Towards Filtering of SMS Spam Messages Using Machine Learning Based Technique

DR. USMAN AIJAZ N[1], SUFIA ANJUM[2], SYED SAMEER AHMED[3], TASMIYA ILYAS[4], SYED HANNAN[5]

[1]*Faculty, ISE Department, HKBK College of Engineering, Bangalore*

[2, 3, 4, 5] *Students, ISE Department, HKBK College of Engineering, Bangalore*

Abstract— The popularity of mobile devices is increasing day by day as they provide a large variety of services by reducing the cost of services. Short Message Service (SMS) is considered one of the widely used communication service. However, this has led to an increase in mobile devices attacks like SMS Spam. In this paper, we present a novel approach that can detect and filter the spam messages using machine learning classification algorithms. We study the characteristics of spam messages in depth and then found ten features, which can efficiently filter SMS spam messages from ham messages. Our proposed approach achieved 96.5% true positive rate and 1.02% false positive rate for Random Forest classification algorithm.

I. INTRODUCTION

Message Service 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. This rise attracted attackers which have resulted in SMS Spam problem. A spam message is generally any unwanted message that is sent to 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 message there is no need of internet connection and it is simple and efficient. The SMS Spam problem is increasing day by day with the increase in the use of text messaging. There are various security measures available to control SMS Spam problem but they are not so mature. Many android apps are also on play store to block spam messages but people are not aware of these apps due to lack of knowledge.

Other than apps the filtering techniques available mainly focuses on email spam as email spam is one of the oldest problem but with the popularity of mobile devices, SMS spam is the one of the major issue these days. SMS is one of the cheapest ways to communicate and can be considered as the simplest way to perform phishing attacks as mobile devices contain sensitive and personal information like card details, username, password, etc. Attackers are finding different ways to steal this information from mobile devices.

SMS spammers can purchase any mobile number with any area code to send spam messages so that it becomes difficult to identify the attacker. US tatango learning center provided the list of top 25 SMS Spam area codes used by spammers National Fraud Intelligence Bureau (NFIB) published a media report about the latest scams which was analyzed by action fraud in 2016.

Spammers are targeting bank customers these days by sending spam messages for asking their bank account details, ATM pin number, password, etc. and the customer thinks that the message is coming from the bank and he/she may give all the details to the spammer. A report was published by ACMA that how bank customers are becoming the victim of SMS Spam attacks. In our proposed approach main aim is to filter the spam and ham SMS using machine learning algorithms. We have used a feature set of 10 features for classification. These features can differentiate a spam SMS from ham SMS.

Machine learning techniques were effective in email spam filtering as it helps in preventing zero-day attacks and provides the high level of security. The Same approach is being used for mobile devices in order to prevent from SMS Spam problem but in the case of SMS Spam features will be different from email spam as the size of the text message is small and

the user uses less formal language for text messages. And text message is simple without any graphic content and attachments.

II. BACKGROUND STUDY (LITERATURE)

After the evolution of Machine Learning algorithm and its usage in document classification, a lot has berm research done on identifying the features of text. In this section authors have described the work done hy researchers in field of identifying Texts features by limiting their study solely on the field of SPAM identification. M. Nivaashini et al has used various Deep Neural Network (DNN) techniques in identifying the SPAM and HAM after collecting the dataset from UCT Machine Leaming Repository. Authors have compared all the used algorithms based on their accuracy, False-Positives, False Poutives and high chances for identifying SPAM with low Fahe Positive rates, in order to identify the best algorithm.

De. Dipak R Kawade, De. Kavita S. Oza" have identified SMS SPAM using spam filtering techniques, by uning open source python software, they have achieved 95% accuracy. For studying and preprocessing they have used WEKA 100 P. Navancy" et. Al has used various supervised based machine learning algorithms like Naive Bayes, Support Vector Machine Algorithm and Maximum Entropy Algorithm, and they have done an accuracy comparison, and it was found that SVM was having more accuracy.

Bichitrananda" et al have used various ML algorithm like SVM (Supporn Vector Machine), Decision Tree, KNN (K-Nearest Neighbor), Neural Network (including Back-Propagation, Perceptron, Stochastic Gradient) for automatic classifying text documents on Datasets obtained from 20Newsgroup, IMDB, BBC News & BBC Sports, also they have compared the performance of all the Algorithms using metrics such as Kappa Statistics, Error Rate, Precision Call, Accuracy, F-Measure Bichitrananda et al have built an automated documen classifier for biomedical data seta (like TREC 2006 genetic Track, Farm-Das, Bio Creative Corpus (1) using ML. algorithma. All the Algorithms used for the task were evaluated and compared on the basis of ML. Classification metrics like accuracy, precision, rocall & -messure. Leila Armas" et al have demonstrated a method to extract the abstract out from the document asing Machine Learning Algorithms like Convolutional Neural Network & SVM Classifier. Francis M Kale et al have proposed a framework for performing text mining & text clustering used the K. Means algorithm and its application in various areas.

This paper gives guidance to rescarchers for sest clustering being the state of the art of text mining Ting SL et al performed text mining on vast and large datasets using various classification-based Machine Learning algorithms like decision tree, neural network, SVM (support vecur machines) and also compared each of the classifiers on the basis of computational efficiency and accuracy. Naive Bayes wa found to be the best & efficient classifier amongst all other classifiers.

III. METHODOLOGY

Data Collection: In this phase authors have collected a dataset based on which they have performed the experimentation from Kaggle Repository".

Data Cleaning: In this phase the authors have cleansed all the data which were taken into consideration. Authors have removed all the white-spaces, lowered the alphabet so that words like Equal and equal become the same, remove the remaining punctuation, like! is not that much important, tokenize each message, to represent the message as a list of words and done stemming, converting all the words to their root word, like floor, floored to floor.

