First International Conference on Computer Engineering
International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
Volume 5 Issue 8, November-December-2020

Flight Ticket Price Prediction Using Machine Learning

Komal Kalane[1], Shivam Ghorpade[1], Omkar Jawale[1], Snehal More[1], Prof. Monika Dangore[1]

[1]Computer Department, Savitribai Phule Pune University, D.Y. Patil School of Engineering Charholi, Pune, Maharashtra, India

ABSTRACT

Nowadays, airline ticket prices can vary dynamically and significantly for the same flight, even for nearby seats within the same cabin. Customers are seeking to get the lowest price while airlines are trying to keep their overall revenue as high as possible and maximize their profit. Airlines use various kinds of computational techniques to increase their revenue such as demand prediction and price discrimination. From the customer side, two kinds of models are proposed by different researchers to save money for customers: models that predict the optimal time to buy a ticket and models that predict the minimum ticket price. In this paper, we present a review of customer side and airlines side prediction models. Our review analysis shows that models on both sides rely on limited set of features such as historical ticket price data, ticket purchase date and departure date. Features extracted from external factors such as social media data and search engine query are not considered. Therefore, we introduce and discuss the concept of using social media data for ticket/demand prediction.

Keywords : Survey, Ticket price prediction, Demand prediction, Price discrimination, Social media.

I. INTRODUCTION

The airline industry is considered as one of the most sophisticated industry in using complex pricing strategies. Nowadays, ticket prices can vary dynamically and significantly for the same flight, even for nearby seats as given in paper[119-128] and paper [28-42]. The ticket price of a specific flight can change up to 7 times a day.The Cheapest available ticket for a given data gets more or less expensive over time.This usually happens as an attempt to maximize revenue based on following things-

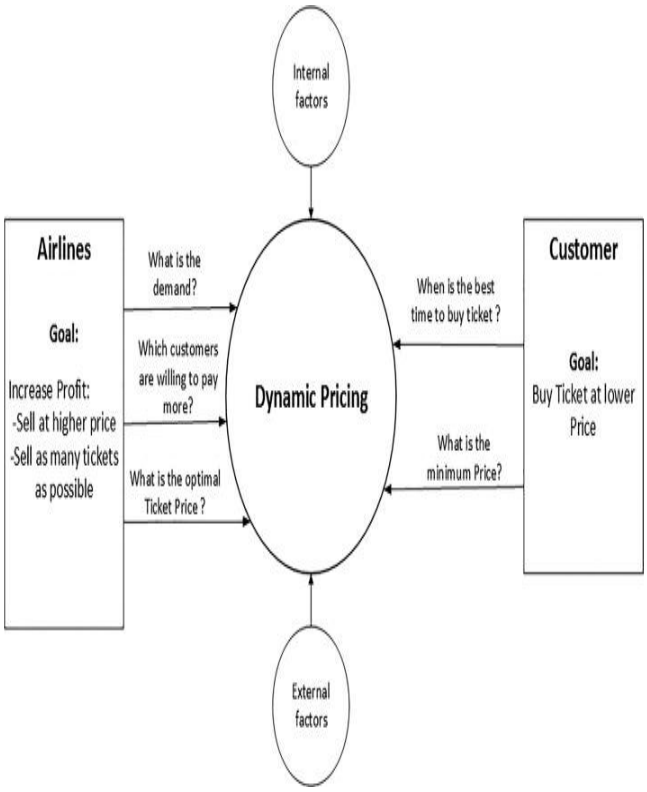
1.Time of purchase patterns(i.e. last minute purchases are expensive)

2.Keeping the flight as full as they want.

The objectives of the project can broadly be laid down by following questions:

1. Flight Trends
2. Best time to buy
3. Verifying myths

From the customer point of view, determining the minimum price or the best time to buy a ticket is the key issue. The conception of "tickets bought in advance are cheaper" is no longer working. It is possible that customers who bought a ticket earlier pay more than those who bought the same ticket later. Moreover, early purchasing implies a risk of commitment to a specific schedule that may need to be changed usually for a fee. The ticket price may be affected by several factors thus may change continuously. To address this, various studies were conducted to support the customer in determining an optimal ticket purchase time and ticket price prediction according to papers [267-276],[10-17], ,[87-90],[84-89], [964-969].



II. RELATED WORK

In the paper[1] author proposes reports on a pilot study in the domain of airline ticket prices where we recorded over 12,000 price observations over a 41 day period. When trained on this data, Hamlet — our multi-strategy data mining algorithm — generated a predictive model that saved 341 simulated passengers $198,074 by advising them when to buy and when to postpone ticket purchases. Remarkably, a clairvoyant algorithm with complete knowledge of future prices could save at most $320,572 in our simulation, thus Hamlet's savings were 61.8% of optimal. The algorithm's savings of $198,074 represents an average savings of 23.8% for the 341 passengers for whom savings are possible. Overall, Hamlet saved 4.4% of the ticket price averaged over the entire set of 4,488 simulated passengers.

In the Paper [2] author proposes that the prices have become increasingly available on the worls wide web.They are able to comparison shop efficiently and to track prices over time.

In the paper [3] author proposes A detailed discussion on the choice of forecast models and forecast variables is reported. A salient feature of the reported methods is that the forecast models can take into account the influence of market power on electricity prices.

In the paper [4] author proposes paper presents general patterns in airline pricing behaviour and a methodology for analysing different routes and/or carriers. The purpose is to provide customers with the relevant information they need to decide the best time to purchase a ticket, striking a balance between the desire to save money and any time restraints the buyer may have.

