OPTIMIZING SPAM FILTERING WITH MACHINE LEARNING

S.NO	CONTENT	PAGE NO
01	Introduction	02
02	Problem Definition and Design Thinking	06
03	Result	08
04	Advantages and Disadvantages	10
05	Applications	13
06	Conclusion	16
07	Future Scope	18
08	Appendix	20

INTRODUCTION

In recent times, unwanted commercial bulk emails called spam has become a huge problem on the internet. The person sending the spam messages is referred to as the spammer. Such a person gathers email addresses from different websites, chatrooms, and viruses. Spam prevents the user from making full and good use of time, storage capacity and network bandwidth. The huge volume of spam mails flowing through the computer networks have destructive effects on the memory space of email servers, communication bandwidth, CPU power and user time . The menace of spam email is on the increase on yearly basis and is responsible for over 77% of the whole global email traffic. Users who receive spam emails that they did not request find it very irritating. It is also resulted to untold financial loss to many users who have fallen victim of internet scams and other fraudulent practices of spammers who send emails pretending to be from reputable companies with the intention to persuade individuals to disclose sensitive personal information like passwords, Bank Verification Number (BVN) and credit card numbers.

OVERVIEW:

Text messages are essential these days; however, spam texts have contributed negatively to the success of this communication mode. The compromised authenticity of such messages has given rise to several security breaches. Using spam messages, malicious links have been sent to either harm the system or obtain information detrimental to the user. Spam SMS messages as well as emails have been used as media for attacks such as masquerading and smishing (a phishing attack through text messaging), and this has threatened both the user and service providers.

Therefore, given the waves of attacks, the need to identify and remove these spam messages is important. This dissertation explores the process of text classification from data input to embedded representation of the words in vector form and finally the classification process. Therefore, we have applied different embedding methods to capture both the linguistic and semantic meanings of words. Static embedding methods that are used include Word to Vector (Word2Vec) and Global Vectors (Glove), while for dynamic embedding the transfer learning of the Bidirectional Encoder Representations from Transformers (BERT) was employed.

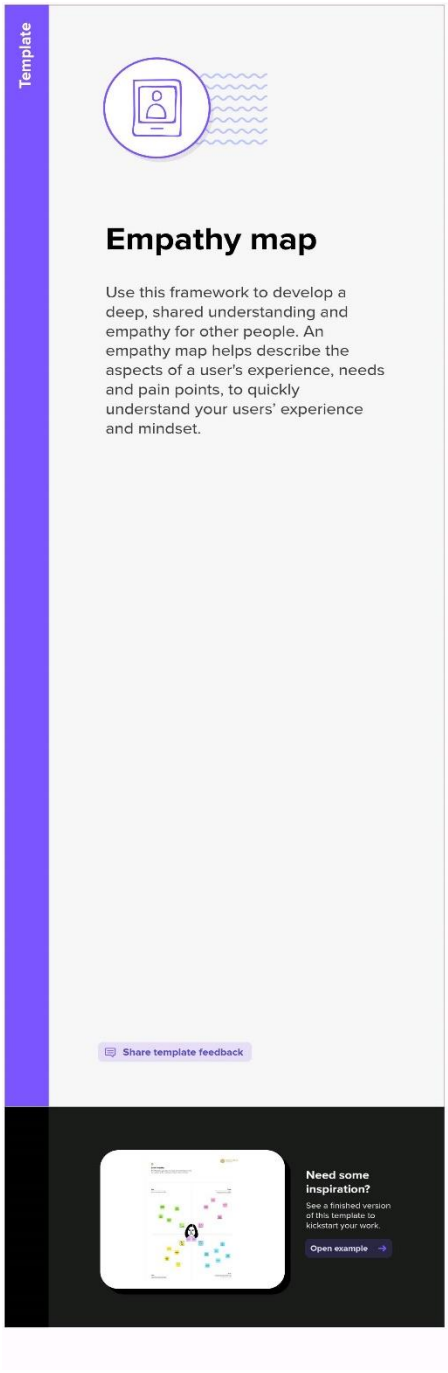
For classification, both machine learning and deep learning techniques were used to build an efficient and sensitive classification model with good accuracy and low false positive rate. Our result established that the combination of BERT for embedding and machine learning for classification produced better classification results than other combinations.

