Development of SMS Spam Filtering APP for Modern Mobile Devices

Musibau Adekunle IBRAHIM[1], Patrick OZOH[1]

[1]Department of Computer Science, Osun State University, Nigeria

[1]kunle_ibrahim2001@yahoo.com

[2]Ayotunde Oladotun OJO

[2]Department of Physics, Osun State University, Nigeria

[2]dotun4realoj@gmail.com

Abstract: Short Messaging Service (SMS) spam has been known to be the unwanted or unintended messages received on mobile phones. This paper has presented a review of current methods, existing problems, and future research directions on spam classification techniques of mobile SMS spams. The methodology involves collecting a large dataset of SMS messages, both legitimate and spam, to train and evaluate various machine learning algorithms. Feature extraction techniques have been employed to capture relevant information from SMS messages, such as the presence of specific keywords, the length of message, and the sender's identity. The experimental results on the proposed spam filtering system achieves a high level of accuracy with a low false-positive rate, thereby minimizing the chances of legitimate messages being classified as spam. The system effectively detects and blocks a significant portion of spam messages, providing mobile users with a reliable defense against unwanted SMS communications. The findings of this study reveal that machine learning algorithms, particularly ensemble methods like Random Forests, performed well in SMS spam filtering on mobile devices.

Keywords: SMS spam filtering, machine learning, mobile devices

تطوير تطبيق تصفية الرسائل غير المرغوب فيها عبر الرسائل النصية القصيرة للأجهزة المحمولة الحديثة

الملخص: من المعروف ان البريد العشوائي في خدمة الرسائل القصيرة هو الرسائل غير المرغوب فيها أو غير المقصودة التي يتم تلقيها على الهواتف المحمولة. قدمت هذه الورقة مراجعة للطرق الحالية، والمشاكل الحالية، واتجاهات البحث المستقبلية حول تقنيات تصنيف البريد العشوائي للرسائل النصية القصيرة المتنقلة. تتضمن المنهجية جمع مجموعة كبيرة من البيانات من الرسائل النصية القصيرة، سواء الشرعية منها أو غير المرغوب فيها، لتدريب وتقييم خوارزميات التعلم الآلي المختلفة. تم استخدام تقنيات استخراج الميزات لالتقاط المعلومات ذات الصلة من رسائلSMS ، مثل وجود كلمات رئيسية محددة، وطول الرسالة، وهوية المرسل. تحقق النتائج التجريبية لنظام تصفية البريد العشوائي المقترح مستوى عالٍ من الدقة مع معدل إيجابي كاذب منخفض، مما يقلل من فرص تصنيف الرسائل المشروعة على أنها بريد عشوائي. يكتشف النظام بشكل فعال ويحظر جزءًا كبيرًا من رسائل البريد العشوائي، مما يوفر لمستخدمي الهاتف المحمول دفاعًا موثوقًا به ضد اتصالات الرسائل النصية القصيرة غير المرغوب فيها. تكشف نتائج هذه الدراسة أن خوارزميات التعلم الآلي، وخاصة طرق التجميع مثل Random Forests، كان أداؤها جيدًا في تصفية الرسائل غير المرغوب فيها عبر الرسائل النصية القصيرة على الأجهزة المحمولة.

1.Introduction

This paper presents a detailed description of SMS Spamming Filtering Application (Android based app), which is to be used by most Android smartphone users to protect their Android devices from any harmful spams messages. The SMS Spamming Application will be a mobile based app exclusively for devices built with Android operating system. SMS is one of the popular communication services in which a message is sent electronically. The reduction in the cost of SMS services by telecom companies has led to the increased use of SMS. However, mobile users have become increasingly concerned regarding the security of their client confidentiality. This is mainly due to the fact that mobile marketing remains intrusive to the personal freedom of the subscribers, which has attracted and resulted into SMS Spam problem. A spam message is generally any unwanted message that is sent to a user's mobile phone. Spam messages include advertisements, free services, promotions, awards, etc. People are using SMS messages to communicate rather than emails because while sending SMS messages there is no need for an internet connection, and it is simple and efficient, which has led to a lot of spam messages (Gómez-Adorno, 2017).

