Low Level Design (LLD)

Flight Fare Prediction

Gautam Sharma

Document Version Control

Date Issued	Version	Description	Author
05 September 2020	V3	Deployed	Gautam Sharma

Contents

Document Version Control 2
1 Introduction 5
1.1 Scope 5
1.2 Constraints 6
2 Technical specifications 6
2.1 Dataset 6
2.1.1 Diabetes dataset overview 6
2.1.2 Experimental Results 7
2.2 Correlations 8
2.3 Database 8
3 Technology stack 9
4 Proposed Solution 9
5 Model training/validation workflow 9
6 User I/O workflow 10
7 Exceptional scenarios 11

Abstract

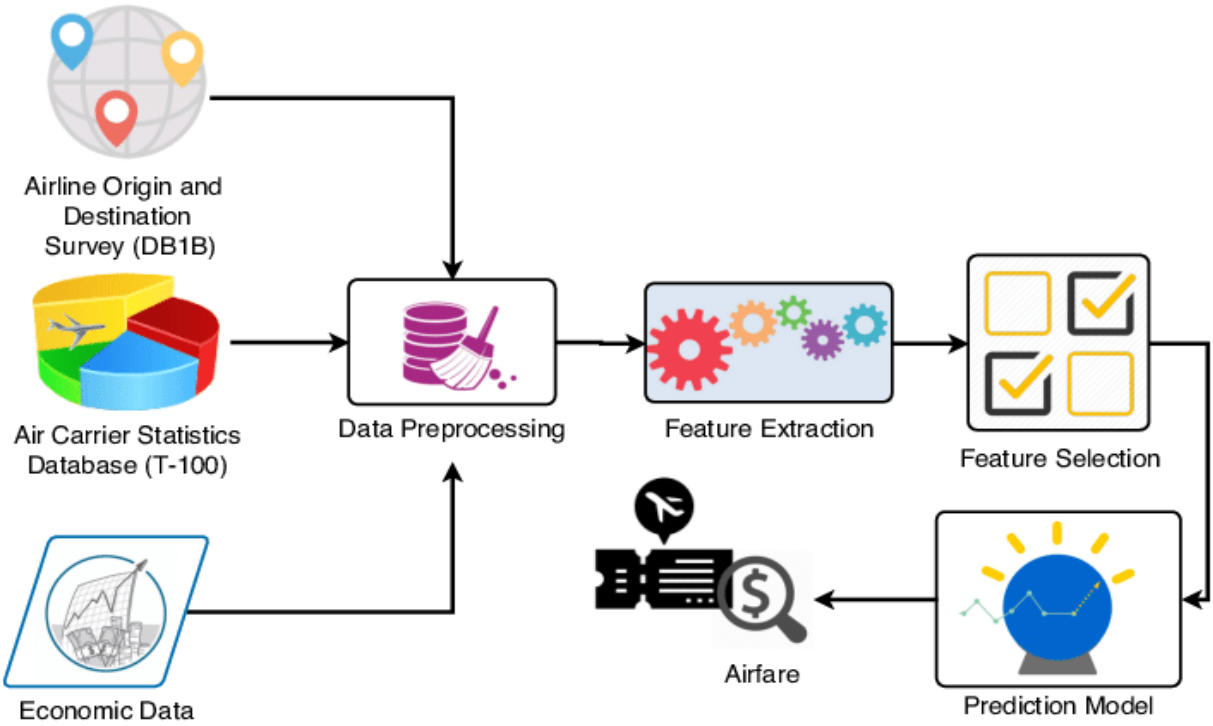
Travelling through flights has become an integral part of today's lifestyle as more and more people are opting for faster travelling options. The flight ticket prices increase or decrease every now and then depending on various factors like timing of the flights, destination, duration of flights. various occasions such as vacations or festive season. Therefore, having some basic idea of the flight fares before planning the trip will surely help many people save money and time. In the proposed system a predictive model will be created by applying machine learning algorithms to the collected historical data of flights. This system will give people the idea about the trends that prices follow and also provide a predicted price value which they can refer to before booking their flight tickets to save money. This kind of system or service can be provided to the customers by flight booking companies which will help the customers to book their tickets accordingly.

1 Introduction

This project aims to develop an application which will predict the flight prices for various flights using machine learning model.

