiers, no error was observed on the training set, but on the test set it was 3,9 %.

– For AdaBoost algorithm, the error on the training set was 7 %, and on the test set – 7,2 %.

– For «decision trees with gradient boosting» the error on the training and test samples was 5,5 % and 5,6 %, respectively.

Thus, the best result was shown by the «random forest» classifier (error on the test sample – 3,9 %).

In addition to the analysis performed, in the second part of the study[3], a neural network was trained, which demonstrated an error on the test data set of 3,7 % (training the neural network took almost seven hours).

Regarding the above analysis, let us make several clarifications and comments.

Firstly, in the cited source Airline passenger satisfaction (Part 2)[4], the researcher made a test sample from the training sample (train.csv file), using 90 % for training and 10 % for testing. This, in our opinion, is not of fundamental importance, since the volume of the entire data is large enough to obtain stable models.

Secondly, there were gaps in the original data, coded as «0», and here the researcher filled the gaps with the median values of the features. In our opinion, filling in gaps in data should be treated very carefully. This can be useful only in the case of small samples, and even then, only under the condition of fairly simple data structures that have single-mode distribution densities of values.

Thirdly, and this is the most important thing, in the above study (as well as in a few others published on the Kaggle portal), in our opinion, the wrong emphasis was placed in statement of the problem itself. Here, attempts are made to build the most accurate model possible, connecting the characteristics of flights, passengers and their assessments of particular service characteristics, but the main goal is

[3] Airline Passenger Satisfaction. (Part 2). [Electronic resource]: https://www.kaggle.com/code/frixinglife/airline-passenger-satisfaction-part-2. Last accessed 07.01.2024.

Table 2

Summary table of model building results using various machine learning methods [performed by the authors]

Method	Error in the model in %	Time to create a model, sec	Time to apply a model, sec
Naïve Bayes	15,48	0,48	0,43
Logistic regression	12,85	2,68	0,19
SMO	12,64	1937,76	0,27
KNN (10-NN)	7,42	0,07	249,75
Multilayer perceptron	4,35	218,33	0,24
Decision trees J48	4,24	11,44	0,19
AdaBoost (J48)	4,12	128,02	0,32
Random forest	3,78	47,85	2,06

missing – to find out as accurately as possible the reasons why passengers are dissatisfied with the airline's services.

We tried to fill this gap by repeating, on the one hand, data analysis using another library of machine learning programs. On the other hand, we complement the results of our research with high-precision patterns characteristic of a group of dissatisfied and neutral air passengers, which are identified using our innovative SRMD (Structural Resonance in Multidimensional Data) technology [11], developed by Deep Patterns[4].

In our study, we used popular algorithms included in WEKA (Waikato Environment for Knowledge Analysis) free software for data analysis and machine learning, developed by University of Waikato (New Zealand) and distributed under the GNU GPL license [12]:

– naïve Bayes classifier;
– multilayer perceptron using an error backpropagation algorithm;
– k-nearest neighbour method (KNN);
– decision trees;
– random forest;
– logistic regression;
– adaptive boosting algorithm (AdaBoost);
– support vector machine (SVM).

When using these methods, the default parameters set in the WEKA package were mainly applied. However, some clarifications should be made. As one of the possible implementations of the support vector machine algorithm, the SMO (Sequential Minimal Optimization) algorithm described in [13] was used. In this case, a linear kernel was used. In a multilayer perceptron, the number of layers was determined by the formula (number of variables + number of classes) / 2, so the number of layers was 12. We also note that when constructing decision trees, the J48 algorithm was used, which is a Java analogue of the well-known C4.5 algorithm [14]. The AdaBoost algorithm uses decision trees built by J48 as classifiers. We previously considered examples of the use of these algorithms in the transport industry in articles [15; 16].

[4] https://deeppatterns.com.

Table 3

Highly accurate logic rules found in data of air passenger questionnaire [performed by the authors]

