Low Level Design (LLD)

Flight Fare Prediction

Written By	Diana Laveena DSouza
Document Version	0.1
Last Revised Date	10-Oct-2022

Document Version Control

Change Record:

Version	Date	Author	Author
0.1	10/10/2022	Diana Laveena DSouza	Introduction & Architecture Defined

Reviews:

Version	Date	Reviewer	Comments
0.1	10/10/2022		Document Content, Version Control and Unit Test Cases to be added

Approval Status:

Version	Review Date	Reviewed By	Approved By	Comments

Contents

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, and duration of flights on 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. The main goal is to predict the fares of the flights based on different factors available in the provided dataset.

1 Introduction

1.1 Why is Low-Level Design Document?

The goal of LLD or a low-level design document (LLDD) is to give the internal logical design of the actual program code for Facebook Status Prediction. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so the programmer can directly code the program from the document.

1.2 Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. This process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

2 Technical specifications

2.1 Dataset

DataSet	Finalized	Source
Air Flight	yes	https://www.kaggle.com/datasets/nikhilmittal/flight-fare-prediction-mh

2.1.1 Dataset Overview

The Dataset consists of Airline, Date_of_Journey, Source, Destination, Route, Dep_Time, Arrival_Time, Duration, Total_Stops, Additional_Info with Price.

There are a total of 8369 records in the training set and 2093 in the validation set.

Air Flight

Airline	Date_of_Journey	Source	Destinatio	Route	Dep_Time	Arrival_Time	Duration	Total_Stop	Additional_Info	Price
GoAir	27/05/2019	Delhi	Cochin	DEL â†' AM	7:25	13:35	6h 10m	1 stop	No info	6841
Multiple ca	9/3/2019	Delhi	Cochin	DEL â†' BC	7:30	19:45	12h 15m	1 stop	No info	12615
Jet Airway	12/3/2019	Banglore	New Delhi	BLR â†' BC	5:45	3/13/2022 8:15	26h 30m	1 stop	No info	12547
SpiceJet	18/06/2019	Chennai	Kolkata	MAA â†' C	8:20	10:35	2h 15m	non-stop	No check-in baggage included	3543
Air India	15/05/2019	Delhi	Cochin	DEL â†' HY	13:05	5/16/2022 9:25	20h 20m	2 stops	No info	10975
Vistara	3/6/2019	Banglore	Delhi	BLR â†' DE	7:00	9:40	2h 40m	non-stop	No info	4668
IndiGo	27/06/2019	Delhi	Cochin	DEL â†' BL	5:10	10:05	4h 55m	1 stop	No info	6496
Multiple ca	15/05/2019	Delhi	Cochin	DEL â†' BC	8:45	19:00	10h 15m	1 stop	No info	9794
Air India	3/4/2019	Mumbai	Hyderabac	BOM â†' H	13:55	15:25	1h 30m	non-stop	No info	2575
Jet Airway	1/5/2019	Kolkata	Banglore	CCU â†' DI	20:25	5/2/2022 9:45	13h 20m	1 stop	No info	11467
Jet Airway	12/6/2019	Delhi	Cochin	DEL â†' BC	11:30	6/13/2022 4:25	16h 55m	1 stop	In-flight meal not included	10262
Jet Airway	3/3/2019	Delhi	Cochin	DEL â†' BC	9:00	18:50	9h 50m	1 stop	No info	17024
SpiceJet	9/6/2019	Delhi	Cochin	DEL â†' MA	15:45	22:05	6h 20m	1 stop	No info	6297
Jet Airway	9/5/2019	Kolkata	Banglore	CCU â†' BO	6:30	18:15	11h 45m	1 stop	No info	14781
Jet Airway	1/3/2019	Banglore	New Delhi	BLR â†' BC	11:40	3/2/2022 7:40	20h	1 stop	No info	22270

2.1.2 Input schema

Feature Name	Datatype	Size	Null/Required
Airline	varchar	20	Not Required
Date_of_Journey	date		Not Required
Source	varchar	20	Not Required
Destination	varchar	20	Not Required
Route	varchar	20	Not Required
Dep_Time	time		Not Required
Arrival_Time	time		Not Required
Duration	time		Not Required
Total_Stops	varchar	20	Not Required
Additional_Info	varchar	20	Not Required
Price	double	20	Not Required

2.2 Predicting Flight Fare

- The system presents the set of inputs required from the user.
- The user gives the required information.
- The System should be able to predict the exact flight price based on the user information.

2.3 Database

The database used in this project to store the data is the Cassandra database.

