ARCHITECTURE

SPAM-HAM CLASSIFIER

Written By	Animesh Jain
Document Version	1.0
Last Revised Data	14 Jun 2022

1 Document Version Control:

Version	Date	Author	Comments
1.0	14/06/2022	Animesh Jain	

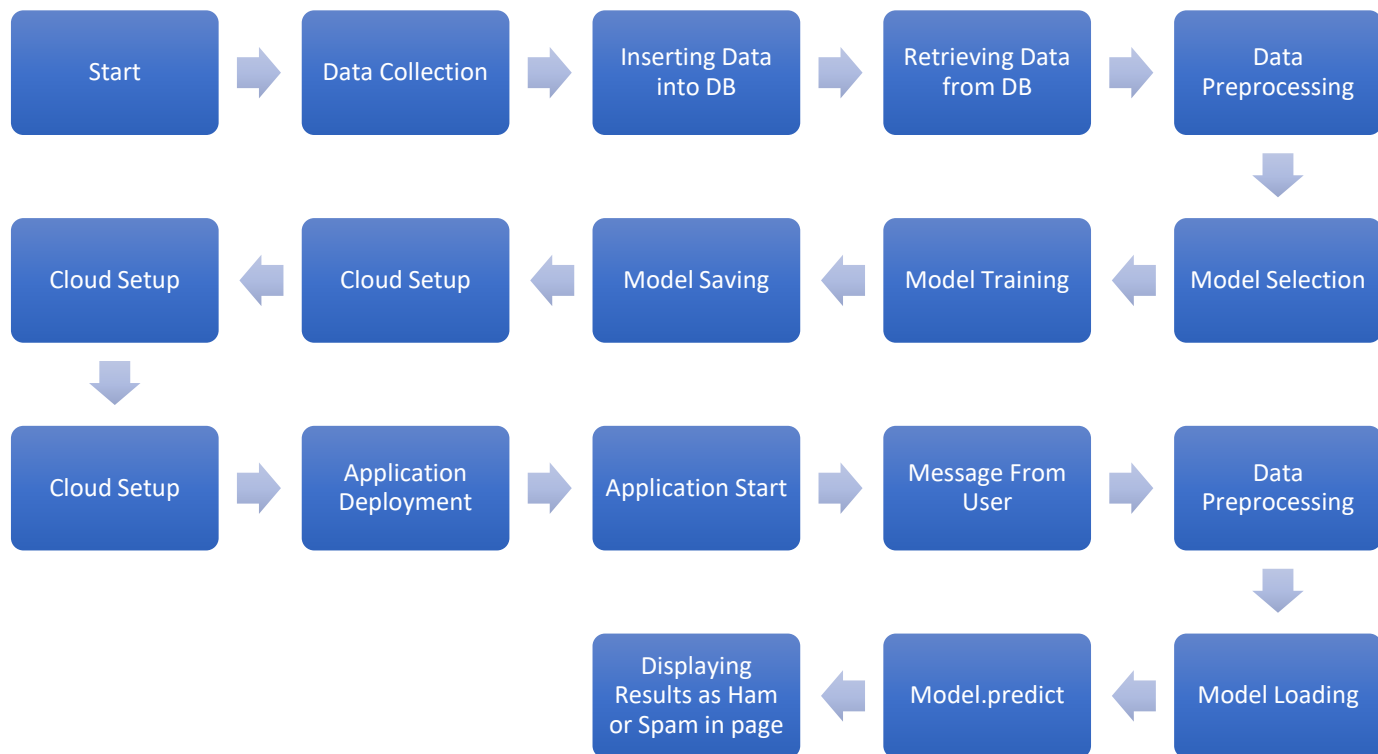# Table of Contents

1.Introduction

1.1 What 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 flight fare estimation System. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that 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. Architecture

3. Dataset

3.1 Dataset Overview

The training dataset consists of 5 columns and every column datatype is string except the

