IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 6, June 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

SMS Spam Detection

[1] Sheeba Selvapattu [2] Mrs. Pallavi V Patil

[1] Department of Msc IT, Jain (Deemed-to-be-University) Bangalore.

[2] Assistant professor, School of CS & IT, Jain (Deemed-to-be-University) Bangalore.

Abstract

The development of the cell phone clients has prompted an emotional increment in SMS spam messages. The developing volume of spam messages has expanded the interest for accurate and effective spam solutions. SMS spam filtering is a generally new tasks which inherits numerous issues and arrangements from email spam separating. Many spam arrangements have been proposed in the recent days. The one which we address in this postulation, treats spam detection as a basic two class document classification problem. The classification will comprise of classification algorithm with feature extractions. Classification and feature helped us in improving the performance as far as exactness and has lesser computational time and capacity prerequisites. We proposed the comparison between different classifiers by establishing different machine learning techniques. Various classifiers are used for best result. The different classifier used in this work are K Nearest Neighbour, Naïve Bayes, Decision Tree.

I. INTRODUCTION

SMS is a text communication platform that allows users to exchange messages. Users may also receive a commercial advertisements to their mobile through text messages. The amount of spam is depends from region to region.

There is a huge difference between spam-filtration in text messages and emails-filtration. For email large dataset are available whereas in SMS the dataset are limited. SMS that allows phone users to exchange short text messages. Since the length of text is small, the number of features that are used for classification is comparatively smaller then email. Mostly text Messages are full of abbreviation and have less formal language.

SMS spam is pretty complicated by several factors, it include lower rate of SMS Spam, which is allowed many users and service providers to ignore issue. The spam-filtering software for mobile are very limited. Filtering SMS spam at the recipient device will not be a perfect solution. In commercial markets users are paid to receive a messages. This problem however come by comparing the users that are not charged to receive messages.

Spams are undesirable and unwelcomed messages which are sent electronically. These messages are sent by spammers for different reason. This are done to take user's personal data. Mobile spam are not same in very region it may vary. The phone number cannot send more than 200 messages per hour and not more than 1,000 messages per day. Thus the SMS spammers have adapted this strategies in innovative ways. More effective approaches are need to be used in order to filter SMS spam accurately.

Types of Mobile Messaging Attacks

☐SMS Spam:

This is one of the most basic form of attack where the unwanted messages are sent by the attackers to the end users for bulk marketing and Social Engineering Viral Hoaxes.

☐Premium rate fraud:

These are uninvited messages which are sent to trick customers to call premium rate numbers or sign up for the subscription services that are charged to the bill for people.

"Congratulation! Your cell number has won 1,00,00,00 Rupees in the ongoing sony ericsson mobile promo. For claim call 91-9865948165" ☐ Phishing:

These messages ask the user to call to specific number by which attacker obtain the information of the user and later misuse these information in wrong way. ☐ Malware

Malware is a piece of software that was written with the intent of damaging devices, stealing data, mainly to create a mess. Malware it's a malicious software that penetrates the mobile devices without user knowledge. It comprises of sending unsolicited links to the user and request them to download the executable file, which is dangerous and leads to application abuse.

- Virus
  Viruses attach themselves to clean the files and infect other files, they reproduce itself .virus usually appear in executable file(.exe).they can spread widely, causing damage to the functionality of systems, and deleting or corrupting files and locking user out of their computers.
- Worms
  Worms infect entire networks of devices, it could be local or across the internet, by using network interface. It uses its consecutively infected machine to infect others. This type of malware can infect entire networks of device very easy.
- Trojan horse
  It is one of the most dangerous malware. It represents itself as something useful in order to trick you. Once it enters your system, the attackers behind the Trojan gain unauthorized access to the affected computer. From these it can steal information or install threats like virus and ransomware.