The user will get the predicted values and with its reference the user can decide to book their tickets accordingly. In the current day scenario flight companies try to manipulate the flight ticket prices to maximize their profits.

There are many people who travel regularly through flights and so they have an idea about the best time to book cheap tickets. But there are also many people who are inexperienced in booking tickets and end up falling in discount traps made by the companies where actually they end up spending more than they should have. The proposed system can help save millions of rupees of customers by proving them the information to book tickets at the right time.



1.1 Scope

Currently, there are many fields where prediction-based services are used such as stock price predictor tools used by stock brokers and service like Zestimate which gives the estimated value of house prices. Therefore, there is requirement for

service like this in the aviation industry which can help the customers in booking tickets. There are many researches works that have been done on this using various techniques and more research is needed to improve the accuracy of the prediction by using different algorithms. More accurate data with better features can be also be used to get more accurate results.

1.2 Constraints

We will only be selecting a dataset from kaggle.

2 Technical specifications

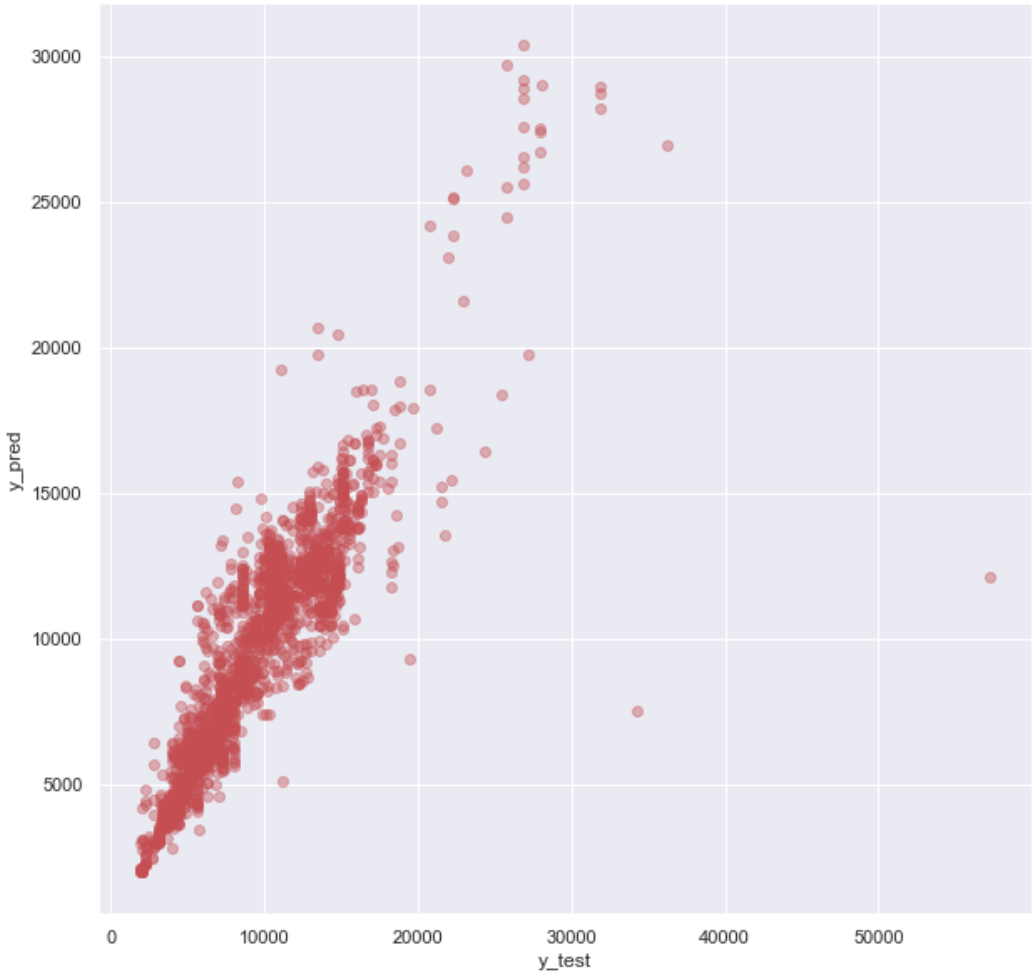
2.1 Dataset

2.1.1 Diabetes dataset overview

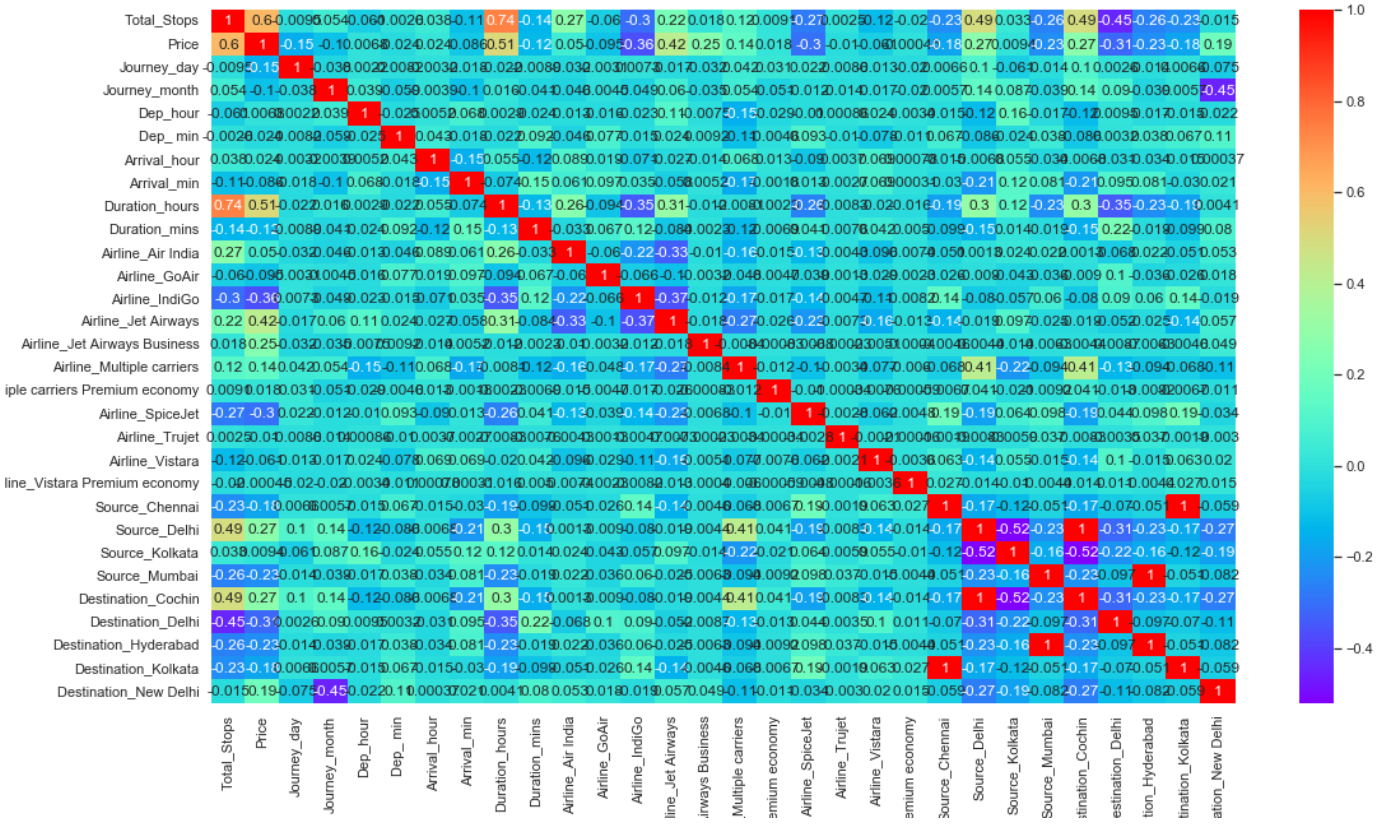
	A	B	C	D	E	F	G	H	I	J	K
1	Airline	e_of_Jour	Source	Destinatior	Route	Dep_Time	rrival_Tim	Duration	otal_Stop	ditional_In	Price
2	IndiGo	24/03/201	Banglore	New Delh	BLR → DEL	22:20	01:10 22 M	2h 50m	non-stop	No info	3897
3	Air India	1/05/2019	Kolkata	Banglore	CCU → IXF	05:50	13:15	7h 25m	2 stops	No info	7662
4	Jet Airway	9/06/2019	Delhi	Cochin	DEL → LKC	09:25	04:25 10 Ju	19h	2 stops	No info	13882
5	IndiGo	12/05/201	Kolkata	Banglore	CCU → NA	18:05	23:30	5h 25m	1 stop	No info	6218
6	IndiGo	01/03/201	Banglore	New Delh	BLR → NA	16:50	21:35	4h 45m	1 stop	No info	13302
7	SpiceJet	24/06/201	Kolkata	Banglore	CCU → BLI	09:00	11:25	2h 25m	non-stop	No info	3873
8	Jet Airway	12/03/201	Banglore	New Delh	BLR → BO	18:55	10:25 13 M	15h 30m	1 stop	In-flight m	11087
9	Jet Airway	01/03/201	Banglore	New Delh	BLR → BO	08:00	05:05 02 M	21h 5m	1 stop	No info	22270
10	Jet Airway	12/03/201	Banglore	New Delh	BLR → BO	08:55	10:25 13 M	25h 30m	1 stop	In-flight m	11087
11	Multiple c	27/05/201	Delhi	Cochin	DEL → BO	11:25	19:15	7h 50m	1 stop	No info	8625
12	Air India	1/06/2019	Delhi	Cochin	DEL → BLF	09:45	23:00	13h 15m	1 stop	No info	8907