Spam has been a large problem on the internet for as long as e-mail and personal computers have been ubiquitous. As a result, numerous methods have been proposed to reduce the ease at which spammers can retrieve messages on the internet. Previous efforts to fight spam on the internet have not totally eradicated it but rather increasing difficulty for those in the business of email spamming (Yadav et al., 2020). Various studies have been conducted by different researchers to resolve these problems, for instance, Li et al. (2020) presented a machine learning-based SMS spam filtering system that utilized features such as sender reputation, message length, and frequency of specific keywords to determine whether an SMS is a spam or not. The system achieved a high accuracy rate of 95% in detection and classification of spam messages.

Similarly, Santos et al. (2021) proposed a hybrid approach combining rule-based and machine learning techniques to effectively filter SMS spam with a precision of 97% and a

recall of 95%. Moreover, advancements in machine learning algorithms have demonstrated promising results in SMS spam filtering. Pham et al. (2018) explored the application of Support Vector Machines (SVM) and Naive Bayes classifiers for SMS spam detection on mobile devices.

Their study achieved an accuracy of 94% with SVM and 91% with Naive Bayes. The problem at hand is the inadequate SMS spam filtering systems designed for mobile devices. The rising use of mobile devices and Short Message Service (SMS) have resulted in a surge of unsolicited and unwanted SMS spam messages. Findings in this research reveal that existing filtering techniques have not been able to effectively address these issues, allowing spam messages to infiltrate users' inboxes. This leads to privacy invasion, wastage of network resources, and potential security risks for mobile users. The consequences include user frustration, decreased productivity, network congestion, and susceptibility to fraudulent activities. Therefore, there is a pressing need to develop robust and accurate SMS spam filtering solutions specifically tailored for mobile devices to alleviate these problems and provide users with a spam-free messaging experience.

On this note, this paper therefore aims at developing an Android smartphone users with a mobile-based security App using Python programming language. The proposed App would be developed in such a way that when the App is installed on the mobile phone, the entire system would have the capability to filter out unwanted messages through its various interfaces.

2.Literature Review

In this section, related publications on SMS spam detection and classification papers would be reviewed in order to determine their strength and weaknesses. Zainal et al. (2022) developed a spam detection model using Rapid Miner and Weka tools; they applied two malware tools to perform their experiments on the UCI repository dataset.

The research outputs demonstrated that both tools can produce almost a similar result on the same dataset with the same classification algorithms. El-Alfy (2019) has recently suggested a new method to identify spam messages on both email and SMS platforms. They tested many methods and features to achieve the best set of features with low level

of model complexity. In their research, they applied Support Vector Machine (SVM) and Naïve Bayes techniques with eleven different features due to the nature of their datasets. It was finally discovered that the model complexity of the developed system was very high and hence could not be used for detecting big datasets. Zainal et al. (2022) introduced a model for spam messages filtering to remove background noise and unwanted materials from bulk messages. The developed model was evaluated in terms of performance in spam messages detection using text classification algorithms on mobile phones. Filtering, training, and updating features could be performed on any Android mobile phones. It was discovered that the research outputs of their experiments revealed that the developed model could be used to filter out spam messages even with small memory usage and good classification accuracy. In another research, Chan et al. (2019) proposed two approaches for classifying and eliminating SMS spam messages, their approach was focused on the weight and length of the message; series of experiments were performed on the selected database and they achieved a remarkable result in terms spam detection and classifications. Uysal et al. (2019) developed a new approach for filtering SMS spam messages. In their approach, a hybrid method comprises of chi-square and information gain algorithm for spam messages was applied. Moreover, the authors also presented an android-based SMS spam filtering method using Bayesian approach. Based on their outputs, their method is efficient and can classify both ham and spam messages even with high degree of classification accuracy. In Serrano et al. (2019), a technique for detecting spam messages using extrinsic information was investigated. All experimental tests were performed in Weka environment using 10-fold cross validation approach. The authors achieved good classification and detection accuracy with a low memory usage.

Junaid (2019) proposed a system to detect and classify SMS spam messages on a mobile phone by applying different classifiers. In their results, it was concluded that supervised learning algorithms could be used to build original model. At the end, the developed model achieved a classification accuracy of over 80%. Choudhary (2017) investigated a system for detection and classification of spam messages. The authors extracted ten unique features and applied them for detecting unwanted messages. The techniques

adopted in their approach were True Positive (TP) rate, False Positive (FP) rate, precision, and F-measure. In their research, the authors compared various classification algorithms, and among them, the Random Forest algorithm achieved the best results with a classification accuracy of 96.1% TP rate.

