

Multi-Mailing and Spam Detection System using ML, Python & GUI

Preetam Raj¹, Neelam Kumari², Tushar Raina³, Pronika Chawla⁴, Madhumita Kathuria⁵

¹Preetam Raj: Department of Computer Science Engineering, MRIIRS, Faridabad, Haryana, India

²Neelam Kumari: Department of Computer Science Engineering, MRIIRS, Faridabad, Haryana, India

³Tushar Raina: Department of Computer Science Engineering, MRIIRS, Faridabad, Haryana, India

⁴Pronika Chawla: Professor, Department of Computer Science Engineering, MRIIRS, Faridabad, Haryana, India

⁵Madhumita Kathuria: Professor, Department of Computer Science Engineering, MRIIRS, Faridabad, Haryana, India

Abstract - Mailing system is the necessity of each and everyone in today's world, as unlike chatting it is a formal form of conversation. Mailing system, it is a strategy for utilizing innovation to unite individuals and thoughts notwithstanding the topographical boundaries. The objectives of this desktop application are to send and receive mail and along with that it will also detect spam mails. Traditional Spam detection is done on the servers of the provided mail but the objective of our prototype is to skip that and provide spam detection on the user pc only so that the data is absolutely safe and unread before a user by any application on the server. We are implementing machine learning to understand the difference between spam and ham mails which will successfully eliminate spam mails.

Key Words: OPM- Online Project Management System, POP- Post Office Protocol, IMAP- Internet Messages Access Protocol, ML- Machine Learning, GUI- Graphical User Interface, SMTP- Simple Mail Transfer Protocol

1. INTRODUCTION

Multi-mailing system and spam detection is to detect spam mail, send and receive mail through a desktop application conventional better and easy access. Mails can be sent to multiple people simultaneously by creating a mailing list. Text filters will be applied to judge mails accordingly. This project is a prototype for the mail management system and it is restricted within the college premises. This project is useful for all the users because they can communicate to other person via internet. It is an system to help in project Management. It is increasingly helpful in current market circumstance where an association isn't near entryway or a city or a country. For this situation sharing record and information identified with venture starting with one corner of the world then onto the next by utilizing web makes our work for simple. Yet still it is unmanaged, to deal with this work we are making this system online project management system (OPM).

We are implementing machine learning to understand the difference between spam and ham mail which will successfully eliminate spam mails. The mailing system uses a desktop application which takes input a mail address and will also have feature to use a database to forward the mail

to the selected mailing address. The system will also include inbox feature and along that a spam detection module which will scan the mails and pick up spam mails and separate them from the other such that the user can have a clean experience of reading only the important mail.

Then the spam detection module is responsible for declaring whether the mail is spam or not. A machine learning algorithm is use to predict whether a mail is spam or not. This module uses stopword which is a data set which contains all the data and information about the stopwords, email.csv is our chosen Data set which contains spam mail and normal mails.

2. EXISTING SYSTEM

Email is a way of interchange of messages between user using devices having Internet connectivity. Previously, mailing systems required the sender and the receiver to both have to be online at the same time, in common with the quick messaging. Now, mailing systems are like save-and-forward model. Now, Email serves accept, store, forward and deliver messages. Neither the system nor the users, no one is need to be online simultaneously. They need to relate just rapidly, generally to mail server or a webmail interface for whatever time span that it takes to send or get messages. Gmail, Google created this mailing system. Users can access Gmail tool on the web and utilizing outsider projects the synchronize email content through POP or IMAP conventions. Google's mail servers naturally examine sends for various purposes, counting to channel spam and malware, and to add setting tricky promotions near messages. The notification practice has been in a general sense investigated by security advocates in view of stress over unfathomable data support, effortlessness of checking by outcasts, customers of other email providers not having assented to the course of action in the wake of sending messages to Gmail addresses, and the potential for Google to change its ways to deal with moreover lessen insurance by getting information together with other Google data use. The association has been the subject of cases concerning the issues. Google has communicated that email customers must "indispensable expected" their email to be at risk to modernized planning and assurance that the organization quit indicating promotions near perhaps sensitive message, for instance, those referencing race, religion, sexual bearing, prosperity, or monetary rundowns. In June 2017, Google announced the best in class end to the

usage of applicable Gmail content for advancing purposes, contingent rather upon data amassed from the use of its various organizations.