II. Related Work

Pradeep Kumar Roy et al [1] discussed on filtering sms spam efficiently, deep learning model was proposed to filter sms spam such as CNN and LSTM along with machine learning classifiers such as NB, RF, GB, LR, and SGD classifier are tested. And the comparative study with two different models of deep learning were performed such as CNN and LSTM and the

result confirmed that the CNN model outperformed the LSTM model. The work was fully dependent on text messages that was written in English only.so the future research can employ other languages in order to filter spam and not-spam text messages.

Naresh Kumar Nagwani et al [2] discussed about the problem of sms spam detection and thread identification. The art clustering-based algorithm are used in this work. It has two stages, in first stages the binary classification technique such as NB, SVM, LDA and NMF is used to categorize the sms into spam or ham sms, the second stages sms clusters are created for ham sms using non negative matrix factorization and K-means clustering techniques. The sms spam detection and thread identification are used in many of sms activities such are SMS folder classification, SMS classification and SMS thread summarization. sms threads use two levels, the first is classification and second is clustering. Sms threads consists of sms messages, so it can recognize the previous communication in a message. NMF clustering technique performs better than K-means clustering techniques in terms of number of SMS messages participating in threads identified.

Miloud Aboubakeur El Sadek Mokri [3] Proposed a different method know as heuristic technique based on the natural function of the octopod to filter and detect spams. octopods are most defensive animal that have the ability to protect from predators, similar to these ability the intelligent system of security has been performed. Adapting the sms spam problem to the behaviour of the octopods. And it has been proven the octopods are best mechanism to detect sms spam messages.

Shafi Muhammad Abdulhamid, Muhammad Shafie Abd Latiff [4] The sms spam causing lot of issues in marketing environment by losing of subscribers due to sms spam. several different solution are proposed to detect and filter sms spam. Comparison is done on different sms spam technique it has been proven that support vector machine and Bayesian network can be used as classifiers for sms spam. And also compared with bio-inspired algorithms such as Monkey Search, Cat Swarm, Magneto Tactic Bacteria Optimization based on Moment Migration, Chicken Swarm Optimization, the Bat algorithm, the Cuckoo search algorithm, the Bees algorithms, and Particle Swarm Optimization have concluded this algorithm are not able to use for sms spam classifiers.

Neelam Choudhary [5] proposed different machine learning algorithm in order to detect and filter sms spam messages. Have found 10 features in sms spam messages in order to filter sms spam efficiently from ham messages. A comparison on 5 different machine learning algorithm namely Naïve Bayes, Logistic Regression, J48, Decision Table and Random Forest. Out of 5 classifiers random forest algorithm gives the best accuracy.

Dima Suleiman [6] proposed a new classifier know as H2O platform to make a comparison between different machine learning algorithm such as random forest, naïve bayes and deep learning. Different parameters were used to compare among the algorithm such are accuracy, f-measure, precision, recall and runtime. And from the result naïve bayes has high performance in term of runtime when compared with other two algorithm. Whereas for accuracy, f-measure, precision, recall the random forest has highest performance.

Hao Chen [7] discussed a comparative study on different algorithm such are Recurrent neural network, convolutional neural network, naïve bayes on the tensorflow platform.in order to improve the accuracy for the spam filteration the cyclic neural network is used. And comparison is done.

Nurulhuda Firdaus Mohd Azmi [8] proposed a method for filtering spam messages since sms classification are becoming more challenging due to the complexities of the spammers. The methods of term frequency-inverse document frequency (TF-IDF) and Random Forest Algorithm will be applied on data and found the accuracy among them. Only accuracy cannot determine the performance of the algorithm. Hence determining the precision, recall and fmeasure of the algorithms are been observed. Performance of the algorithm various based on the features used in the data set.