	Id	target	text	V1	V2	V3
1	Id	target	text	V1	V2	V3
2	0	ham	Go until jurong point, crazy.. Available only in bugis n great world la e buffet... Cine there got amore wat...			
3	1	ham	Ok lar... Joking wif u oni...			
4	2	spam	Free entry in 2 a wkly comp to win FA Cup final tkts 21st May 2005. Text FA to 87121 to receive entry question(std txt rate)T&C's apply 08452810075over18's			
5	3	ham	U dun say so early hor... U c already then say...			
6	4	ham	Nah I don't think he goes to usf, he lives around here though			
7	5	spam	FreeMsg Hey there darling it's been 3 week's now and no word back! I'd like some fun you up for it still? Tb ok! XxX std chgs to send, ÃƒÂ£ÃƒÂ¢Ã‚Â£ÃƒÂ¢Ã‚Â£1.50 to rcv			
8	6	ham	Even my brother is not like to speak with me. They treat me like aids patent.			
9	7	ham	As per your request 'Melle Melle (Oru Minnaminunginte Nurungu Vettam)' has been set as your callertune for all Callers. Press *9 to copy your friends Callertune			
10	8	spam	WINNER!! As a valued network customer you have been selected to receivea ÃƒÂ£ÃƒÂ¢Ã‚Â¥ÃƒÂ¢Ã‚Â£900 prize reward! To claim call 09061701461. Claim code KL341. Valid 12 hours only.			
11	9	spam	Had your mobile 11 months or more? U R entitled to Update to the latest colour mobiles with camera for Free! Call The Mobile Update Co FREE on 08002986030			
12	10	ham	I'm gonna be home soon and i don't want to talk about this stuff anymore tonight, k? I've cried enough today.			
13	11	spam	SIX chances to win CASH! From 100 to 20,000 pounds txt> CSH11 and send to 87575. Cost 150p/day, 6days, 16+ TsandCs apply Reply HL 4 info			
14	12	spam	URGENT! You have won a 1 week FREE membership in our ÃƒÂ£ÃƒÂ¢Ã‚Â¥ÃƒÂ¢Ã‚Â£100,000 Prize Jackpot! Txt the word: CLAIM to No: 81010 T&C www.dbuk.net LCCLTD POBOX 4403LDNW1A7RW18			
15	13	ham	I've been searching for the right words to thank you for this breather. I promise i wont take your help for granted and will fulfil my promise. You have been wonderful and a blessing at all times.			
16	14	ham	I HAVE A DATE ON SUNDAY WITH WILL!!			
17	15	spam	XXXMobileMovieClub: To use your credit, click the WAP link in the next txt message or click here>> http://wap. xxxmobilemovieclub.com?n=QJKGIGHJJGCBL			
18	16	ham	Oh k...i'm watching here:)			
19	17	ham	Eh u remember how 2 spell his name... Yes i did. He v naughty make until i v wet.			
20	18	ham	Fine if thatÃƒÂ£ÃƒÂ¢Ã‚Â¥ÃƒÂ£ÃƒÂ¢Ã‚Â•s the way u feel. ThatÃƒÂ£ÃƒÂ¢Ã‚Â¥ÃƒÂ¢ÃƒÂ£ÃƒÂ¢Ã‚Â•s the way its gota b			
21	19	spam	England v Macedonia - dont miss the goals/team news. Txt ur national team to 87077 eg ENGLAND to 87077 Try:WALES, SCOTLAND 4txt/ÃƒÂ£ÃƒÂ¢Ã‚ÂŒÃƒÂ¢Ã‚Â¼1.20 POBOXox36504W45WQ 16+			
22	20	ham	Is that seriously how you spell his name?			
23	21	ham	IÃƒÂ£Ã‚Â‰ÃƒÂ£ÃƒÂ¢Ã‚Â›ÃƒÂ£ÃƒÂ¢Ã‚Â·m going to try for 2 months ha ha only joking			
24	22	ham	So ÃƒÂ£ÃƒÂ¢Ã‚ÂŒ_ pay first lar... Then when is da stock comin...			
25	23	ham	Aft i finish my lunch then i go str down lor. Ard 3 smth lor. U finish ur lunch already?			