3. PROPOSED SYSTEM

Proposed system uses other servers as the mean of sending mail what the system is after is not to provide or compete with previously designed systems but to use that system and further add new features to it.

Unlike normal system our system can send to a multiple user same mail as it has a database that is uses to send mail to all the users in that database without letting other users know that the mail is either forwarded or sent to other users. The second module in the proposed system is spam detection system that uses concepts of machine learning to propose whether the mail is spam or ham.

How the machine learning model works is that it trains the machine on a dataset over 1000 mails which are classified on the basis of spam and ham and them created an algorithm to detect whether the mail is spam or ham where spam being the mail which is of no use and ham being mail which are generated by user or some work related lives mails which the users casually send to each other. This project is a prototype for the mailing management system and it is restricted within the college premises. This project is useful for the entire user as they can communicate with each other using this technology via internet.

Traditional spam detection is done on the servers of the provided mail but the objectives of our prototype is to skip that and provide spam detection on the use pc only so that the data is absolutely safe and unread before a user by any application on the server. The spam detection Algorithm is implemented through machine learning therefore can be improved accordingly with the help of well-defined training set.

The product model utilized by our framework is the cascade model. Cascade model is a precise and successive way to deal with the product improvement. This incorporates framework designing and displaying which sets up necessities for all the framework components and apportioning some subset of these prerequisites to programming. Framework building and examination incorporate prerequisite social affair at the framework level with modest quantity of top-level plan. Examination information building consolidate need assortment at the key business level and at the business area level.

4.1. Problem definition:

The primary stage of product development. Really characterizes point, the idea of venture. Goal, the project is to connect people from different regions through emails via internet. It can be use by project team in their project management along with spam detection with the help of machine learning.

This project helps in differentiating between spam and ham messages and eliminating spam messages and only passing the legitimate messages to client.

This project is based on machine learning. This project require some kind of tool chain and development environment already installed and ready like jupyter or Spyder. Some of the libraries such as Sckit-learn, Pandas, Numpy of matplotlib are used.

4.2. Analysis:

Existing system is the level product and we don't want to compete with it. But we just want to add some feature to it.

4.3. Designing and coding

This stage is most important one as it provides all the structure of the product.

4.4. Testing:

Testing looks for any fault in the structure of the product.

4.5. Maintenance:

Product maintenance to look whether it sends the mail to right person or group in the list and forward the mail without any trace of other recipients and detecting the spam accurately.

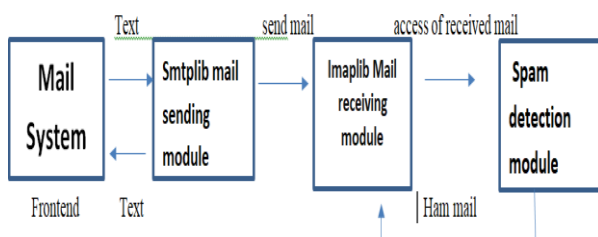


Fig-1: Component Design

4. METHODOLOGY

The life span of creation of this product includes models ie, object-oriented model, Prototype model, waterfall model etc. for creating the proper system. Waterfall model, the oldest model of creation of proper system.

5. DESIGN AND IMPLEMENTATION

We are developing Multi-Mailing and spam detection tool. ER diagram, use case diagram, Class diagram. These diagrams are below in Fig. 3, Fig.4 and Fig.5 respectively.

