

Flight Fare Prediction using Random Forest Algorithm

Prof. Ms. Archana Dirgule¹, Shubham Agarwal², Ram Agrawal³, Neha Singh⁴, Kiran Adsul⁵

Faculty, Department of Computer Engineering¹
Students, Department of Computer Engineering^{2,3,4,5}
Sinhgad College of Engineering, Pune, Maharashtra, India

Abstract: *Accurate flight delay prediction is fundamental to establish the more efficient airline business. Recent studies have been focused on applying machine learning methods to predict the flight delay. Most of the previous prediction methods are conducted in a single route or airport. This paper explores a broader scope of factors which may potentially influence the flight delay, and compares several machine learning-based models in designed generalized flight delay prediction tasks. To build a dataset for the proposed scheme, automatic dependent surveillance broadcast (ADS-B) messages are received, pre-processed, and integrated with other information such as weather condition flight schedule, and airport information. The designed prediction tasks contain different classification tasks and a regression task. Experimental results show that long short-term memory (LSTM) is capable of handling the obtained aviation sequence data, but overfitting problem occurs in our limited dataset. Compared with the previous schemes, the proposed random forest-based model can obtain higher prediction accuracy (90.2% for the binary classification) and can overcome the overfitting problem.*

Keywords: Flight delay prediction, ADS-B, machine learning, LSTM neural network, random forest.

REFERENCES

- [1]. J. Stavins, "Price discrimination in the airline market: The effect of market concentration," *Review of Economics and Statistics*, vol. 83, no. 1, pp. 200–202, 2001.
- [2]. B. Mantin and B. Koo, "Dynamic price dispersion in airline markets," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 6, pp. 1020–1029, 2009.
- [3]. P. Malighetti, S. Paleari, and R. Redondi, "Has ryanair's pricing strategy changed over time? an empirical analysis of its 2006–2007 flights," *Tourism Management*, vol. 31, no. 1, pp. 36–44, 2010.
- [4]. T. H. Oum, A. Zhang, and Y. Zhang, "Inter-firm rivalry and firm-specific price elasticities in deregulated airline markets," *Journal of Transport Economics and Policy*, vol. 7, no. 2, pp. 171–192, 1993.
- [5]. B. Burger and M. Fuchs, "Dynamic pricing – A future airline business model," *Journal of Revenue and Pricing Management*, vol. 4, no. 1, pp. 39–53, 2005.
- [6]. T. M. Vowles, "Airfare pricing determinants in hub-to-hub markets," *Journal of Transport Geography*, vol. 14, no. 1, pp. 15–22, 2006.
- [7]. K. Rama-Murthy, "Modeling of united states airline fares– using the official airline guide (OAG) and airline origin and destination survey (DB1B)," Ph.D. dissertation, Virginia Tech, 2006.
- [8]. B. Derudder and F. Witlox, "An appraisal of the use of airline data in assessing the world city network: a research note on data," *Urban Studies*, vol. 42, no. 13, pp. 2371–2388, 2005.
- [9]. A. Mottini and R. Acuna-Agost, "Deep choice model using pointer networks for airline itinerary prediction," in the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 2017, pp. 1575–1583.
- [10]. K. Tziridis, T. Kalampokas, G. A. Papakostas, and K. I. Diamantaras, "Airfare prices prediction using machine learning techniques," in the 25th IEEE European signal processing conference, 2017, pp. 1036–1039.
- [11]. Y. Chen, J. Cao, S. Feng, and Y. Tan, "An ensemble learning based approach for building airfare forecast service," in the IEEE international conference on big data, 2015, pp. 964–969.

- [12]. T. Liu, J. Cao, Y. Tan, and Q. Xiao, "ACER: An adaptive context-aware ensemble regression model for airfare price prediction," in the international conference on progress in informatics and computing, 2017, pp. 312–317.
- [13]. V. Pai, "On the factors that affect airline flight frequency and aircraft size," *Journal of Air Transport Management*, vol. 16, no. 4, pp. 169–177, 2010.
- [14]. M. S. Ryerson and H. Kim, "Integrating airline operational practices into passenger airline hub definition," *Journal of Transport Geography*, vol. 31, pp. 84– 93, 2013.
- [15]. H. Baik, A. A. Trani, N. Hinze, H. Swingle, S. A