III Methodology

This segment describes the general structure of work process of the experiment. In this examination AI instrument is utilized for the analysis and classification of the dataset. At the principal level information is assembled from various sources to make a decent dataset of ham and spam in text format and give that information as the input for the model. At the second degree of the investigation we changed over the informational collection which is prior in the text format to CSV (Comma Separated Value). At that point pre-processing is accomplished for a superior quality info either by removing of unrequired words or by performing stemming on them. Then the pre-processed data information is changed into a machine readable form or non-contextual form by changing over to vector or by doing discretization. The labeled data is opened and the attributes are recorded. The attributes that are utilized for the investigation intention are text and class in this dataset. From that point forward, a classifier is applied to the dataset we have used. Hence the information is trained utilizing the dataset. Testing is performed on the testing data to get the conclusive results. At the last step of the experiment, Confusion Matrix are acquired from the dataset and the results of the applied classifier are investigated and talked about. The flow of this work is given below Fig 1

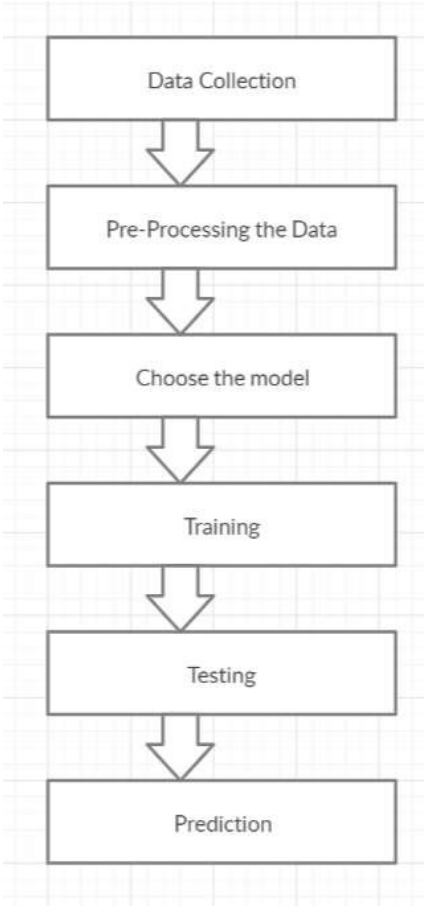


Fig 1 process of spam filtering

A. KNN Classification

K Nearest Neighbors (KNN) is straightforward and its basic, adaptable and is one of the highest machine learning algorithms. KNN utilized in many ways, for example, account, medicinal services, political theory, picture recognition and video recognition. In Credit evaluations, budgetary organizations will anticipate the FICO assessment of clients. In credit rating, banking organizations will decide whether the loan is protected or dangerous. In political theory, arranging potential voters in two classes will cast a vote or won't cast a vote. For classification and regression problem KNN algorithm is used. KNN calculation dependent on the similarity approach.

B. Decision Tree Algorithm

A decision tree is a flowchart-like tree structure where an internal node represents attribute, decision rule is represented as branch, and each leaf node represents as the outcome.

The root node is the topmost node in a decision tree. The partition is done based on the attribute value. The partition is done in recursively way. This flowchart-like structure encourages you in decision making. It's perception like a flowchart chart which effectively mimics the human level reasoning. That is the reason decision trees are straightforward and translate.

C. Naive Bayes

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 7 Issue 6, June 2020
ISSN (Online) 2348 – 7968 | Impact Factor (2020) – 6.72
www.ijiset.com

Naive Bayes is technique based on Bayes Theorem it is statistical classification technique. It is one of the most straightforward supervised learning algorithms. . Naive Bayes classifier is the quick, precise and reliable algorithm. Naive Bayes classifiers have high exactness and speed on enormous datasets.

Naive Bayes classifier accept that the impact of a specific component in a class is autonomous of other features. For instance, a credit candidate is alluring or not relying upon his/her salary, past advance and exchange history, age, and area. Regardless of whether these features are associated, these features are as yet considered autonomously. This assumption improves calculation, and that is the reason it is considered as naive.

IV RESULTS AND DISCUSSION

A. Data set used

The Dataset used in this study are of dataset prepared in UCI [9] and the dataset includes 5574 SMS. It has labeled into two groups ie., 4827 SMS messages Ham and 747 SMS messages are spam that is knows as unwanted messages. As shown in below Fig 1

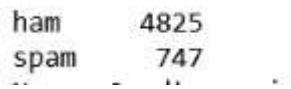


Fig 1

B. Implementation

In this paper for implementation, Data preprocessing are been used in order to clean the dataset. Different steps are been used for data preprocessing such as Data Cleaning, Data Integration, Data Transformation, Data Reduction. The below Fig 2 shows the Data After pre-processing.