2.5 Deployment

3 Technology Stack

Front End	HTML/CSS
Backend	Python Flask
Database	Cassandra
Deployment	GCP

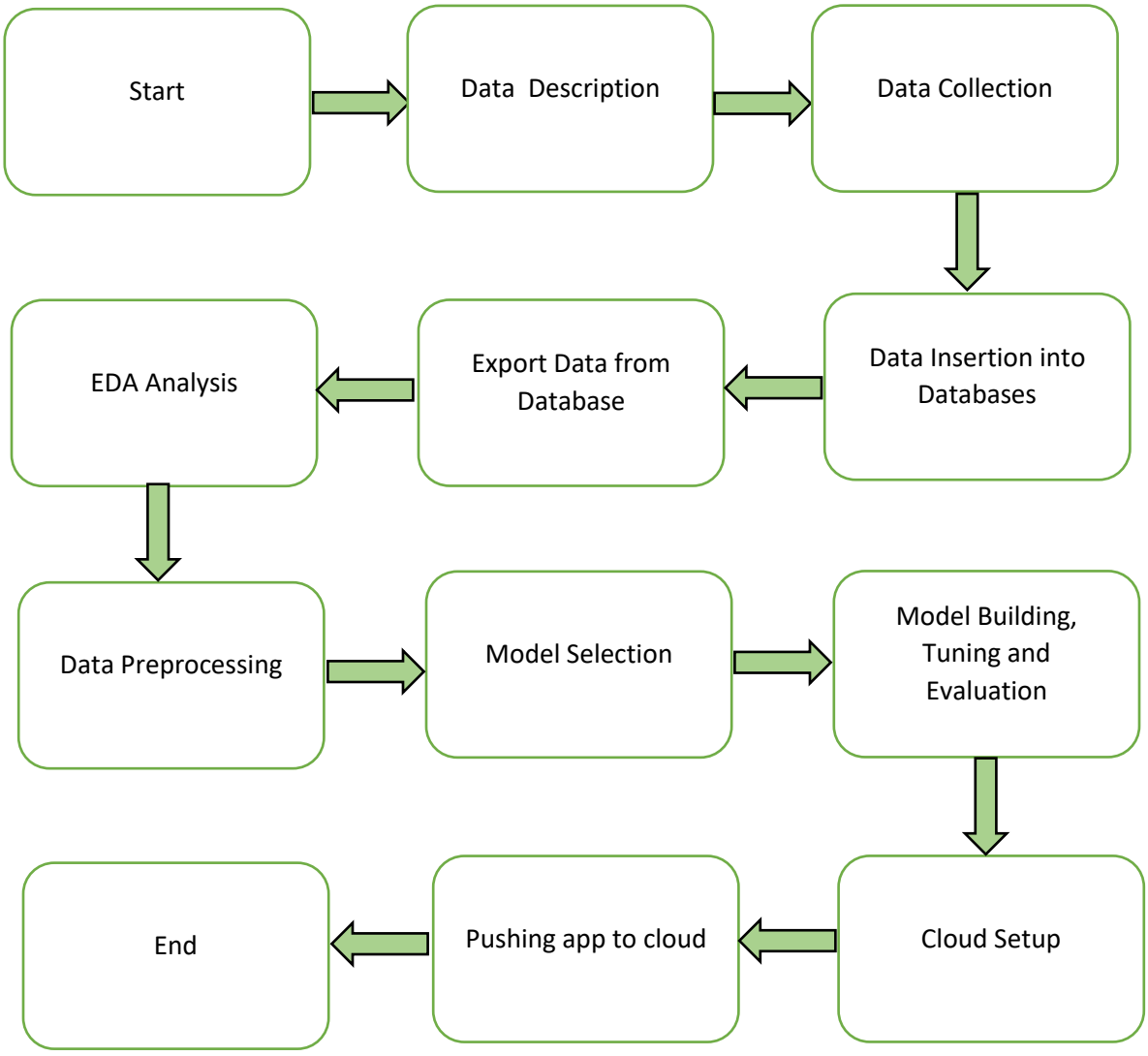
4 Proposed Solution

Refer: https://www.ijraset.com/research-paper/flight-fare-prediction-system-using-ml

