



Document Information

Analyzed document	FinalReport.pdf (D123456410)
Submitted	2021-12-22T05:54:00.0000000
Submitted by	sagar kolekar 1
Submitter email	sagar@iiitdwd.ac.in
Similarity	8%
Analysis address	sagar.iiitdh@analysis.urkund.com

Sources included in the report

W	URL: https://core.ac.uk/download/pdf/55305133.pdf Fetched: 2021-06-21T09:22:57.3600000	 1
W	URL: https://aakashtechsupportdocs.readthedocs.io/en/latest/declaration.html Fetched: 2019-09-30T07:30:31.5500000	 1

Entire Document

A project report

88%

MATCHING BLOCK 1/2

W

<https://core.ac.uk/download/pdf/55305133.pdf>

submitted in partial fulfilment of the requirement for the degree of Bachelor of

Technology in

Computer Science Engineering On Decision AI Framework for Airline Price Prediction Submitted by Y. Mokshith ramendra(18bcs112) K Gireesh Kumar(18bcs035) Godina Pranav(18bcs028) Mareedu sai rama Krishna (18bcs050) Under the guidance of Dr. Uma.S Professor in Computer Science & Engineering IIIT Dharwad INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD

1 Certificate This is to certify that the work contained in the project report titled "Decision AI Framework for Airline Price Prediction" by Y. Mokshith ramendra(18bcs112), K Gireesh Kumar(18bcs035), Godina Pranav(18bcs028), Mareedu sai rama Krishna (18bcs050) was completed during the VII semester - IV Year as a Minor Project under the guidance of Dr.Uma.S Signature of Supervisor Dr.Uma.S Professor in Computer Science & Engineering

2

100%

MATCHING BLOCK 2/2

W

[https://aakashtechsupportdocs.readthedocs.io/e ...](https://aakashtechsupportdocs.readthedocs.io/e...)

Declaration We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Y. Mokshith ramendra(18bcs112) K Gireesh Kumar(18bcs035) Godina Pranav(18bcs028) Mareedu sai rama Krishna (18bcs050)

3 Approval Sheet This project report entitled Decision AI Framework for Airline Price Prediction by Y. Mokshith ramendra(18bcs112), K Gireesh Kumar(18bcs035), Godina Pranav(18bcs028), Mareedu sai rama Krishna (18bcs050) of Indian Institute of Information Technology, Dharwad is approved for the degree of Bachelor of Technology in Computer Science and Engineering . Supervisor Uma.S Professor Computer Science & Engineering, IIIT Dharwad Head of Department Dr. Uma S Computer Science & Engineering, IIIT Dharwad

4 Table of contents Chapter-1 : Introduction Chapter-2 : Abstract Chapter-3 : Literature work Chapter-4 : Problem statement Chapter-5 : Working model 5.1: Introduction to model 5.2: Model diagram 5.3: Requirements 5.4: Steps involved in training Chapter-6 : Results Chapter-7 : Future scope Chapter-8 : Conclusion Chapter-9 : References

5 Chapter-1: Introduction In today's fast-growing world transportation is the key for development of a company or nation. In order to meet its(world) requirement people has started preferring fastest way of transport i.e., Airlines. As number of people choosing airlines for transport has increased in order to meet the demand number of routes wee increased so does the prices. In airlines, ticket prices vary rapidly depending on various factors i.e., sometimes they are high and sometimes low, which makes difficult for passengers to choose when they have to purchase ticket and end paying more than actual cost of that travel. So, we tried to develop an AI based Airline Price prediction model, so that a passenger can know when to purchase ticket.

6 Chapter-2: Abstract There are various projects done on airline price prediction but many failed to use multiple factors that effect the actual price of a ticket. So, in our project we tried to include as many as useful features to include in training of our model and achieved a better percentage of success.

7 Chapter-3: Literature Review In the modern era air travel is considered to be fastest mode of transportation. However, the major factor effecting is prices. Especially in case of flight tickets the price won't be constant as in bus fare/train fare it varies abnormally. The main reason for it is that airline uses a technique called dynamic pricing with factors such as demand for a airline ticket over a specific route or type of seat and cat what price customers are willing to pay with a goal of selling maximum number of tickets at the same time making maximum profit possible for its airline company. Where as the customer willing to pay lowest price available for ticket. Because of these factors in same category of tickets one passenger pay less while other for same ticket end up paying higher. Although previously many projects are done towards development of model for prediction but many skipped few features in training model, we are working on a model random forest regression tree using various useful features to develop an efficient model/application so that we can predict the price of an airline ticket.

