

Low Level Design (LLD)

Spam Ham Classifier

Geetha S

Akanksha Venkatesh V B

Ashok Kumar S

M C Shravan

Document Version Control

Date Issued	Version	Description	Author
29 th September 2021	1.0	Complete Architecture	Geetha S Akanksha Venkatesh V B Ashok Kumar S M C Shravan

Contents

Document version control	2
Abstract	4
1. Introduction	5
1.1 Why this Low-Level Design Document	5
1.2 Scope	5
1.3 Constraints	5
1.4 Risks	5
1.5 Out of Scope	6
2. Technical specifications	6
2.1 Dataset	6
2.2 Predicting Spam or Ham	9
2.3 Logging	9
2.4 Deployment	9
3 Technology stack	9
4 Proposed Solution	10
5 Model training/validation workflow	10
6 User I/O workflow	11
7 Key Performance Indicators	11

Abstract

The high number of unsolicited emails commonly called spam has necessitated the development of an increase in reliable and robust antispam filters. Using the recent machine learning approach we have been successful in detecting and filtering spam emails. We are classifying Spam or ham from the dataset which is a set of SMS tagged messages that have been collected for SMS Spam research.

1 Introduction

1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the Spam Ham Classifier. It will explain the purpose and features of the system, the interfaces of the system, what the system will do, the constraints under which it must operate and how the system will react to external stimuli. This document is intended for developers of the system and will be proposed to the higher management for its approval.

The main objective of the project is to classify spam and ham messages.

Our datasets consist of information, such as:

1. Label: spam, ham.
2. Message: spam, ham.

1.2 Scope

This software system will be a Web application. This system will be designed to classify spam or ham messages using Machine Learning Techniques.

1.3 Constraints

Variation in accuracy and limited to the given dataset.

1.4 Risks

Document specific risks that have been identified or that should be considered.

1.5 Out of Scope

Delineate specific activities, capabilities, and items that are out of scope for the project.

2 Technical specifications

2.1 Dataset

2.1.1 Dataset overview

Consists of 2 different tables.

We have a train (5284) and test (1322) datasets, We need to classify Spam or ham from the dataset which is a set of SMS tagged messages.

ham	Do you know what Mallika Sherawat did yesterday? Find out now @ <URL>
spam	Congrats! 1 year special cinema pass for 2 is yours. call 09061209465 now! C Suprman V, Matrix3, StarWars3, etc all 4 FREE! bx420-ip4-5we. 150pm. Dont miss out!
ham	Sorry, I'll call later in meeting.
ham	Tell where you reached
ham	Yes..gauti and sehvag out of odi series.
ham	Your gonna have to pick up a \$1 burger for yourself on your way home. I can't even move. Pain is killing me.
ham	Ha ha ha good joke. Girls are situation seekers.
ham	Its a part of checking IQ
ham	Sorry my roommates took forever, it ok if I come by now?
ham	Ok lar i double check wif da hair dresser already he said wun cut v short. He said will cut until i look nice.
spam	As a valued customer, I am pleased to advise you that following recent review of your Mob No. you are awarded with a £1500 Bonus Prize, call 09066364589
ham	Today is "song dedicated day.." Which song will u dedicate for me? Send this to all ur valuable frnds but first rply me...
spam	Urgent UR awarded a complimentary trip to EuroDisinc Trav, Aco&Entry41 Or £1000. To claim txt DIS to 87121 18+6*£1.50(moreFrmMob. ShrAcomOrSglSupt)10, LS1 3AJ
spam	Did you hear about the new "Divorce Barbie"? It comes with all of Ken's stuff!
ham	I plane to give on this month end.
ham	Wah lucky man... Then can save money... Hee...
ham	Finished class where are you.
ham	HI BABE IM AT HOME NOW WANNA DO SOMETHING? XX
ham	K..k..)where are you?how did you performed?
ham	U can call me now...
ham	I am waiting machan. Call me once you free.