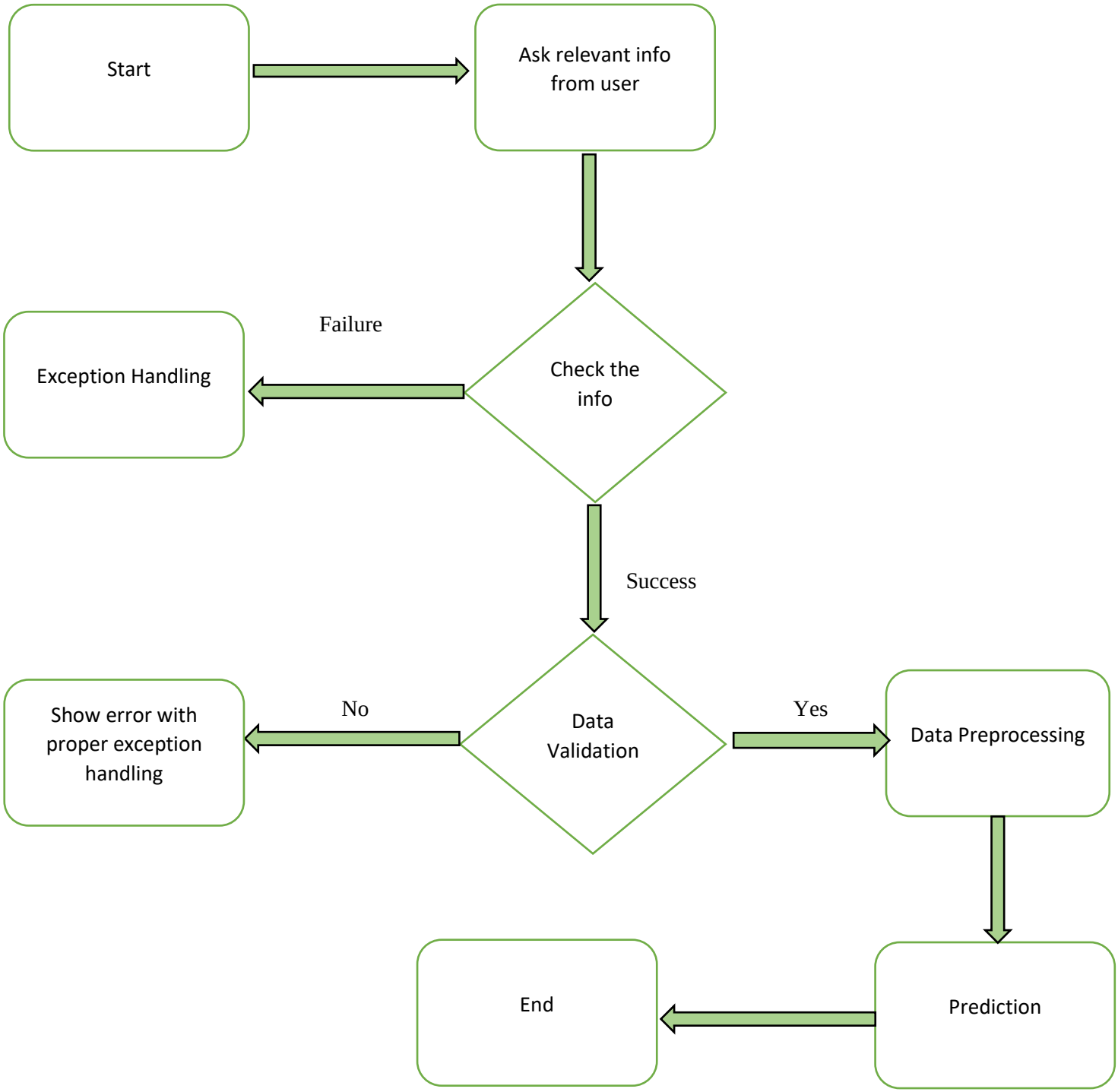
The actual research paper proposed the method of Machine Learning. Finally, we selected Machine learning-based methods. Under Machine learning-based methods, we choose Supervised. We decided on Supervised methods such as tree-based methods, Linear models, K-nearest neighbour and SVR. But the tree-based method XGBOOST gives better results than others.

Baseline Model: XGBOOST, since this is a regression problem.

5 Model Training/Validation Workflow



6 User I/O workflow



7 Test Cases

Test Case Description	Pre-Requisite	Expected Result
Verify whether the Application URL is accessible to the user	1. Application URL should be defined	The application URL should be accessible to the user.
Verify whether the Application loads entirely for the user when the URL is accessed	1. Application URL is accessible 2. Application is deployed	The Application should load entirely for the user when the URL is accessed.
Verify whether the user can input the text in all input fields	1. Application is accessible	The user should be able to input the text in all input fields.
Verify whether the user gets Submit button to submit the inputs.		The user should get Submit button to submit the inputs.
Verify whether the user is presented with results on clicking submit.		The user should be presented with results on clicking submit