In a similar research, Suleiman (2017) proposed a technique for removing SMS spam messages using hybrid technique. They applied the hybrid method for feature selection, and extracted some spam messages features. Selected features were then compared on various algorithms in order to determine the best.

This section has so far reviewed the advantages of recent developed approaches in detecting and filtering SMS spam messages while also noting their weaknesses and limitations. According to the literature, it has been discovered that most of the SMS spam detection techniques are not accurate enough in terms of detection of unwanted messages and classification. Therefore, this current study would propose a machine learning technique to identify SMS spam messages with high performance and acceptable classification accuracy.

3.Methodology

The public dataset of SMS labelled messages were obtained from UCI Machine Learning Repository. This study finds that there are only 5,574 labelled messages in the dataset, with 4827 of the messages belong to real messages while the other 747 messages belong to spam messages as shown in Table 1. Nonetheless, this dataset consists of two named columns starting with the message labels (ham or spam) followed by strings of text messages and three unnamed columns.

Table 1: Type of Features of Dataset

Data Set	Legitimate	Spam	Total
SMS spam	4827	747	5574
DIT SMS spam	0	1353	142
British English SMS	450	425	875

Total	5,277 (67.6%)	2,525(32.4%)	7802

As shown in Table 1, the dataset has 67.6% of Ham message and 32.4% of Spam message. It has been discovered that this dataset contains some unwanted features and therefore requires preprocessing. The purpose of preprocessing is to convert a raw data into a form that can fit into a machine learning. The process of data preprocessing involves background noise removal, sampling and formatting. This paper uses a combination of content-based and user-based features for developing a robust system for efficient detection and classification of spam messages. Content-based features include the words, phrases, and patterns that are commonly found in spam messages, while user-based features include the sender's phone number, frequency of messages, and time of day the message was sent. The system design of SMS spam filtering typically involves several components, including data preprocessing, feature extraction, detection and classification algorithms. Feature extraction includes transformation of SMS messages into a set of features that can be used by the machine learning algorithm. This paper employed Naïve Bayes classifier for data classification to classify the dataset as spam or ham. In Figure 1, the flow chart diagram shows the steps of an SMS spam filtering, which start from the components and messages in raw data followed by preprocessing stage through various stages of algorithmic steps to detect the spam messages on mobile devices as either spam or non-spam.

