Air Ticket Seating Availability With A Structure For Airline Price Prediction Using Machine Learning Approach

Pradeepa V[1], Senthilraja P[2]
[1]Dept of Computer Science and Engineering
[2]Assistant Professor, Dept of Computer Science and Engineering
[1, 2] K.S. Rangasamy college of technology,Tiruchengode,Tamilnadu

Abstract- *Airlines use sophisticated quasi-academic tactics which airlines call "revenue management" or "yield management". The cheapest available ticket on a given flight gets more and less expensive over time depends upon the purchase pattern. This usually happens as a attempt to maximize revenue based on time of purchase patterns. The last minute purchase tickets are more expensive when compare to normal flight tickets. Due to the quick turn of events and wide use of Web innovation, online buys have gotten the fundamental strategy to purchase airfare tickets. Since data about valuing factors is inaccessible, it is hard for customers to purchase airfare tickets at the most minimal expense or even at a generally lower cost.The concept assists with knowing the flow cost of the air tickets, seat accessibility and further more to predict the future air fare price. More specifically, this system is designed to late arriving customers, can save their seats before close to their departure date. The concept is an AI enabled web platform based application. It servers as the day to day companion of Airport Authority individuals. It can intelligently set the fare of flight tickets, can see the seating arrangements and also provide environmental friendly and saving solution.*

Keywords- machine learning,airfare price,Random forest algorthim

I. INTRODUCTION

Airline stend to charge high prices to passengers who search for tickets close to their date of travel.The conventional view is that these are business travellers and that airlines capture their high willingness to pay through price discrimination. Aircrafts additionally change costs on an everyday premise as limit is restricted and the future interest for some random flight is uncertain. They may adjust fare upward to avoid selling out flights in advance, or fares may actually from one day to the next, after a sequence of flow demand realization.

In this concept new flight-level dataset consisting of daily fare and seat availabilities to estimate a model of dynamic airline pricing in which firms face as to chastic arrival of customers.This model to quality the welfare effects of dynamic airline pricing but also to establish important interactions between in markets.The estimate structural model allows to establish key points about the interaction between the pricing forces. dynamic adjustment complements intertemporal price discrimination in the airline industry.This is due to the fact that price inelastic consumers tend to buy tickets close to the departure date.To be able to price discrimated towards these late arriving consumers airlines must successfully save seats until close to their departure date.

II. LITERATURE REVIEW

It is difficult to customer for purchasing their flight ticket in lowest price. For that several techniques are used to obtain the day a which price of air ticket will be minimum.Most of these researchers are used Artifical Intelligence model.Adachar [1] presented total delay minimization and the reduction of the pollution emission..Lei Yang[16] presented a CTM-based network model shown its efficient and accurate simulation methods and air traffic prediction.Yicheng Zhabg[18] presented a distributed strategy can solve fairly large routing and scheduling problems by using Airport traffic flow management.Yuchang[4] had been applied a superior decision arbitrary perception for customers to purchase a air tickets.Zaimingful[5] utilized the Interchangeable virtual Instruments for framework to talk about the quick and savvy programmed boundaries.Abhijit Borah[2] tried to predict the flight fares using kalman filter which was a famous Bayesian estimation techniques.Yongli Wang [13]explained about dynamic evaluating issue thinking about buyer conductset up the game mobel between retailer and customer by using dynamic valuing model.Sachine kumar[7] presented a machine learning approach to analyse the tweets to improve customer experience.Maria Virgina Caccavale[3] described the airbound air traffic over a congested by using PSRA model.

III. DATA COLLECTION

The data collection is most aspect of this project. The different websites have various data are avaiable used to train the models.For viewing the current price ,seat availability, future fare the data sets are taken from kaggle and the models are implemented using python.

3.1 Data collection

The python is take out the various information from the datasets. In the collected data sets contains various parameters such as date of journey, name of the flight, source time, desination time, additional info, stops and prices are available in the data set.