8 Chapter-4: Problem Statement Transportation plays an important role in today's economy out which fastest way of transport is through airlines. One of the major problems when it comes airlines is price as they considered to be highest among all forms of transportation. There are two kinds of travelling passengers namely leisure travellers, who book their travel way ahead of travelling date and secondly business travellers who make their booking in a short time where prices vary in both cases. In order to make their travel economically affordable we need to make a model that could predict the price of an airline ticket.

9 Chapter-5: Working model 5.1: Introduction to model The model used is "Random Forest regression tree". It is a supervised learning algorithm that uses ensemble learning method for regression. Ensemble learning is a technique that combines predictions from various machine learning algorithms to make a more accurate model. In this scenario those machine learning algorithms are decision tree. So, we can say that random forest regression tree forms an output based on various outputs from multiple decision trees. 5.2: Model Diagram 5.3: Requirements • Software Used: Jupyter or google colaboratory • Data set for training

10 5.4: Steps involved in training Step-1 : Data Pre-processing Since the dataset we have consists of factors such as • Airline name • Date of journey • Source and destination • Duration • Route • Departure and arrival times • Total stops and additional info All above mentioned are in string format and cannot be understand by our model so using Exploratory Data Analysis (EDA), one-hot encoding and Categorical data we covert them into int format so that our model could understand and train the data set. Image of the dataset at beginning of pre-processing: Image of the dataset at after pre-processing:

11 Step-2: Removing extra columns In the previous step we formed new columns in our required format for model training so we will be removing previous column from dataset through "drop" feature (data_train.drop()). Step-3: Training the model

12 With the pre-processed data set we will train the dataset using Random Forest Regression tree model. Step-4: After training the model we test our model with our test data set pre-processed in same way as training dataset to check our accuracy. Step-5 : We plot the graph based on test predictions and results as (y_test, y_pred).

13 Chapter-6: Results We have trained our dataset using randon forest regression tree and tested with test dataset and we have achieved an training accuracy of 96.526% and a test accuracy/prediction accuracy of 79.603

14 Chapter-7: Future scope As a fast growing economy, changes happens very quickly in now-days so does the pricing techniques followed by various airline industries. So we are looking forward for further development of our model towards greater accuracy also for developing a web based application for easier access to all to find a better price for airline ticket.

15 Chapter-8: Conclusion The aim of this project is to develop a working model that will be able to predict the price of an airline ticket price and to discuss about various factors that are used in deciding the price of an airline ticket by airline industry and factors that can be used to predict the price.

16 Chapter-9: References 1. https://www.kaggle.com/vinayshaw/airfare-price-prediction?select=Data_Train.xlsx 2. <https://www.analyticsvidhya.com/blog/2021/06/flight-price-prediction-a-regression-analysis-using-lazy-prediction/> 3. <https://medium.com/analytics-vidhya/regression-flight-price-prediction-6771fc4d1fb3> 4. <https://www.geeksforgeeks.org/random-forest-regression-in-python/> 5. <https://towardsdatascience.com/machine-learning-basics-random-forest-regression-be3e1e3bb91a> 6. <https://simpleflying.com/how-airline-ticket-pricing-works/>

Hit and source - focused comparison, Side by Side

Submitted text As student entered the text in the submitted document.
Matching text As the text appears in the source.

1/2	SUBMITTED TEXT	15 WORDS	88% MATCHING TEXT	15 WORDS
	submitted in partial fulfilment of the requirement for the degree of Bachelor of		Submitted in partial fulfillment of the requirements for the degree of Bachelor of	
	W https://core.ac.uk/download/pdf/55305133.pdf			

2/2	SUBMITTED TEXT	107 WORDS	100% MATCHING TEXT	107 WORDS
	Declaration We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.		DECLARATION¶ We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included. We have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.	
	W https://aakashtechsupportdocs.readthedocs.io/en/latest/declaration.html			