№	Rule	Recall	Accuracy
1.	**If** (not business class) **AND** (Assessment of Wi-Fi connection on board) <= 3 **AND** (Assessment of ease of online booking) <= 4 **AND** (Assessment of online boarding pass service) <= 3 **Then** (Overall assessment is neutral or negative)	0,538	0,986
2.	**If** (not business class) **AND** (Assessment of Wi-Fi connection on board) <= 3 **AND** (Assessment of ease of online booking) <= 3 **AND** (Assessment of servicing during flight) > 2 **Then** (Overall assessment is neutral or negative)	0,467	0,99
3.	**If** (personal trip) **AND** (Assessment of Wi-Fi connection on board) <= 3 **Then** (Overall assessment is neutral or negative)	0,431	1
4.	**If** (personal trip) **AND** (assessment of ease of online booking) <= 3 **Then** (Overall assessment is neutral or negative)	0,402	0,992
5.	**If** (personal trip) **AND** (Assessment of Wi-Fi connection on board) <= 4 **AND** (Assessment of ease of online booking) <= 4 **AND** (Assessment of online boarding pass service) <= 3 **AND** (Assessment of additional comfort for legs) > 0 **Then** (Overall assessment is neutral or negative)	0,357	0,989
6.	**If** (Assessment of Wi-Fi connection on board) <= 2 **AND** (Assessment of gate location for boarding) > 2 **Then** (Overall assessment is neutral or negative)	0,357	0,987
7.	**If** (age) <= 35 **AND** (not business class) **AND** (Assessment of Wi-Fi connection on board) <= 3 **Then** (Overall assessment is neutral or negative)	0,315	0,994
8.	**If** (Assessment of Wi-Fi connection on board) <= 3 **AND** (assessment of departure/arrival delay) > 3 **AND** (Assessment of ease of online booking) <= 3 **AND** (Assessment of servicing on board during flight) > 2 **Then** (Overall assessment is neutral or negative)	0,267	0,991
9.	**If** (Assessment of Wi-Fi connection on board) <= 2 **AND** (Assessment of ease of online booking) <= 2 **AND** (Assessment of gate location for boarding)> 2 **Then** (Overall assessment is neutral or negative)	0,263	0,996

Table 2 shows the results of applying the mentioned methods to build a model for predicting the values of the target variable «Satisfaction». In addition to model errors calculated on the test sample (test.csv), the Table 2 shows the time spent on creating and applying the model.

Overall, the results obtained are very similar to those obtained in Airline Passenger Satisfaction. (Part 1)[3]. At the same time, we note the large (almost fourfold) difference between the accuracy of the «naïve Bayes classifier» and the maximum accuracy achieved using the «random forest» method. This phenomenon is typical for heterogeneous data structures that cannot be adequately described by a general model in a simple (for example, linear) interpretable form. However, more complex models provide high accuracy, but are poorly interpretable, if at all. For example, the «decision tree» built using the J48 algorithm, in our case, has 1378 leaves, which creates problems for formation of a holistic perception and understanding of the relationships identified in the data.

The method of structural resonance in multidimensional data (SRMD), which we applied at the next stage of analysis, is characterised by the fact that it is aimed at searching in the data for logical «if–then» rules that, for a given accuracy, have the maximum recall of objects of their own class. This property of SRMD ensures good interpretability of data analysis results. Additionally, it is worth noting that SRMD does not require any artificial filling of gaps in the data table – the gaps are simply not processed.

Table 3 shows some high-precision logical rules found in experimental data from a questionnaire on passenger satisfaction with airline services.

The rules given in Table 3 with high accuracy (the error in the aggregate is 1,9 %) cover 82 % of air passengers who gave a negative or neutral assessment to the airline's services. Moreover, most of them (57 %) are a group of people making personal (non-business) travel. And, in turn, almost all these 82 % express dissatisfaction due to poor quality of Wi-Fi connection on board the aircraft. Apparently, the airline that conducted the survey could improve passenger satisfaction quite significantly by increasing the quality of this in-flight service. In addition, an obvious (based on the data in Table 3) resource for improving the quality of service lies in improving the ease of procedures for online booking and passenger online check-in.

CONCLUSION

1. Machine learning methods make it possible to obtain models that link the general characteristics of flights, the characteristics of passengers and their assessments of variety of particular services with overall satisfaction of passengers with a flight.

2. The models with the best accuracy were built using the random forest algorithms (error on the test sample was of 3,9 %) and the neural network approach (error on the test sample was of 3,7 %). At the same time, these algorithms do not allow us to explicitly identify factors characteristic of passengers who either negatively or neutrally evaluate the entire airline's services.

3. Precise logical rules with sufficiently high recall completeness, which are patterns of passengers who either negatively or neutrally evaluate the entire airline's services, were identified in the data using an algorithm based on the «structural resonance» method in multidimensional SRMD data.

4. From the identified patterns for the case considered, it follows that the airline can significantly increase passenger satisfaction by increasing the quality of Wi-Fi service in the aircraft cabin. In addition, a significant resource for improving the quality of service for air passengers lies in improving online booking and check-in procedures.

REFERENCES

1. Dike, S. E., Davis, Z., Abrahams, A., Anjomshoae, A., Ractham, P. Evaluation of passengers' expectations and satisfaction in the airline industry: an empirical performance analysis of online reviews. *Benchmarking: An International Journal*, 2024, Vol. 31, Iss. 2, pp. 611–639. DOI: 10.1108/BIJ-09-2021-0563.

2. Fodness, D., Murray, B. Passengers' expectations of airport service quality. *Journal of Services Marketing*, 2007, Vol. 21, Iss. 7, pp. 492–506. DOI: 10.1108/08876040710824852.