In the paper [5] author proposes Airline ticket price depends on many different dynamic factors, such as airline pricing policies, flight distance, class of service, airline, global population mobility, all of which define the travel demand. Ticket costs can vary significantly for the same flight, even for nearby seats. The model of airline tickets market varies from country to country and depends on the volume and structure of supply (number of airlines and air flights) and demand (number of passengers, seasonal peculiarities).

In The Paper [6] author proposes Airlines are one branch of these companies, having the available seats on a plane as their standing inventory. They divide these seats into several buckets, where each bucket has its own fare price. Airlines rearrange these seats across the buckets to make more money out of them, this creates changes in the prices which customers have to pay for the flight. Such that different customers pay different prices for tickets of the same flight.

Machine Learning is an idea to learn from examples and experience, without being explicitly programmed. Instead of writing code, you feed data to the generic algorithm, and it builds logic based on the data given. The process of training an ML model involves providing an ML algorithm (that is, Machine Learning Algorithm) with training data to learn from. The data classification can be performed on structured or unstructured data. The main goal of classification is to identify the category/class to which a new data will fall under. We can use these ML models to get predictions on new data for which target is unknown. II. GENRAL STRUCTURE Fig : General Structure of ML Model The above figure represents the general structure of an ML model. The data are split into Training and Test Data respectively. The Training Data is passed to trough the ML algorithms for enabling the machine to learn and apply it on the test data to predict the solutions. The such predicted solutions can be used for comparisons, calculate the accuracy.

III. LITERATURE REVIEW SUMMARY TABLE:

Paper no.	Paper title	Methodology	Advantages	Drawbacks
1	“Mining airfare data to minimize ticket purchase price”-ETZIONIETAL (2017)	Rule learning (Ripper), Reinforcement learning (Q-learning), time series methods, and combinations of these	An average of 61.8% savings achieved as compared to optimal saving.	Limitations in data set

2	"An agent for optimizing airline ticket purchasing"-WILLIAMGRO AND MARIAGIN (2017)	PLS regression Decision tree,nu-SVRRidge Regression	75.3% saving for the as compared to optimal saving.	Does not consider heterogeneous flights
3	"APPROACH TO PRICE FORCASTING"-WOHLFARTHETAL(2018)	Marked point processes (MPP) for Preprocessing	Data collected for 28 days for 6 routes	55% performance
4	"OPTIMALPURCHASE TIMING"-DOMINGUEZ-MENCHEROET (2017)	Non-parametric isotonic regression	2 months daily price information extracted 30 days	NULL
5	"DATA DRIVEN MODELLING"-ANASTASIA (2018)	Regression Model	Ticket price data collected for 75 days	The dataset set is limited
6	"A LINEAR QUANTILE REGRESSION MODEL" -T.JANSSEN (2018)	Linear quantile mixed regression model	Performs well for shorter period	Inefficient for longer period

IV. CONCLUSION:

After surveying the papers published before we conclude that the majority of the methods used in papers above made use of traditional prediction models from the computational intelligence research field known as Machine Learning. It was difficult for the customer to purchase air ticket due to the high complexity of the pricing models applied by the airlines because the prices change dynamically. Many features that can vary the base price of a ticket was not considered. There were problems that was encountered during manual data collection and previous works done in the field of airfare price prediction. From the experiments we concluded which features influence the airfare prediction at most. Future this project could be extended to predict the airfare prices with higher performance considering few other features.

V. REFERENCES

[1]. 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

[2]. 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

[3]. C. Koopmans and R. Lieshout, "Airline cost changes: To what extent are they passed through to the passenger?" Journal of Air Transport Management, vol. 53, pp. 1–11, 2016.

[4]. P. Malighetti, S. Paleari and R. Redondi, "Pricing strategies of low-cost airlines: The Ryanair case study," Journal of Air Transport Management, vol. 15, no. 4, pp. 195-203, 2009.

[5]. A regression model for predicting optimal purchase timing for airline tickets William Groves and Maria Gini October 18, 2011

[6]. Predicting Airfare Prices Manolis Papadakis, 2014

[7]. Linear Quantile Mixed Regression Model for Prediction of Airline Ticket Prices, 2014

[8]. To Buy or Not to Buy: Mining Airfare Data to Minimize, 2003

[9]. Random Forests, January 2011

[10]. Dingli, A., Mercieca, L., Spina, R., Galea, M., 2015. Event detection using social sensors. In: 2015 2nd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), Rennes, 2015, pp. 35–41.

[11]. Porshnev, A., Redkin, I., Shevchenko, A., 2013. Machine Learning in Prediction of Stock Market Indicators Based on Historical Data and Data from Twitter Sentiment Analysis. In: 2013 IEEE 13th International Conference on Data Mining Workshops, Dallas, TX, 2013, pp. 440–444.

[12]. Bo An, Haipeng Chen, Noseong Park, V.S. SubrahmanianMAP: Frequency-Based Maximization of Airline Profits based on an Ensemble Forecasting Approach Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), ACM, New York, NY, USA (2016), pp. 421-430

[13]. Bhavuk Chawla, Ms Chandandeep KaurAirfare Analysis And Prediction Using Data Mining And Machine Learning Int. J. Eng. Sci. Invention, 6 (11) (2017), pp. 10-17

[14]. Chieh-Hua Wen, Po-Hung ChenPassenger booking timing for low-cost airlines: a continuous logit approach J. Air Transport Manage., 64 (2017), pp. 91-99

[15]. Estimating dynamic demand for airlines Econ. Lett., 124 (1) (2014), pp. 26-29