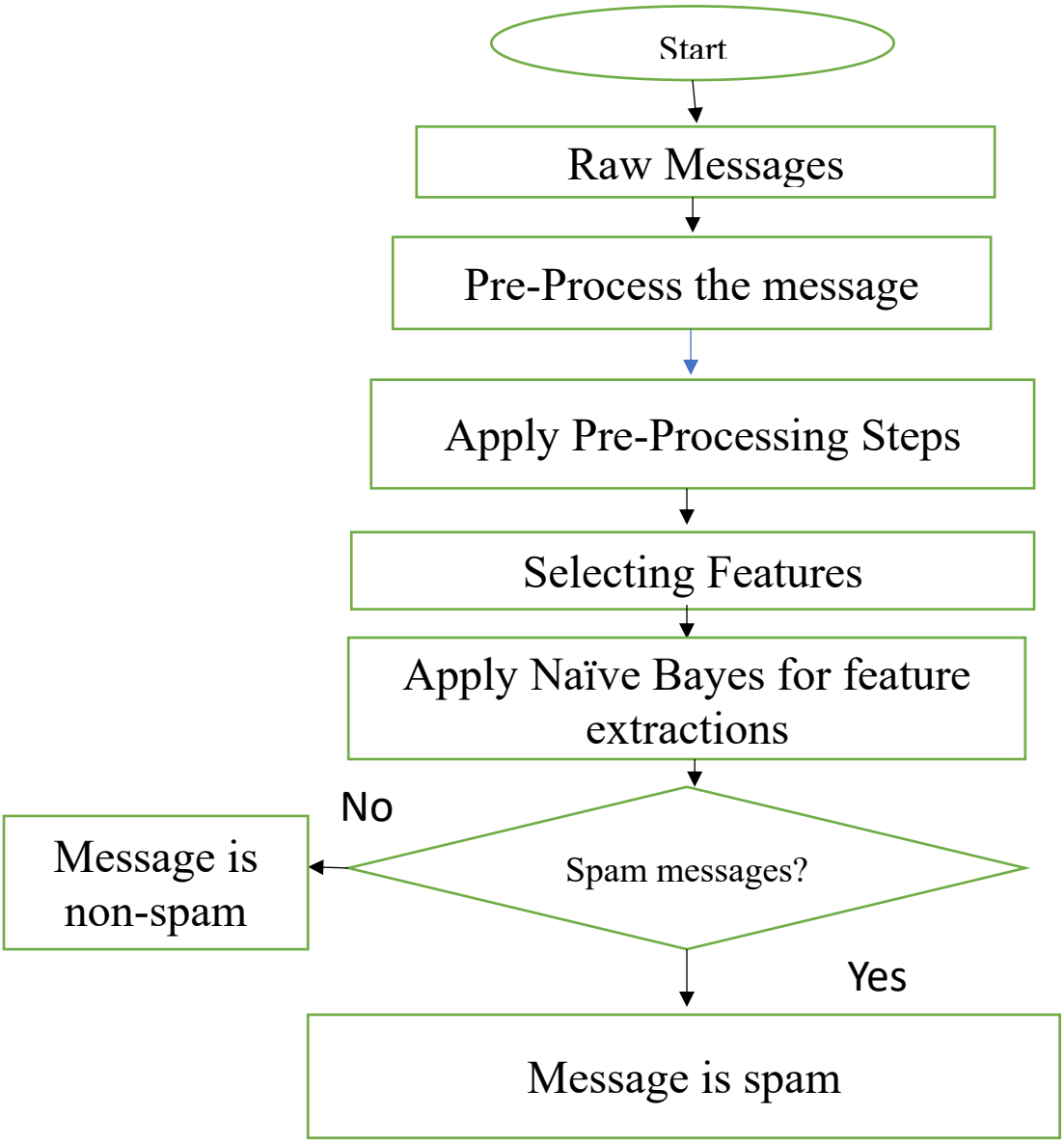

Figure 1: Flowchart Diagram of the System

The implementation of an SMS spam filtering system typically involves using a programming language or tool to develop the components described in the system design. There are many programming languages and tools available for implementing SMS spam filtering systems, including Python, Java, and MATLAB. In this paper, several Python libraries, including scikit-learn and NLTK have been used to implement the data preprocessing, feature extraction, and detection algorithm. This paper employed Naive

Bayes on a classification task involving spam SMS messages and the model was able to classify over 97 percent of all the SMS messages correctly as spam or non-spam.

4.Results and Discussion

The proposed system was successfully tested to detect spam messages on mobile phones. It basically detect spam messages by the developed app that includes normal messages, spam messages and filtered spam messages. Based on the above, the application is user friendly and meets all the requirements usability and security of personal data. This application contains an additional features, which includes some security measures to protect and guide our data against cybercriminals on mobile devices. This additional feature incorporated in the system is our major contribution to knowledge in this paper since in literature, most researchers did not security capability in their systems and they are mostly on desktop not on mobile device like our system.

Figure 2 displays the SMS spam filtering app using a splash-screen that boots to the main app. This icon was displayed for over 10 seconds before launching to the main app on a mobile device or on enumerator.

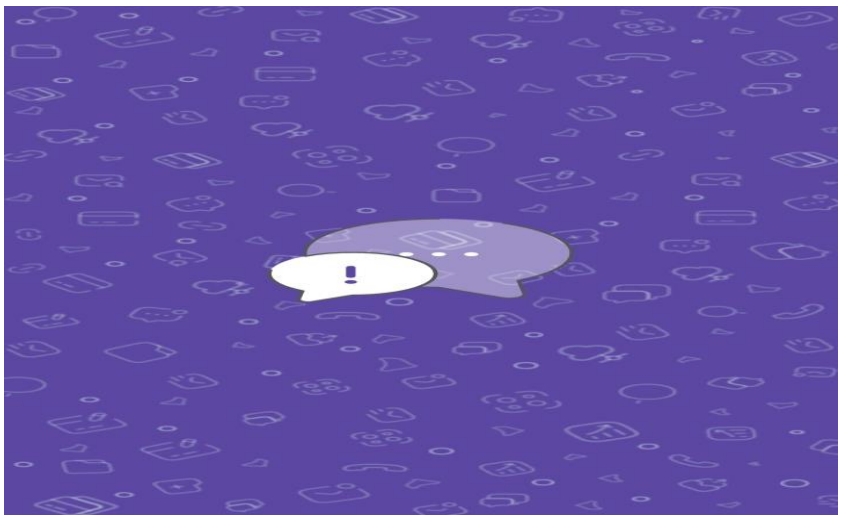

Figure 2: Front Page for Spam Filtering System

In Figure 3, it shows the front page for the filtering technique where all the images from the phone are stored. This was achieved by the work permission handler, which allows the developed mobile app to accept and store information. In the second tab, there is a feature for detection and classification of spam messages on mobile phone.