	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
2	IndiGo	24/03/2019	Banglore	New Delhi	BLR → DEL	22:20	01:10 22 Ma	2h 50m	non-stop	No info	3897
3	Air India	1/05/2019	Kolkata	Banglore	CCU → IXR → BBI → BLR	05:50	13:15	7h 25m	2 stops	No info	7662
4	Jet Airways	9/06/2019	Delhi	Cochin	DEL → LKO → BOM → C	09:25	04:25 10 Jur	19h	2 stops	No info	13882
5	IndiGo	12/05/2019	Kolkata	Banglore	CCU → NAG → BLR	18:05	23:30	5h 25m	1 stop	No info	6218
6	IndiGo	01/03/2019	Banglore	New Delhi	BLR → NAG → DEL	16:50	21:35	4h 45m	1 stop	No info	13302
7	SpiceJet	24/06/2019	Kolkata	Banglore	CCU → BLR	09:00	11:25	2h 25m	non-stop	No info	3873
8	Jet Airways	12/03/2019	Banglore	New Delhi	BLR → BOM → DEL	18:55	10:25 13 Ma	15h 30m	1 stop	In-flight meal nc	11087
9	Jet Airways	01/03/2019	Banglore	New Delhi	BLR → BOM → DEL	08:00	05:05 02 Ma	21h 5m	1 stop	No info	22270
10	Jet Airways	12/03/2019	Banglore	New Delhi	BLR → BOM → DEL	08:55	10:25 13 Ma	25h 30m	1 stop	In-flight meal nc	11087
11	Multiple carriers	27/05/2019	Delhi	Cochin	DEL → BOM → COK	11:25	19:15	7h 50m	1 stop	No info	8625
12	Air India	1/06/2019	Delhi	Cochin	DEL → BLR → COK	09:45	23:00	13h 15m	1 stop	No info	8907
13	IndiGo	18/04/2019	Kolkata	Banglore	CCU → BLR	20:20	22:55	2h 35m	non-stop	No info	4174
14	Air India	24/06/2019	Chennai	Kolkata	MAA → CCU	11:40	13:55	2h 15m	non-stop	No info	4667
15	Jet Airways	9/05/2019	Kolkata	Banglore	CCU → BOM → BLR	21:10	09:20 10 Ma	12h 10m	1 stop	In-flight meal nc	9663
16	IndiGo	24/04/2019	Kolkata	Banglore	CCU → BLR	17:15	19:50	2h 35m	non-stop	No info	4804
17	Air India	3/03/2019	Delhi	Cochin	DEL → AMD → BOM → C	16:40	19:15 04 Ma	26h 35m	2 stops	No info	14011
18	SpiceJet	15/04/2019	Delhi	Cochin	DEL → PNQ → COK	08:45	13:15	4h 30m	1 stop	No info	5830
19	Jet Airways	12/06/2019	Delhi	Cochin	DEL → BOM → COK	14:00	12:35 13 Jur	22h 35m	1 stop	In-flight meal nc	10262
20	Air India	12/06/2019	Delhi	Cochin	DEL → CCU → BOM → C	20:15	19:15 13 Jur	23h	2 stops	No info	13381
21	Jet Airways	27/05/2019	Delhi	Cochin	DEL → BOM → COK	16:00	12:35 28 Ma	20h 35m	1 stop	In-flight meal nc	12898
22	GoAir	6/03/2019	Delhi	Cochin	DEL → BOM → COK	14:10	19:20	5h 10m	1 stop	No info	19495
23	Air India	21/03/2019	Banglore	New Delhi	BLR → COK → DEL	22:00	13:20 19 Ma	15h 20m	1 stop	No info	6955
24	IndiGo	3/04/2019	Banglore	Delhi	BLR → DEL	04:00	06:50	2h 50m	non-stop	No info	3943
25	IndiGo	1/05/2019	Banglore	Delhi	BLR → DEL	18:55	21:50	2h 55m	non-stop	No info	4823

figure**1:**Collected Data set

figure 1 shows the original data set that taken from kaggle websites for predicting the future fare.

3.2 Cleaning and Preparing the data

After collecting the data,cleaning unwanted data from the data set and prepare the data for the further process.cleanind the data is a most important process and time consuming process.In this cleaning and preparing process different logic in python to clean and set up the information in correct manner.

3.3 Analyzing data

Followed by cleaning process data analyzing is occur by uncovering the hidden data trends and then applying various machine learning approaches.Total departure days is calculated by difference between the departure date and the source date.In this Analyzing data times play an important role for calculating prices.