13	IndiGo	18/04/201	Kolkata	Banglore	CCU → BLI	20:20	22:55	2h 35m	non-stop	No info	4174
14	Air India	24/06/201	Chennai	Kolkata	MAA → CC	11:40	13:55	2h 15m	non-stop	No info	4667
15	Jet Airway	9/05/2019	Kolkata	Banglore	CCU → BO	21:10	09:20 10 M	12h 10m	1 stop	In-flight m	9663
16	IndiGo	24/04/201	Kolkata	Banglore	CCU → BLI	17:15	19:50	2h 35m	non-stop	No info	4804
17	Air India	3/03/2019	Delhi	Cochin	DEL → AM	16:40	19:15 04 M	26h 35m	2 stops	No info	14011
18	SpiceJet	15/04/201	Delhi	Cochin	DEL → PN	08:45	13:15	4h 30m	1 stop	No info	5830
19	Jet Airway	12/06/201	Delhi	Cochin	DEL → BO	14:00	12:35 13 Ju	22h 35m	1 stop	In-flight m	10262
20	Air India	12/06/201	Delhi	Cochin	DEL → CC	20:15	19:15 13 Ju	23h	2 stops	No info	13381
21	Jet Airway	27/05/201	Delhi	Cochin	DEL → BO	16:00	12:35 28 M	20h 35m	1 stop	In-flight m	12898
22	GoAir	6/03/2019	Delhi	Cochin	DEL → BO	14:10	19:20	5h 10m	1 stop	No info	19495
23	Air India	21/03/201	Banglore	New Delh	BLR → CO	22:00	13:20 19 M	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
26	Jet Airway	6/05/2019	Kolkata	Banglore	CCU → BO	18:55	08:15 07 M	13h 20m	1 stop	In-flight m	7757
27	Jet Airway	9/06/2019	Delhi	Cochin	DEL → IDR	21:25	12:35 10 Ju	15h 10m	2 stops	No info	13292
28	IndiGo	1/06/2019	Delhi	Cochin	DEL → LKC	21:50	03:35 02 Ju	5h 45m	1 stop	No info	8238