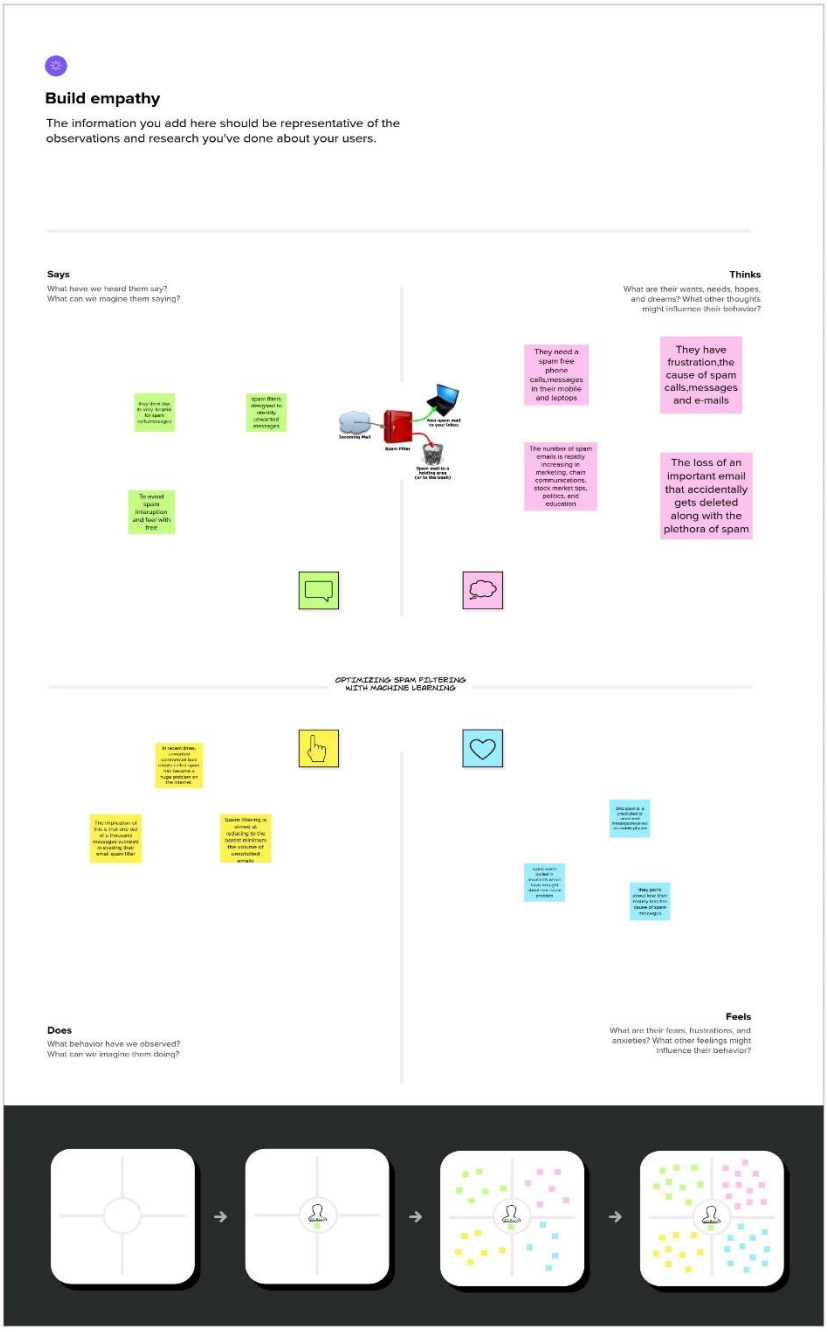
PURPOSE:

Machine learning algorithms have been extensively applied in the field of spam filtering. Substantial work has been done to improve the effectiveness of spam filters for classifying emails as either ham (valid messages) or spam (unwanted messages) by means of ML classifiers. They have the ability to recognise distinctive characteristics of the contents of emails. Many significant work have been done in the field of spam filtering using techniques that does not possess the ability to adapt to different conditions; and on problems that are exclusive to some fields e.g. identifying messages that are hidden inside a stage image. Most of the machine learning algorithms used for classification of tasks were designed to learn about inactive objective groups.