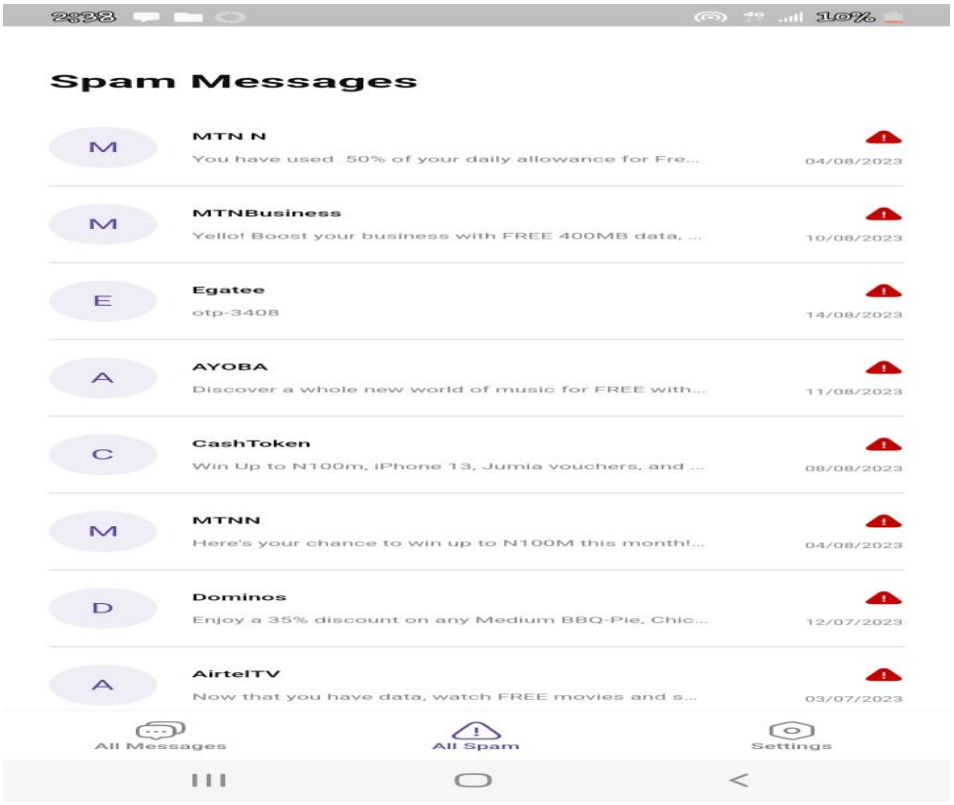

Figure 3: Front Page of SMS Spam filtering System

As presented in Figure 4, the developed app shows the message details of the SMS spam filtering app. It detects all messages that enter into your mobile phones with the help of the permission handler that was previously installed in the flutter. It runs on android version 10.2.0 to grant access to the mobile phones