Sheet1

2.1.2 Experimental Results



2.2 Correlations



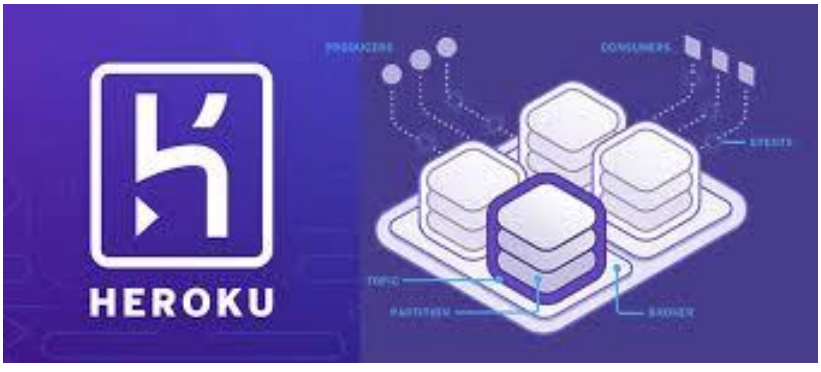
2.3 Database

System needs to store every request into the database and we need to store it in such a way that it is easy to retrain the model as well.

1. The User chooses the time and Amount
2. The User gives required information.

2.4 Deployment

1. Heroku



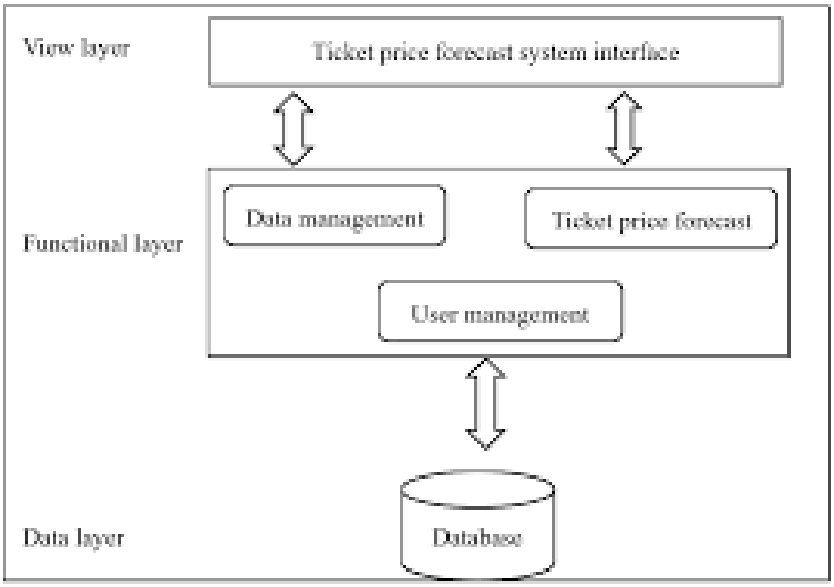
3 Technology stack

Front End	HTML/CSS/
Backend	Python Flask
Database	Kaggle
Deployment	Heroku

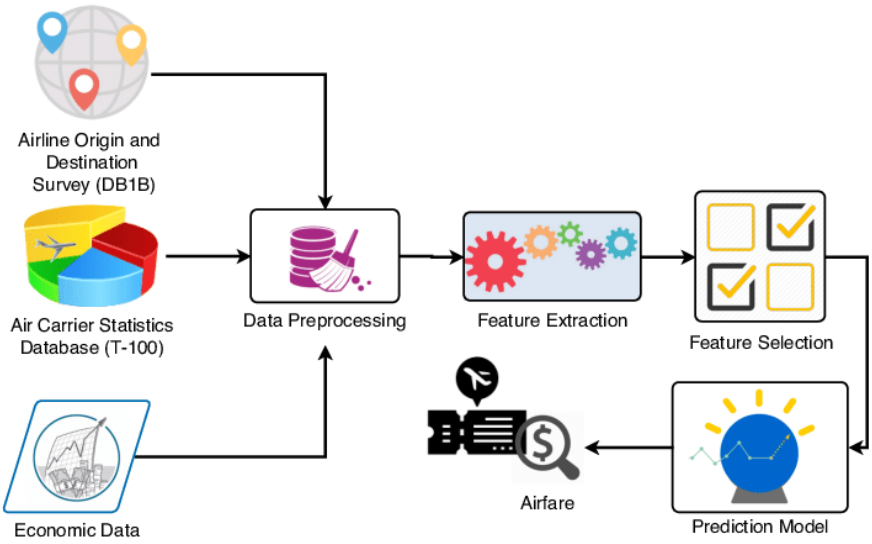
4 Proposed Solution

For this project, we have implemented the machine learning life cycle to create a basic web application which will predict the flight fare by applying machine learning algorithm to historical flight data using python libraries like Pandas, NumPy, Matplotlib, seaborn and sklearn.

5 Model training/validation workflow



6 User I/O workflow



7 Exceptional scenarios

Step	Exception	Mitigation	Module
5 September 2021	1.1	Deploy	Gautam Sharma