Testing Data is just a message.

4. Logging

Logging is very important to keep track of the activities performed by our application. I have used logging module to do so. All the logs either it is train or test, both will be present inside All_logs folder. Logging helps us in debugging process also so it is mandatory to do.

5. Database

The database I am using is Cassandra db. 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. The system stores each and every data given by the user or received on request to the database.

6. Deployment

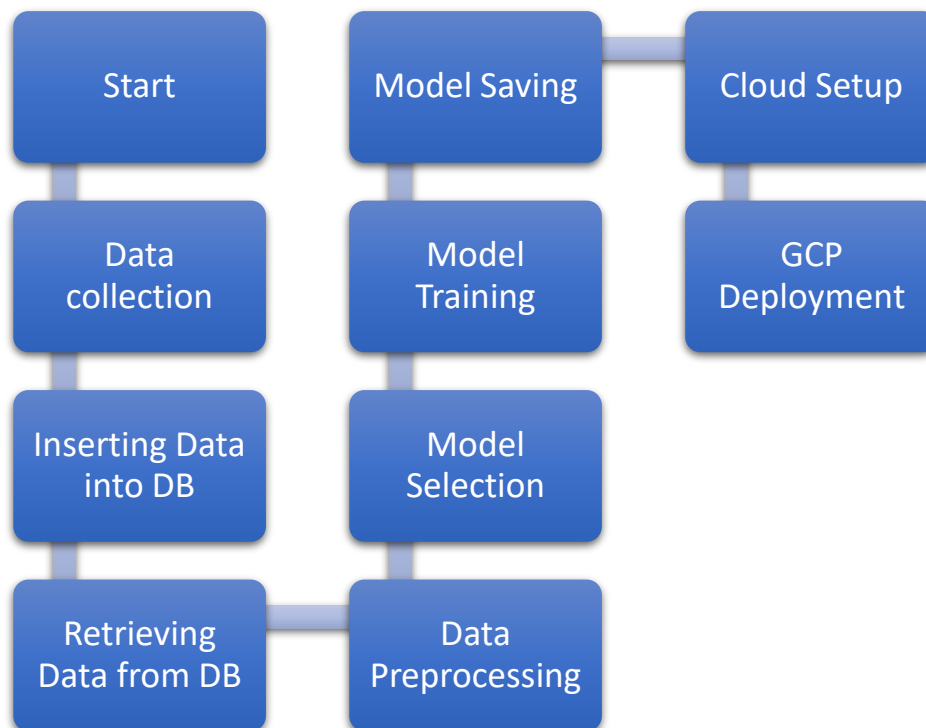Deployment is done in AWS and it's a production server.