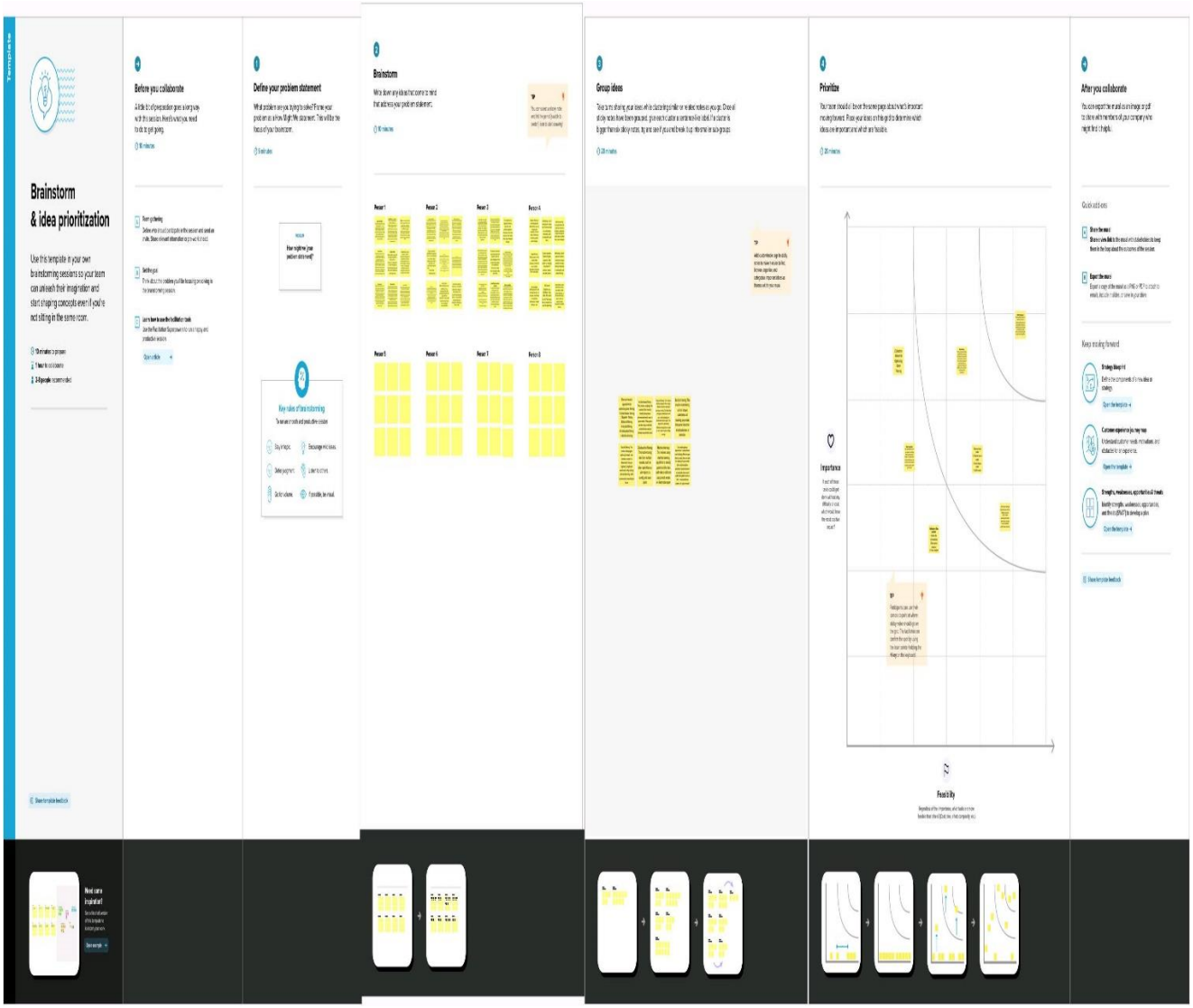
PROBLEM DEFINITION AND DESIGN THINKING

Empathy Map :