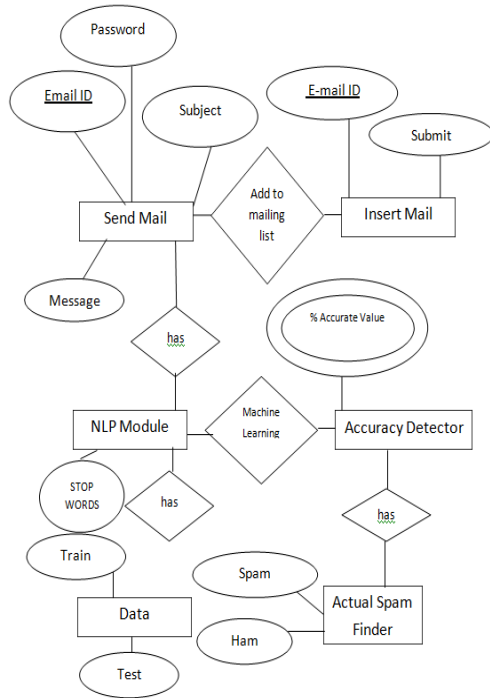


Fig-3: ER Diagram

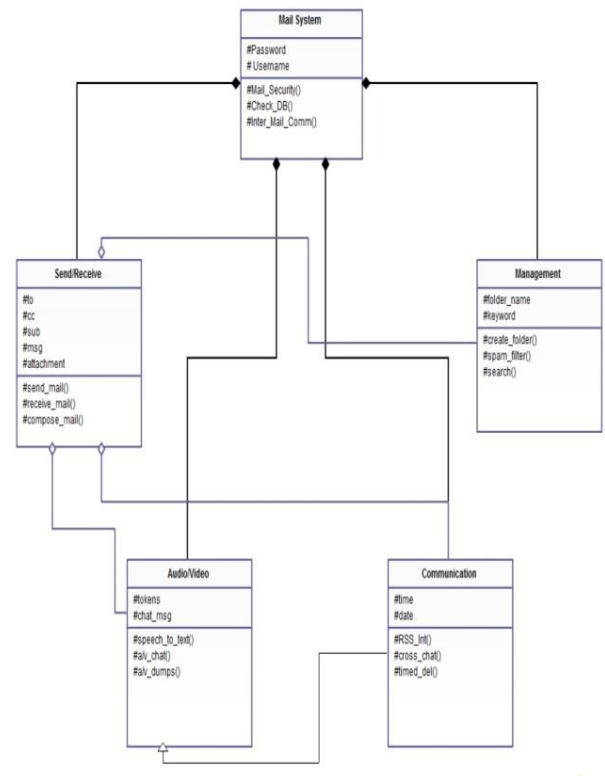


Fig-5: Class Diagram

Modules of the product are given below:

5.1. Registration:

The main module of the Product. The user who needs to get benefits of our product, they needs to enroll themselves. Product will incite what subtleties should be entered. As needs be, client will give input through text of info. Before submitting inputs, framework will affirm the data by provoking.

5.2. Login:

When the enlistment is done, client can login to framework by giving username and secret phrase to getting to the record. This module is utilized for confirmation. It will acknowledge username and secret key in content arrangement. At that point this content will be utilized by framework to choose whether that client is substantial or not. On the off chance that they is approved client then framework will divert that client to next GUI.

5.3. Compose Mail:

The module, User can make mail they needs to send. It fills in as follows:

- 1) Product will request the required information
- 2) Users can give this data by text.

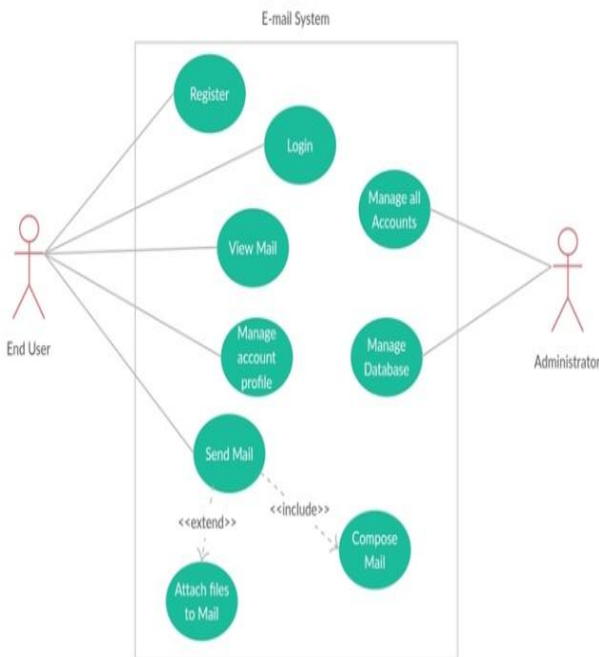


Fig-4: Use Case Diagram

- 3) Then framework will peruse beneficiary name, subject, and message body and request affirmation.
- 4) Once the affirmation is done mail will be sent to particular beneficiary/s.