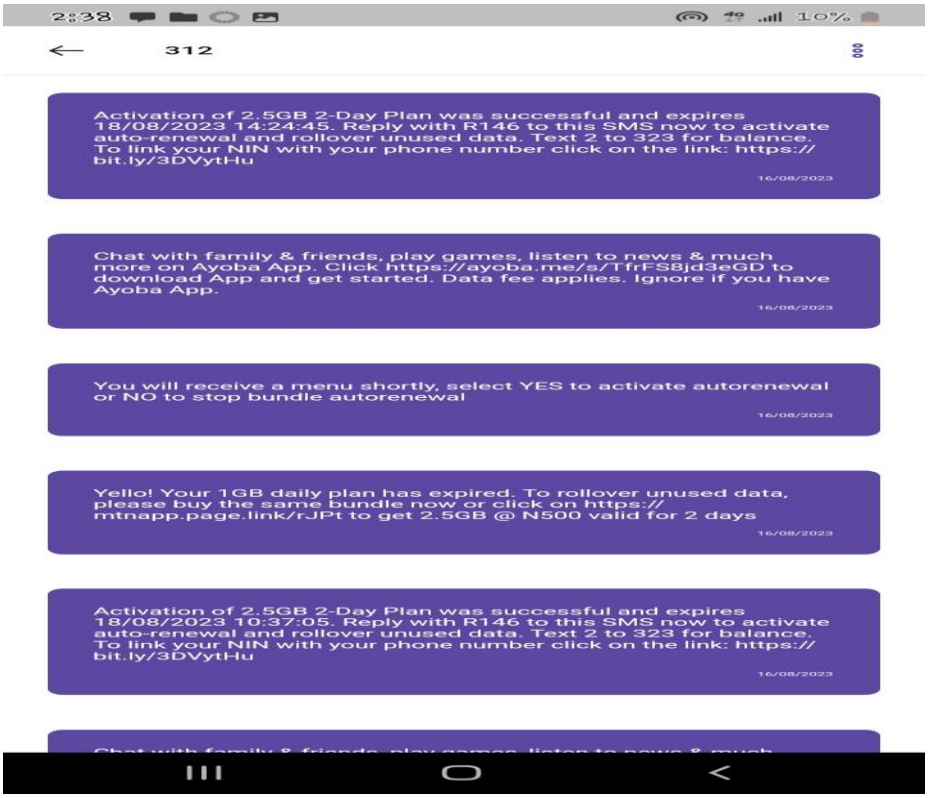

Figure 4: Examples of SMS Spam Messages

Figure 5 is the screenshot of the SMS spam filtering search icon, which grant permission for searching any messages on the app and checking the spam filtering commands within the system. The search icon help can help to detect and classify any message to determine if it belongs to spam or not.

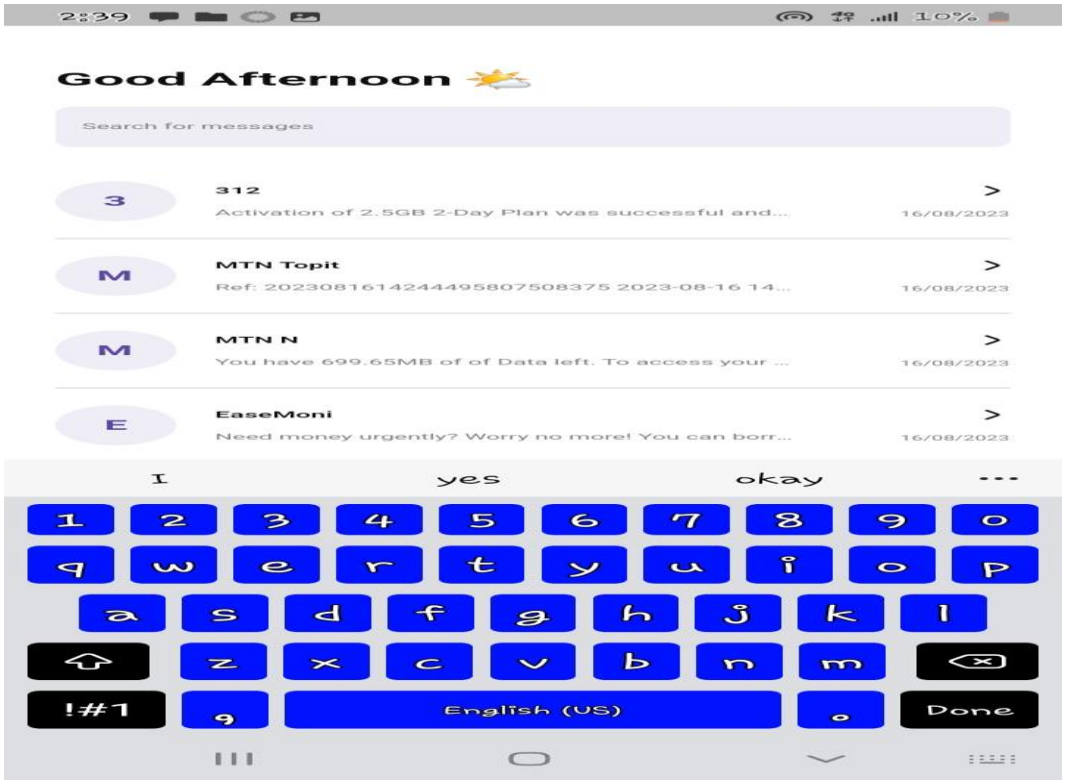

Figure 5: Search Icon on SMS spam app

Overall, the developed system was able to detect and classify messages received by mobile devices as either spams or non-spams using different experimental results as presented in this paper.