Template
Empathy map
Use this framework to develop a deep, shared understanding and empathy for other people. An empathy map helps describe the aspects of a user's experience, needs and pain points, to quickly understand your users' experience and mindset.
Share template feedback
Need some inspiration?
See a finished version of this template to kickstart your work.
Open example


Build empathy
The information you add here should be representative of the observations and research you've done about your users.
Says
What have we heard them say?
What can we magine them saying?
Thinks
What are their wants, needs, hopes, and dreams? What other thoughts might influence their behavior?
They need a spam free phone calls,messages in their mobile and leptops
They have frustration,the cause of spam calls,messages and e-mails
The number of spam emails is rapidly increasing in marketing, chain communications, stock market tips, politics, and education
The loss of an important email that accidentally gets deleted along with the plethora of spam
OPTIMIZING SPAM FILTERING WITH MACHINE LEARNING
Does
What behavior have we observed?
What can we imagine them doing?
Feels
What are their fears, frustrations, and anxieties? What other feelings might influence their behavior?

Ideation & Brainstorming map:



RESULT

Spam Detector For Emails!!
Enter Your Message Here
Predict
Great! This is NOT a spam message.

Advantages and Disadvantages

Advantages:

Protection against Viruses

Spam emails are not just innocent marketing tools they can be carriers of dangerous computer viruses. Just one click on the wrong email can debilitate your network. Filters can provide a great firewall.

Keeping Hackers at Bay

In addition to dangerous viruses, hackers can also gain access to your system through a benign-looking email. A filter that blocks spam emails from reaching your inbox can save your important data.

Saving Time

Spam filtering can save time. Business employees do not have to go through numerous emails to decide which ones are spam, as sometimes that can be hard to decide. The time saved can be used to increase productivity.

Keeping your Reputation Intact

Spam filters can help keep a company maintains its reputation. They can block viruses from reaching consumers' data and prevent any spam mail from accidentally being forwarded to consumers.

Customized Services

Anti-spam software and programs can be tailored to your needs. You can create a blacklist of email addresses that often send you spam. A whitelist contains all the email addresses of your important associates.

Lets you Be Sure

Many anti-spam filters offer the service of keeping the spam emails saved for a few days. It allows you to make sure that no useful emails are being deleted together with the junk mail.

Disadvantages:

Most people naturally become nervous as they see they got an email that is unwanted and that is instantly labeled as being spam. If you are a heavy email user there is a pretty good possibility you actually get hundreds of spam emails per week, or even per day, based on activity.

The common approach is to select the unwanted messages and then delete them. However, this is not always something that you want to do. What happens if you mistakenly identified a message as being spam? If this is the case, it is possible you will end up with some serious problems. As an example, if the boss sends you an email and you do not even open it, you might get fired.

Why Not Use Spam Filters

The biggest disadvantage of using an email filter is that you may end up with messages being identified as being spam through a mistake of the algorithm that is used.

While missing out on important emails is a nuisance, we need to think about the fact that you can also miss the same emails if you receive a lot of spam. How can you see that message from the boss if there are hundreds of emails sent every single day? You can be highly attentive and still miss out on some emails.

Application

When reading spam email to obtain keywords to create application level filters, remember to keep your *Internet* connection turned off or firewall locked so that images in the email don't display and broadcast your address availability back to the spammers (fortunately some email applications like *Thunderbird* have this protection built-in). The following guidelines describe how to set up application-level spam filters:

- Mailbox. Create a special mailbox called "Junk-filter", or something similar, into which you will direct the filtered mail instead of deleting it. Then every once in awhile scan the junk mailbox to gauge the success of your filters and make sure no legitimate email is being trapped. You can also use this archive to help tune your filters as described under *Selection* below.
- Filters. Now you are ready to set up a set of filters to recognize common spam keywords and then transfer the email into the Junk-filter mailbox. First set up an initial set of filters based on the spam you have been getting most recently, and constructed to trap the most common spam keywords and phrases.
- Maintenance. After you set up your initial set of filters, monitor the spam you still receive and add a couple of rules each time you check your mail to continually increase the efficiency of your filter engine. Add a few rules each time to catch the most common examples still making it through the filter engine. Over time, the amount of spam that makes it through will become less and less, and will be of increasingly unusual nature (ex: weird subject lines) that they will be easily recognizable as spam by the human eye and easily deleted.
- Selection. There are two basic goals when drafting a spam filter: make the rule broad enough to be effective at catching spam, and make the rule specific enough to avoid trapping legitimate email by mistake. The following guidelines assist in creation of good spam filter rules:
    - Tuning. You can occasionally go through your trash, which consists largely of spam you had to delete manually, sort the mailbox by subject, and look for common patterns and keywords. You can then create a few new filters to catch similar email, fine-tuning your engine to catch more of the spam that have been escaping your filter engine.
    - Efficiency. If you wonder if a rule is worth creating, you can use the filter mailbox as a useful archive to test the rule. Search the spam mailbox for the filter condition you are considering. If none (or very few) of the spam you've so far received match the condition, the rule is probably ineffective and not worthwhile.

- Safety. Always use guard rules as described above to minimize the chance of trapping legitimate email. However, you still want to remain accessible to the world and new correspondents.
- Fields. With most email applications, filters can target the sender, subject, message body, and other fields. However, because most fields can be faked, the subject and message body are the best for spam filters

CONCLUSION

we reviewed machine learning approaches and their application to the field of spam filtering. A review of the state of the art algorithms been applied for classification of messages as either spam or ham is provided. The attempts made by different researchers to solving the problem of spam through the use of machine learning classifiers was discussed. The evolution of spam messages over the years to evade filters was examined. The basic architecture of email spam filter and the processes involved in filtering spam emails were looked into. The paper surveyed some of the publicly available datasets and performance metrics that can be used to measure the effectiveness of any spam filter. The challenges of the machine learning algorithms in efficiently handling the menace of spam was pointed out and comparative studies of the machine learning technics available in literature was done. We also revealed some open research problems associated with spam filters. In general, the figure and volume of literature we reviewed shows that significant progress have been made and will still be made in this field. Having discussed the open problems in spam filtering, further research to enhance the effectiveness of spam filters need to be done. This will make the development of spam filters to continue to be an active research field for academician and industry practitioners researching machine learning techniques for effective spam filtering. Our hope is that research students will use this paper as a spring board for doing qualitative research in spam filtering using machine learning, deep leaning and deep adversarial learning algorithms.

FUTURE SCOPE

Efficient pattern detection in spam mail filtering plays crucial role. Using RFD model spam detection gives the spam patterns, non –spam patterns and general patterns which easily identify the whether the mail is spam or ham. The current method which uses the pattern detection method does not include the general patterns. RFD gives the general patterns of which user can decide to determine whether he wants to put the mail as spam or non-spam to avoid the loss of important mails. The images which are in forms of spams are also detected using File Properties, Histogram and Hough Transform. The current proposed system is for English language mails but as future scope we can design the system for multiple languages.

Hence there is scope for complete automation of spam detection systems with maximum efficiency. With grow- ing popularity of online stores, the competition also increases.

APPENDIX

SOURCE CODE: **..\Downloads\spam-filtering-ml-code**

D.M. Fonseca, O.H. Fazzion, E. Cunha, I. Las-Casas, P.D. Guedes, W. Meira, M. Chaves

Measuring characterizing, and avoiding spam traffic costs
IEEE Int. Comp., 99 (2016)

Y. Cao, X. Liao, Y. Li
An e-mail filtering approach using neural network
International Symposium on Neural Networks, Springer Berlin Heidelberg (2004), pp. 688-694

A. Bhowmick, S.M. Hazarika
Machine Learning for E-Mail Spam Filtering: Review, Techniques and Trends
arXiv:1606.01042v1 [cs.LG] 3 Jun 2016