3. Ban, H.-J., Kim, H.-S. Understanding Customer Experience and Satisfaction through Airline Passengers' Online Review. *Sustainability*, 2019, Vol. 11(15), pp. 1–17. DOI: 10.3390/su11154066.

4. Awadh, M. Assessing the Quality of Sustainable Airline Services Utilizing the Multicriteria Decision-Making Approach. *Sustainability*, 2023, Vol. 15(9), pp. 1–19. DOI: 10.3390/su15097044.

5. Namukasa, J. The influence of airline service quality on passenger satisfaction and loyalty: the case of Uganda airline industry. *The TQM Journal*, 2013, Vol. 25, Iss. 5, pp. 520–532. DOI: 10.1108/TQM-11-2012-0092.

6. Tahanisaz, S., Shokuhyar, S. Evaluation of passenger satisfaction with service quality: A consecutive method applied to the airline industry. *Journal of Air Transport Management*, 2020, Vol. 83, pp. 101764. DOI: 10.1016/j.jairtraman.2020.101764.

7. Tsaur, S.-H., Chang, T.-Y., Yen, C.-H. The evaluation of airline service quality by fuzzy MCDM. *Tourism Management*, 2002, Vol. 23 (2), pp. 107–115. DOI: 10.1016/S0261-5177(01)00050-4.

8. Ayriev, R. S., Kudryashov, M. A. Quality Indices of Public Transportation Services. *World of Transport and Transportation*, 2018, Vol. 16, Iss. 4 (77), pp. 140–149. DOI: https://doi.org/10.30932/1992-3252-2018-16-4-11.

9. Matantseva, O. Y., Aredova, A. K., Shchegoleva, I. V. Study of the Influence of Factors on Passenger Service Quality and Efficiency of Rolling Stock Use. *World of Transport and Transportation,* 2022, Vol. 20, Iss. 4 (101), pp. 98–104. DOI: https://doi.org/10.30932/1992-3252-2022-20-4-8.

10. Sokolov, Yu. I. Service Quality Should be Assessed by the Clients Themselves. *World of Transport and Transportation*, 2015, Vol. 13, Iss. 4 (59), pp. 100–109. [Electronic resource]: https://mirtr.elpub.ru/jour/article/view/490/761. Last accessed 07.01.2024.

11. Duke, V. A. Logical methods of machine learning (tools and practical examples) [*Logicheskie metody mashinnogo obucheniya (instrumentalnie sredstva i prakticheskie primery)*]. St. Petersburg, Publishing and Printing Association of Higher Educational Institutions, 2020, 248 p. ISBN 978-5-91155-087-5.

12. Frank, E., Hall, M. A., Witten, I. H. The WEKA Workbench. Online Appendix for «Data Mining: Practical Machine Learning Tools and Techniques», Morgan Kaufmann, Fourth Edition, 2016. [Electronic resource]: https://www.scirp.org/reference/referencespapers?referenceid=2477648. Last accessed 07.01.2024.

13. Platt, C. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. In: Advances in Kernel Methods – Support Vector Learning, ed. by B. Schölkopf and C. J. C. Burges and A. J. Smola. Cambridge, MA, MIT Press. 1999, pp. 185–208. [Electronic resource]: https://www.researchgate.net/publication/2624239_Sequential_Minimal_Optimization_A_Fast_Algorithm_for_Training_Support_Vector_Machines. Last accessed 07.01.2024.

14. Quinlan, J. R. C4.5 Programs for Machine Learning, San Mateo, CA, Morgan Kaufmann, 1992, 302 p. ISBN 978-1558602380.

15. Duke, V. A., Malygin, I. G. Comparison of algorithms for recognition of vehicle types by parameters of their silhouettes. *Marine intellectual technologies*, 2018, Vol.4, Iss. 4 (42), pp. 197–201. EDN: YXSDNR.

16. Duke, V. A., Malygin, I. G. Comparative study of machine learning algorithms in the problem of predicting the dynamics of bike sharing [*Sravnitelnoe issledovanie algoritmov mashinnogo obucheniya v zadache prognozirovaniya dinamiki velosheringa*]. *Transport: science, technology, management. Scientific information collection*, 2023, Iss. 6, pp. 40–45. EDN: BVMMSV. ●

Information about the authors:

Duke, Vyacheslav A., *D.Sc. (Eng), Chief Researcher at the Laboratory of Intelligent Transport Systems of Solomenko Institute of Transport Problems of the Russian Academy of Sciences (IPT RAS), St. Petersburg, Russia, v_duke@mail.ru.*

Malygin, Igor G., *D. Sc. (Eng), Professor, Director of Solomenko Institute of Transport Problems of the Russian Academy of Sciences (IPT RAS), St. Petersburg, Russia, malygin_com@mail.ru.*

Article received 01.02.2024, approved 28.02.2024, accepted 15.03.2024.