Low Level Design (LLD)

ham	Annoying isn't it.
ham	Goodmorning, Today i am late for & min.
ham	There's no point hangin on to mr not right if he's not makin u happy
ham	All will come alive.better correct any good looking figure there itself..
ham	In that case I guess I'll see you at campus lodge
ham	We're done...
ham	Come to my home for one last time i wont do anything. Trust me.
ham	I was up all night too worrying about this appt. It's a shame we missed a girls night out with quizzes popcorn and you doing my hair.
spam	Sex up ur mobile with a FREE sexy pic of Jordan! Just text BABE to 88600. Then every wk get a sexy celeb! PocketBabe.co.uk 4 more pics. 16 £3/wk 087016248
ham	Ok... C ya...
spam	You have 1 new voicemail. Please call 08719181503
ham	What he said is not the matter. My mind saying some other matter is there.
ham	He also knows about lunch menu only da. . I know
ham	Al he does is moan at me if n e thin goes wrong its my fault&al de arguments r my fault&fed up of him of himso y bother? Hav 2go, thanx.xx
ham	NEFT Transaction with reference number & for Rs. & has been credited to the beneficiary account on & at & : &
ham	Otherwise had part time job na-tuition..
ham	I know she called me
ham	Me also da, i feel yesterday night wait til 2day night dear.
ham	Thanks for understanding. I've been trying to tell sura that.
spam	WIN a year supply of CDs 4 a store of ur choice worth £500 & enter our £100 Weekly draw txt MUSIC to 87066 Ts&Cs www.Ldew.com.subs16+1win150ppmx3
ham	The whole car appreciated the last two! Dad and are having a map reading semi argument but apart from that things are going ok. P.
ham	Going for dinner.msg you after.
ham	I'm ok wif it cos i like 2 try new things. But i scared u dun like mah. Cos u said not too loud.
spam	GENT! We are trying to contact you. Last weekends draw shows that you won a £1000 prize GUARANTEED. Call 09064012160. Claim Code K52. Valid 12hrs only. 150ppm
ham	Wa, ur openin sentence very formal... Anyway, i'm fine too, juz tt i'm eatin too much n puttin on weight...Haha... So anythin special happened?
ham	As I entered my cabin my PA said, " Happy B'day Boss !!". I felt special. She askd me 4 lunch. After lunch she invited me to her apartment. We went there.
spam	You are a winner U have been specially selected 2 receive £1000 or a 4* holiday (flights inc) speak to a live operator 2 claim 0871277810910p/min (18+)
ham	Goodo! Yes we must speak friday - egg-potato ratio for tortilla needed!
ham	Hmm...my uncle just informed me that he's paying the school directly. So pls buy food.
spam	PRIVATE! Your 2004 Account Statement for 07742676969 shows 786 unredeemed Bonus Points. To claim call 08719180248 Identifier Code: 45239 Expires
spam	URGENT! Your Mobile No. was awarded £2000 Bonus Caller Prize on 5/9/03 This is our final try to contact U! Call from Landline 09064019788 BOX42WR29C, 150PPM
ham	here is my new address -apples&pairs&all that malarky
spam	Today's Voda numbers ending 7548 are selected to receive a \$350 award. If you have a match please call 08712300220 quoting claim code 4041 standard rates app
ham	I am going to sao mu today. Will be done only at 12
ham	Ü predict wat time u'll finish buying?
ham	Good stuff, will do.
ham	Just so that you know,yetunde hasn't sent money yet. I just sent her a text not to bother sending. So its over, you dont have to involve yourself in anything. I shouldn't have imposed anything on yo
ham	Are you there in room.
ham	HEY GIRL. HOW R U? HOPE U R WELL ME AN DEL R BAK! AGAIN LONG TIME NO C! GIVE ME A CALL SUM TIME FROM LUCYxx
ham	K..k:)how much does it cost?
ham	I'm home.
ham	Dear, will call Tmorrow.pls accomodate.

2.1.2 Input schema

Feature name	Datatype	Null/Required
Message	Object	Required
Label	Object	Required

2.2 Predicting Spam or Ham

- The system displays the output based on entered input.
- The User selects the message.
- The system helps to provide the set of inputs required from the user.
- It gives the required information.
- The system should be able to classify spam and ham.