IV. SYSTEM MODULES

The module description is a high-level description of a functional area,consisting of a group of processes describing the functionality of the module and a group of packages implementing the functionality.

4.1 Flight status and current price

In this concept three different server tasks to achieve all three customer demands.Flight status and current price is the first module. This module provide the detailed information about the current price of the flight. It provides the information like flight name,source, departure time, total stops between source and destination, Date of journey, Duration details.In this module there are two sides client and server side.The client side user can check the price of the flight based upon their requirements.

The request is passed to the server side with help of Goibibo dataset user can check the current price of the flight.The work in background to fetch user required data on the front end the machine learning learning module is converted in to Flask API which gives the current price of the price.From this module user can view their current price of the flight without any hesitate.

4.2 Future fare prediction

In this module used can predict the future fare. It helps to plan their journey based up on the fare. The machine learning modules is converted to Flask API which gives the predicted ticket price.The flight status data is fetched from database from kaggle and all of the instances and databases for the airport monitoring system has been served and hosted on Google cloud and firebase.

Flask is a web framework for Python, meaning that it provides functionality for building web Application including managing HTTP requests and rendering templates .In this sub module, it helps to predict the future fare of the air ticket by using Flash Server Application Programming Interface. This database fetch the Client request and to predict the airfare. For predicting the future fare Random forest algorithm is used to predict the future fare.

4.3 Seat Availablity

This is the second module. It shows the seat availability of the required flight based up on the customer needs.It helps to check the seats availability prior to the customer journey.In this module like first module it also has two sides one is client and server side.

In the client side user send the request to the server side.In the server side has database runs it gives the required output to the customers without any difficulties customer can receive their required data.From this module customer user can book their required seats without any difficulties.

V. CONCLUSION AND FUTURE WORK

In this study a machine learning structured was developed to predict the future fare price, can see the seat availability and see the current fare of flights in the domestic airlines. The machine learning model is working in the back-end to support for future prediction. It is a multi purpose application is versatile in usage. It also built using multiple cross compiled language. Data is stored on cloud to avoid any loss. This web can be very hard in avitation industries. Based upon this concept user can view the seat availabilities ,current and future fare tickets. In future this concept will be implemented in all National and International airline companies and also event information will also collected from various source, which includes social media and news agencies as to complemented the prediction model.

VI. ACKNOWLEDGMENT

I am very proudly rendering our thanks to our principal **Dr.R.GOPALAKRISHNAN, M.E., Ph.D.,** for the facilities and the encouragement given by him to the progress and completion of our project.

I proudly render our immense gratitude to the Head of the Department **Dr.S.MADHAVI M.E., Ph.D.,** for her effective leadership, encouragement and guidance in the project.

I am highly indebted to provide our heart full thanks to our guide Assistant Professor **Dr.P.SENTHILRAJA M.E., Ph.D,** for his valuable ideas, encouragement and supportive guidance throughout the project.

I am extremely grateful to my project co-ordinator Professor **Dr.A.GNANABASKARAN M.E., Ph.D.,** for his heartfelt involvement.

I wish to extend our sincere thanks to all faculty members of our Computer Science and Engineering Department for their valuable suggestions, kind co-operation and constant encouragement for successful completion of this project.

I wish to acknowledge the help received from various Departments and various individuals during the preparation and editing stages of the manuscript.

REFERANCES

[1] Adacher, L., Flamini, M. and Romano, E., (2018). "Airport ground movement problem: minimization of delay and pollution emission." IEEE Transactions onIntelligent Transportation Systems, vol.19, no.12, pp.3830-3839

[2] Boruah, A., Baruah, K., Das, B., Das, M.J. and Gohain, N.B., (2019)." A Bayesian Approach for Flight Fare Prediction Based on Kalman Filter." In progress inAdavancedComputing and Intelligent Engineering ,pp. 191-203, Springer, Singapore.

[3] Caccavale, M.V., Iovanella, A., Lancia, C., Lulli, G. and Scoppola, B., (2014)." A model of inbound air traffic: The application to Heathrow airport. "Journal of Air Transport Management, vol.34, pp.116-122,Elsevier.