5.4. Inbox:

This choice causes client to see all sends got to hello there/her record. In the wake of choosing this choice, latest sends get stacked. At that point framework will incite beneficiary's name and subject of each mail. The client additionally has choice to erase the got mail.

5.5. Sent:

This choice monitors sends sent by the client. So as to get to the sent sends client needs to do activities gave by the brief to explore between sends. It will help the client in proficiently extricating and sending the necessary mail.

5.6. Spam Detection:

This module helps user to detect the spam with the help of machine learning. This module is liable for declaring whether the mail is spam or not. A machine learning rule is employed to predict whether a mail is spam or not. This module uses stopwords which is a data set which contains all the data and information about the stopwords, emails.csv is our chosen data set which contains spam mails and normal mails.

6. FUTURE SCOPE

However, numerous endeavours have been actualized towards tackling the issue of spam SMS utilizing administrative, social and innovative measures, the arrangement proposed are not finished arrangements.

- 1) Achieving precise grouping, with zero percent (0%) misclassification of Ham SMS as spam and spam SMS as Ham.
- 2) The endeavours would be applied to square phishing SMS which conveys the phishing assaults and now-days which is progressively matter of concern.

The framework we are creating will be working just on windows. As expanding utilization of cell phone step by step, there is a need to remember this office as an application for cell phone too.

7. CONCLUSION

The Multi-mailing system construction in a way where future enhancement can be done easily, the following

conclusion can be deduced from the development of the project.

Without opening any browser mail can be sent and received by the user. Mails can be broadcasted to multiple users by creating a directory. Program will access mailing system via SMTP and will send mail to mentioned or desired audience. Spam free inbox which is done on the client pc. The system has adequate scope for modification in future

Bibliography

1. Babu, A. K. (2016). *Send Email using python*. Retrieved from Freecodecamp: <https://www.freecodecamp.org/news/send-emails-using-code-4fcea9df63f/>
2. Dhanashree.D.Zope, P. (2016). *International journal of scientific research in computer science and engineering*. https://www.isroset.org/journal/IJSRCSE/full_paper_view.php?paper_id=444
3. *Intra Mailing System*. (n.d.). Retrieved from Scribd: <https://www.scribd.com/doc/61315817/Intra-Mailing-System>
4. Islam, M. R. (2005). *Spam filtering using ML Algorithm*. Retrieved from Academia: https://www.academia.edu/36614854/Spam_Filtering_Using_ML_Algorithms
5. *Mailing system Python Projects*. (2019). Retrieved from freeprojectz: <https://www.freeprojectz.com/python-django-project/ mailing-system>
6. *Nevonproject*. (n.d.). Retrieved from <https://nevonprojects.com/email-client-project/>
7. *Research on Mail System: Project report for Bachelor in Computer - Rajendra Man Banepali*. (2006). Retrieved from Scribd: <https://www.scribd.com/doc/19518895/Research-on-Mail-System-Project-report-for-Bachelor-in-Computer-Rajendra-Man-Banepali>

REFERENCES

- [1]. https://www.isroset.org/journal/IJSRCSE/full_paper_view.php?paper_id=444
- [2]. <https://www.freecodecamp.org/news/send-emails-using-code-4fcea9df63f/>
- [3]. <https://www.scribd.com/doc/61315817/Intra-Mailing-System>

- [4]. <https://www.scribd.com/doc/19518895/Research-on-Mail-System-Project-report-for-Bachelor-in-Computer-Rajendra-Man-Banepali>
- [5]. <https://nevonprojects.com/email-client-project/>
- [6]. <https://www.freeprojectz.com/python-django-project/mailing-system>
- [7]. https://www.academia.edu/36614854/Spam_Filtering_Using_ML_Algorithms
- [8]. <https://www.pythonforbeginners.com/code-snippets-source-code/using-python-to-send-email>
- [9]. https://blog.mailtrap.io/sending-emails-in-python-tutorial-with-code-examples/#Next_steps_with_emails_in_Python
- [10]. <https://towardsdatascience.com/spam-detection-with-logistic-regression-23e3709e522?gi=7adac30ea6b3>
- [11]. <https://hackernoon.com/how-to-build-a-simple-spam-detecting-machine-learning-classifier-4471fe6b816e>