2.3 Logging

We should be able to log every activity done by the user.

- The System identifies at what step logging required
- The System should be able to log each and every system flow.
- Developers can choose logging methods. You can choose database logging/ File logging as well.
- The system should not be hung even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

2.4 Deployment

1. Heroku



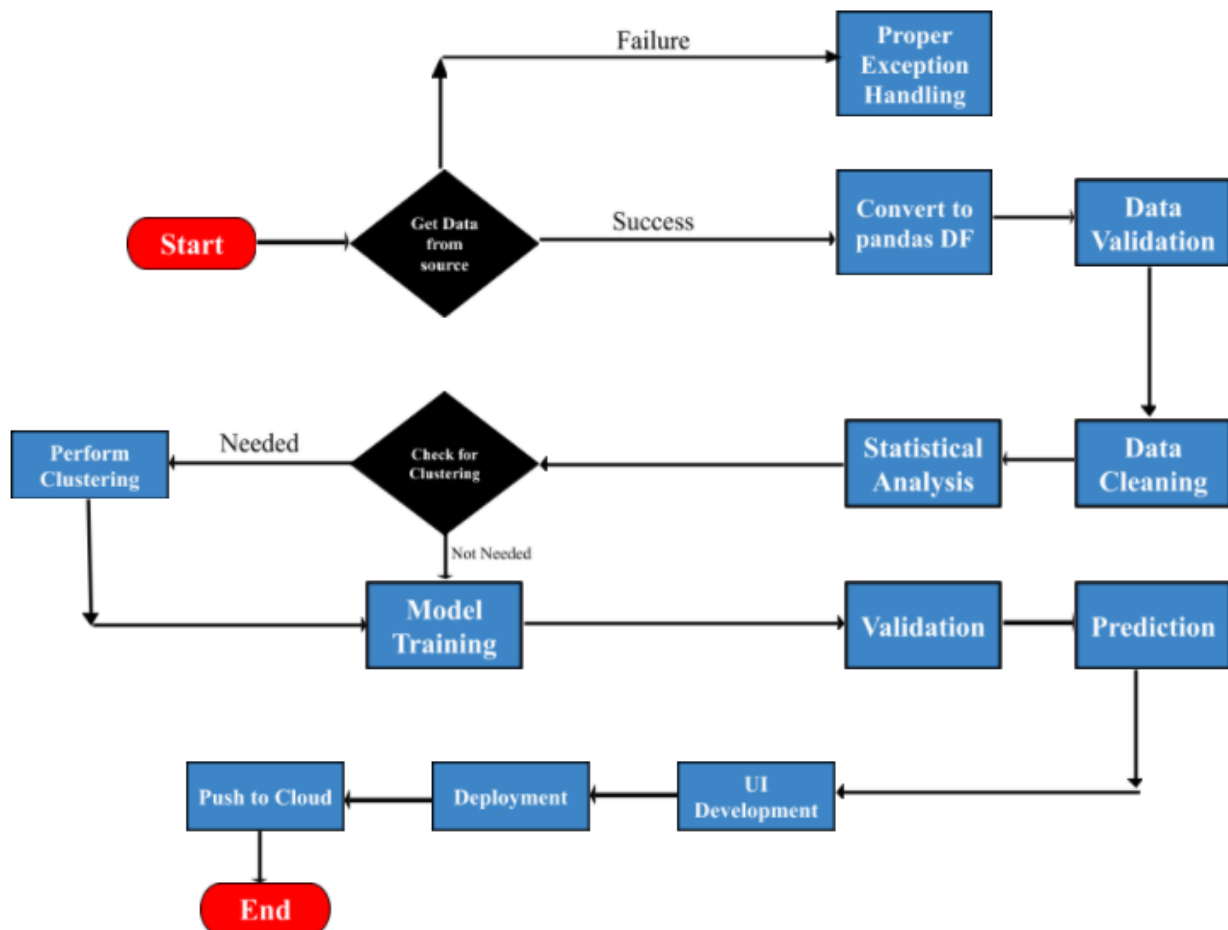
3 Technology stack

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

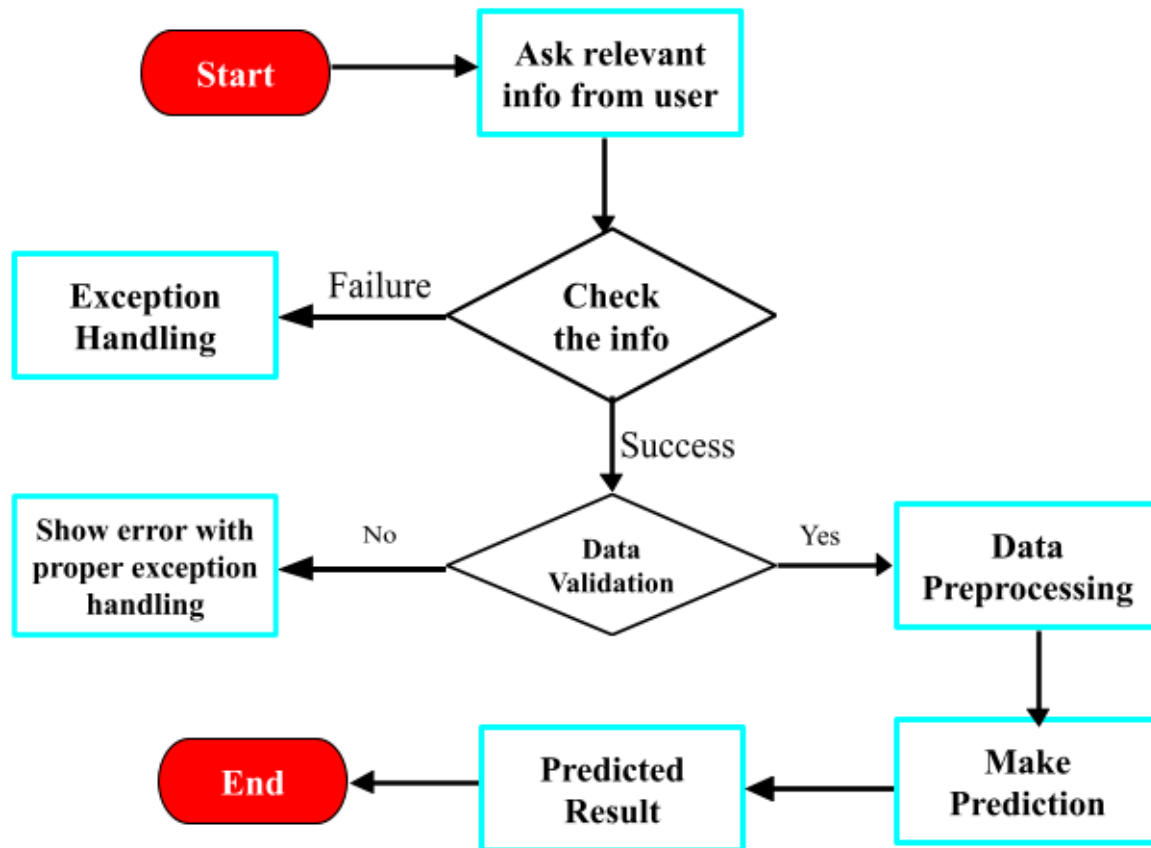
3. Proposed Solution

The solution proposed here is a Spam Ham Classifier which will help to classify Spam or ham from the given dataset. Here we have used Logistic regression, Extra trees classifier, Adaboost classifier, XGBoost classifier and Gradient boosting classifier.

4. Model training/validation workflow



4 User I/O workflow



5 Key performance indicators (KPI)

- The output will be provided within seconds
- Provides good accuracy.
- Easy to access.
- The classification of Spam and Ham can be done efficiently.

5.1 Latency

- The output for the given input will be provided within seconds after clicking on the submit button.