[4] Cai, K.Q., Zhang, J., Xiao, M.M., Tang, K. and Du, W.B., (2017). "Simultaneous optimization of airspace congestion and flight delay in air traffic network flow management." IEEETransactions onIntelligent Transportation Systems, vol.18,no.11, pp.3072-3082.

[5] Fu, Z., Ren, N., Ke, L. and Zhao, Y., (2018). "Research on fast and intelligent calibration method based on automatic test system." In 2018IEEEAUTOTESTCON pp. 1-5. IEEE.

[6] Keimer, A., Laurent-Brouty, N., Farokhi, F., Signargout, H., Cvetkovic, V., Bayen, A.M. and Johansson, K.H., (2018). "Information patterns in the modeling and design of mobility management services". Proceedings of the IEEE, vol.106,no.4, pp.554-576.

[7] Kumar, S. and Zymbler, M., (2019). "A machine learning approach to analyze customer satisfaction from airline tweets."Journal of Big Data, vol.6,no.1, p.62,Springer.

[8] Lancia, C., Guadagni, G., Ndreca, S. and Scoppola, B., (2018)." Asymptotics for the late arrivals problem." Mathematical Methods of Operations Research, vol.88, *no.*3, pp.475-493,Springer.

[9] Li, Q., Zhang, Y. and Su, R., (2016). "A flow-based flight scheduler for en-route air traffic management."IFAC-apersOnLine, vol.49,no.3, pp.353-358,Elsevier.

[10] Martínez, M.E.A., Navarro, J.L.A. and Trinquecoste, J.F., (2017). "The effect of destination type and travel period on the behavior of the price of airline tickets."Research in Transportation Economics, vol.62, pp.37-43,Elsevier.

[11] Ramanujam,V and Balakrishnan,H.,(2015).”Data-driven modeling of the airport configuration selection process”IEEE Transaction on Human-Machine Systems,vol.45,no.4,pp.490-499.

[12] Sandamali, G.G.N., Su, R. and Zhang, Y., (2019). “Flight Routing and Scheduling Under Departure and En Route Speed Uncertainty,” IEEE Transactions on IntelligentTransportation Systems, vol. 21,no.5, pp.1915-1928.

[13] Wang, Y., (2016).” Dynamic pricing considering strategic customers.” In *2016* International Conference on Logistics, Informatics and Service Sciences(LISS) pp.1-5. IEEE.

[14] Wen, C.H. and Chen, P.H., (2017). “Passenger booking timing for low-cost airlines: A continuous logit approach.” Journal of Air Transport Management,vol. 64, pp.91-99,Elsevier.

[15] Yang, H.T. and Liu, X., (2018).” Predictive simulation of airline passenger volume based on three models”. In International Conference of Pioneering ComputerScientists, Engineers and Educators , pp. 350-358.,Springer, Singapore

[16] Yang, L., Yin, S., Han, K., Haddad, J. and Hu, M., (2017). “Fundamental diagrams of airport surface traffic: Models and applications.”Transportation research part B:Methodological, vol.*106*, pp.29-51,Elsevier.

[17] Yeboah, G., Cottrill, C.D., Nelson, J.D., Corsar, D., Markovic, M. and Edwards, P., (2019). “Understanding factors influencing public transport passengers’ pre-travel information-seeking behaviour.” Public Transport,vol.*11,no.*1, pp.135-158,Springer.

[18] Zhang, Y., Su, R., Li, Q., Cassandras, C.G. and Xie, L., (2017). “Distributed flight routing and scheduling for air traffic flow management.” IEEE Transactions onIntelligent Transportation Systems, vol.18,no.1, pp.2681-2692.

[19] Zhang, Y., Li, Q. and Su, R., (2016). “Sector-based distributed scheduling strategy in air traffic flow management. “ IFAC-PapersOnLine, vol.49,no.3, pp.365-370,Elsevier.

[20] Zhang, Y., Su, R., Sandamali, G.G.N., Zhang, Y., Cassandras, C.G. and Xie, L., (2018). “A hierarchical heuristic approach for solving air traffic scheduling and routing problem with a novel air traffic model. “IEEE Transactions on Intelligent Transportation Systems, vol.20,no.9, pp.3421-3434.