

**NAME OF THE PROJECT**

**FLIGHT PRICE PREDICTION PROJECT**

**SUBMITTED BY**

**SONAL MISAL**

**ACKNOLEDGEMENT**

Reference for the project is google. From google project data is scraped and formed the dataset. Google helps out to find the data for the given project.

**INTRODUCTION**

* Business Problem Framing

Our project is flight price prediction. To predict the price of a flight is a big problem because the fares are varied day by day and time by time. So, we can’t predict the actual price of flight. This problem is common for all the peoples around the world. Because flight price is necessary thing for traveler.

* Conceptual Background of the Domain Problem

For predicting the price of flight, we have required some information. For predicting the price, we need the travelling date, route, airline name. With the help of these data, we predicted the price of flight.

**ANALYTICAL PROBLEM FRAMING**

* In this project first we read and load the dataset. After loading dataset, we find out the data types of the columns.
* After finding the datatypes we find the null values in dataset.
* If any columns have null values, then, we have to fill those null values with their mean, median or mode of the columns.
* After filling the null values, we have to visualized the data in graph format. For plotting the graph, we used matplotlib, seaborn libraries.

**DATA INPUTS-LOGIC-OUPUTS RELATIONSHIP**

* After plotting the graphs, we have to described the dataset. After the description we have to find the correlation between the columns of dataset.
* Correlation is the relation between one or more variable. How variables are related to each other are shown in the correlation.
* Some columns are positively correlated and some are negatively correlated.
* Sometimes variables are highly correlated which results into multicollinearity.

**MODEL DEVELOPMENT AND EVALUATION**

* Data conversion is one of the step-in machine learning algorithms.
* In this step we have to convert objective data into numerical form.
* We can’t perform any operation on objective data so, we have to convert objective data into numerical form.
* After converting the data into numerical form, we have to split data into x and y.
* All the independent or feature variables are assigned to x and target or dependent variable are assigned to y.
* After splitting the data, we have to send data for training and testing.