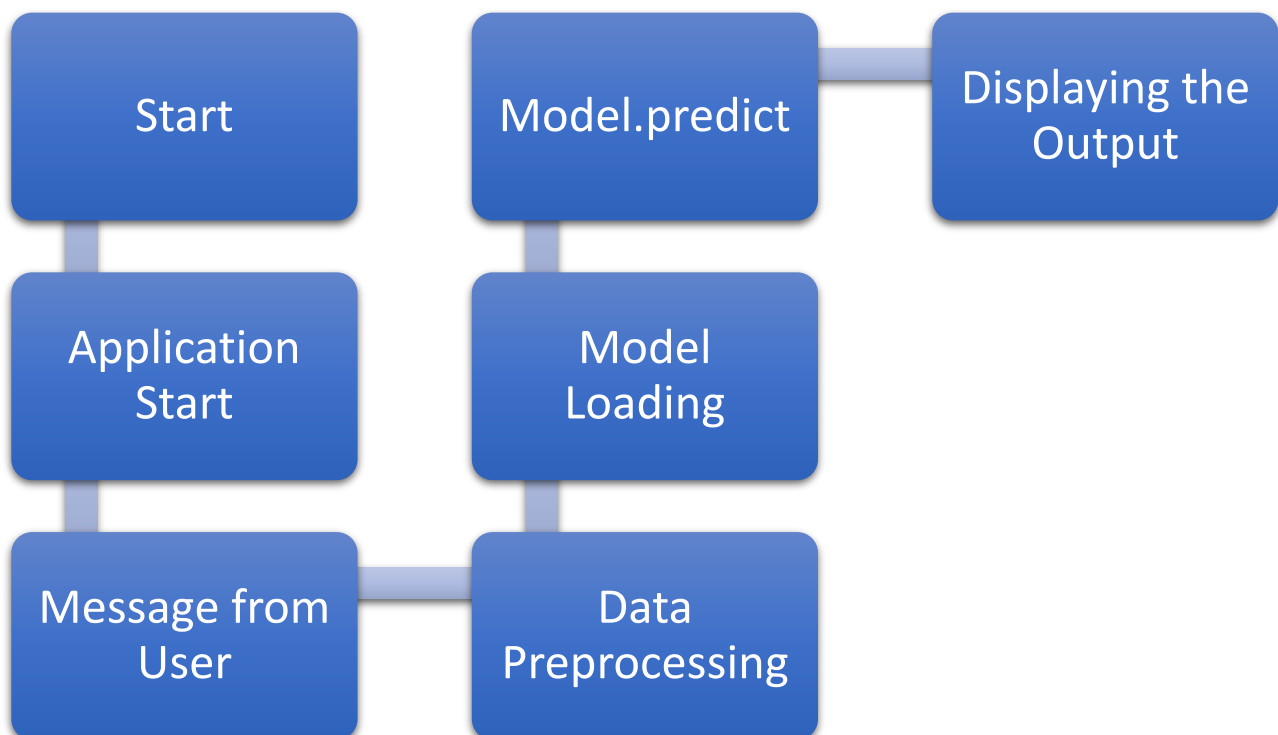Link: http://ec2-3-140-241-66.us-east-2.compute.amazonaws.com:5000/

7. Proposed Solution

Solution is very simple here. I am going to build a simple ML model which will be able to classify the message type based on the message words. Doing some research I got to know that Multinomial Naïve Bayes is the best one.

8. Model Training/Validation workflow

Start	Model Saving	Cloud Setup
Data collection	Model Training	GCP Deployment
Inserting Data into DB	Model Selection	
Retrieving Data from DB	Data Preprocessing	

9. User I/O workflow

Start	Model.predict	Displaying the Output
Application Start	Model Loading	
Message from User	Data Preprocessing	

4 Test Cases

Test cases are given below

Test Case Description	Pre-Requisite	Expected Result
Verify whether the Application URL is accessible to the user	1. Application URL should be defined	Application URL should be accessible to the user
Verify whether the Application loads completely for the user when the URL is accessed	1. Application URL is accessible 2. Application is deployed	The Application should load completely for the user when the URL is accessed
Verify Response time of url from backend model.	1. Application is accessible	The latency and accessibility of application is very faster we got in heroku service.
Verify whether user is giving standard input.	1. Handeled test cases at backends.	User should be able to see successfully valid results.
Verify whether user is able to see input fields on logging in	1. Application is accessible 2. User is logged in to the application	User should be able to see input fields on logging in
Verify whether user is able to edit all input fields	1. Application is accessible 2. User is logged in to the application	User should be able to edit all input fields
Verify whether user gets Predict button to submit the message.	1. Application is accessible 2. User is logged in to the application	User should get predict button to submit the message.

Verify whether user is presented with recommended results on clicking predict button.	1. Application is accessible 2. User is logged in to the application	User should be presented with recommended results on clicking predict.
Verify whether the recommended results are in accordance to the selections user made	1. Application is accessible 2. User is logged in to the application and database	The recommended results should be in accordance to the selections user made