Generating Testing and Training Data Sets: Authors have created the testing and training data on the converted cleansed datasets.

Generating Word Cloud Vector: Authors have used the TF-IDF vectorization for creating the word-vector. On the basis, the spam feature will be classified.

Prediction: Authors have given input messages to check whether the message is SPAM or HAM

IV. IMPLEMENTATION

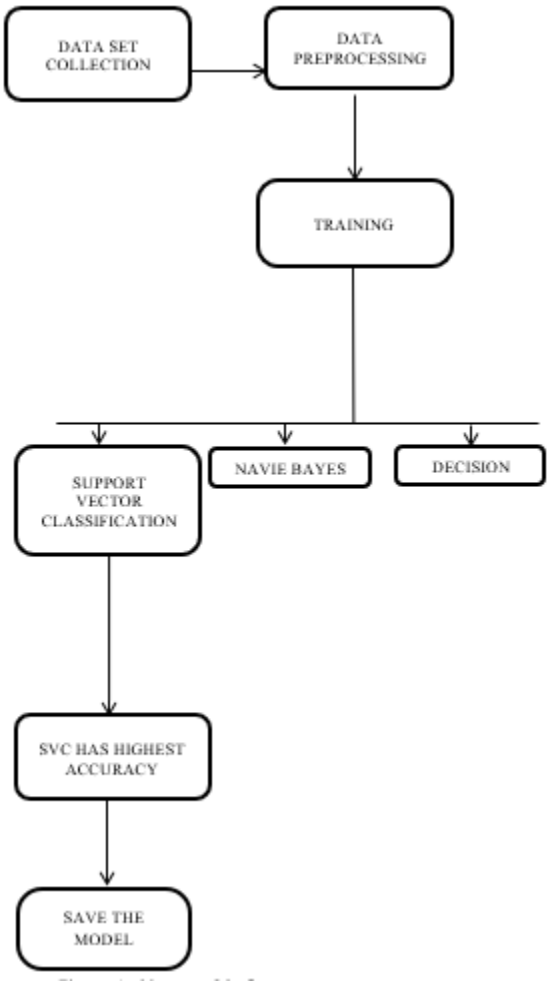

Figure: Architecture of the System

SMS Data Collection involves gathering a dataset of SMS messages, where each message is labeled as either spam or ham (legitimate). The dataset serves as the foundation for training and evaluating the machine learning models.

Once collected, the SMS data undergoes preprocessing, including cleaning and transformation. This step ensures that the data is in a suitable format for further analysis and feature extraction.

Features relevant to distinguishing between spam and ham messages are extracted from the preprocessed data. These features may include word frequency, message length, and the presence of specific keywords associated with spam.

Various machine learning algorithms are considered for classification, including Logistic Regression, XGBoost, Support Vector Classification (SVC), Decision Tree, and Naive Bayes. Each algorithm has its own strengths and weaknesses, making them suitable for different scenarios.

The selected models are trained using the preprocessed data and evaluated using cross-validation techniques. Evaluation metrics such as accuracy, precision, recall, and F1-score are calculated to assess the performance of each model.

Once a satisfactory model is identified, it is deployed for real-time classification of incoming SMS messages. This ensures that spam messages are detected and filtered before reaching the users' inbox.

CONCLUSION

The Abuse word problem is increasing nowadays with the increase in the use of text messaging. SMS filtering is the big challenge these days. We propose a technique for SMS Spam filtering based on 10 feature using three machine learning algorithms namely Naive Bayes, Decision Table and SVC. The dataset that we have used in our work consists of 2608 messages out of which 2408 messages were collected from the SMS Spam Corpus v.0.1 publically available and 200 messages collected manually. Out of all classification algorithms, SVC Algorithm gives best results with 96.1% true positive rate.

ACKNOWLEDGEMENT

We would like to thank, our guide Dr. Usman Aijaz N and HOD Dr. Syed Mustafa for their valuable suggestion, expert advice and moral support in the process of preparing this paper.

REFERENCES

[1] Mobile Commons Blog. https://www.mobilecommons.com/blog/2016/01/how-textmessaging-will-change-for-the-better-in-2016/

[2] SMS Blocker Award. https://play.google.com/store/apps/details?id=co

m.smsBlocker&hl=en

[3] TextBlocker. https://play.google.com/store/apps/details?id=com.thesimpleandroidguy.app.messageclient&hl=en

[4] Androidapp. https://play.google.com/store/apps/details?id=com.mrnumber.blocker&hl=en

[5] Puniškis, D., Laurutis, R., Dirmeikis, R.: An artificial neural nets for spam e-mail recognition. Elektronika ir Elektrotechnika 69, 73–76 (2006)

[6] Jain, A.K., Gupta, B.B.: Phishing detection: analysis of visual similarity based approaches. Secur. Commun. Netw. 2017 (2017). Article ID 5421046. doi:10.1155/2017/5421046

[7] Gupta, B.B., Tewari, A., Jain, A.K., Agrawal, D.P.: Fighting against phishing attacks: state of the art and future challenges. Neural Comput. Appl. 1–26 (2016). doi:10.1007/s00521-016-2275-y

[8] Jain, A.K., Gupta, B.B.: A novel approach to protect against phishing attacks at client side using auto-updated white-list. EURASIP J. Inf. Secur. 1–11 (2016). doi:10.1186/s13635-016-0034-3

[9] Choudhary, N., Jain, A.K.: Comparative Analysis of Mobile Phishing Detection and Prevention Approaches (Accepted)

[10] Tatango Learning Center. https://www.tatango.com/blog/top-25-sms-spam-area-codes/