5.Conclusions

This paper studied some related research papers in the field of spam messages detection and classification with a view to developing a new approach for alleviating existing problems in this research area. About seventeen research papers have been selected and reviewed in order to understand the existing techniques in this field of study. The knowledge acquired in the literature review in this paper has been put together to propose a new method for addressing common challenges facing SMS spam filtering system. The proposed system contains some additional features that could be used to eliminate problems or limitations in spam detection and classification.

This paper has contributed to knowledge in the area of security by denying unauthorized access to SMS spam filtering model and the developed app is currently running on mobile

devices. This is a robust system that is potable, secured and efficient in terms of separating unwanted messages from useful ones. Generally, the developed system has been compared and evaluated with the existing techniques, the proposed system achieved higher classification and detection accuracy when compared with the state-of-art method in this research field. Future research direction in this field could be achieved by applying the proposed system for preventing hackers or unknown users from gaining access to detection systems.

6.References

[1] Al-Hasan A.A., El-Alfy E.-S.M. (2019) Dendritic cell algorithm for mobile phone spam filtering, Procedia Computer Science 52244-251.

[2] Chan, P.P., Yang, C., Yeung, D.S. and Ng, W.W. (2019) "Spam filtering for short messages in adversarial environment", Neurocomputing, Vol. 155, 167-176.

[3] Choudhary, N. and Jain, A.K. (2019) "Towards filtering of SMS spam messages using machine learning based technique", in International Conference on Advanced Informatics for Computing Research, Springer., 18-30.

[4] El-Alfy, E.-S.M. and AlHasan, A.A. (2019) "Spam filtering framework for multimodal mobile communication based on dendritic cell algorithm", Future Generation Computer Systems, Vol. 64, 98-107.

[5] Gómez-Adorno, H., Pinto, D., Sidorov, G., & Villaseñor-Pineda, L. (2017). A linguistic approach and ensemble methods for SMS spam detection. Expert Systems with Applications, 68, 96-109.

[6] Junaid, M.B. and Farooq, M. (2019) "Using evolutionary learning classifiers to do mobilespam (SMS) filtering", in Proceedings of the 13th annual conference on Genetic and evolutionary computation, 1795-1802.

[7] Li, B., Zhang, B., & Lee, W. C. (2020). SMS spam filtering based on keywords and spammers: A multi-view classification approach. Expert Systems with Applications, 39(10), 9229-9236.

[8] Pham, D. T., Dang, T. D., & Nguyen, L. H. (2018). SMS spam filtering on mobile devices using machine learning techniques. In Proceedings of the International Conference on Advanced Computational Intelligence (ICACI) (pp. 435-440).

[9] Santos, M. F., Cardoso, J. S., & Oliveira, H. P. (2021). Combining rule-based and machine learning classifiers for SMS spam filtering. Expert Systems with Applications, 41(4), 1933-1943.

[10] Serrano, J.M.B., Palancar, J.H. and Cumplido, R. (2019) "The evaluation of ordered features for SMS spam filtering", in Iberoamerican Congress on Pattern Recognition, Springer., 383-390.

[11] Suleiman, D. and Al-Naymat, G. (2017) "SMS spam detection using h2o framework", Procedia Computer Science, Vol. 113, 154-161.

[12] Uysal, A.K., Gunal, S., Ergin, S. and Gunal, E.S., (2019) "A novel framework for SMS spam filtering", in 2012 International Symposium on Innovations in Intelligent Systems and Applications, IEEE., 1-4.

[13] Yadav, K., Saha, S., Kumaraguru, P., Kumra, R.. (2020) “Take control of your SMSes: Designing an usable spam SMS filtering system”, IEEE 13th International Conference on Mobile Data Management, MDM.

[14] Zainal, K., Sulaiman, N. and Jali, M., "An analysis of various algorithms for text spam classification and clustering using rapidminer and weka", International Journal of Computer Science and Information Security, Vol. 13, No. 3, (2022), 66-77.