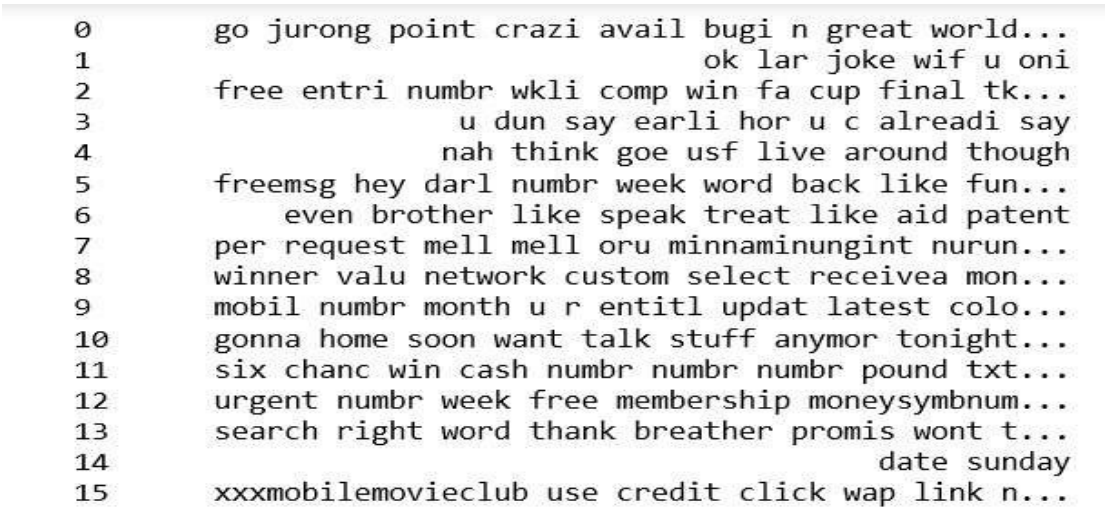


Fig 2

Bag of Words

From the dataset with extracted features we are identify 15 words that are possibly strong indicators for marking a text as spam or non-spam.

From the below Fig shows the Top 15 most commonly used words and also total number of words from the below Fig 3

```
Number of words: 6579
Most common words: [('numbr', 2648), ('u', 1207), ('call', 674), ('go', 456), ('get', 451), ('ur', 391), ('gt', 318), ('lt', 31
6), ('come', 304), ('moneysymbnumbr', 303), ('ok', 293), ('free', 284), ('day', 276), ('know', 275), ('love', 266)]
```

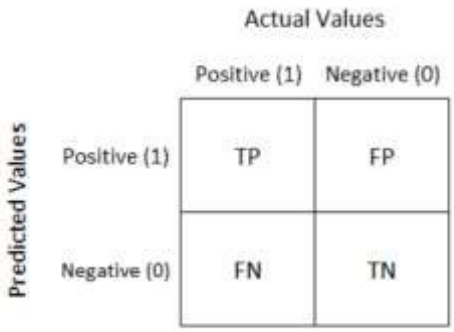
Fig 3

After training the model it was found that the accuracy of K Nearest neighbours was 94.4% , the accuracy of Decision Tree was 97.7% And the accuracy of Naive Bayes was 98.3% as shown in figure 5.3

```
K Nearest Neighbors Accuracy: 94.47236180904522
Decision Tree Accuracy: 97.70279971284997
Naive Bayes Accuracy: 98.34888729361091
```

Confusion Matrix

The performance of the classification of the model are determined by Confusion matrix. Like Precision, Recall, Accuracy and F-Measure are calculated. At the top labels represents the actual class label and the down side predicted class labels are shown.



Confusion matrix

True Positive (TP): The Observed label is positive, and the prediction is positive.

False Negative (FN): The Observed label is positive, but the prediction is negative.

True Negative (TN): The Observed label is negative, but the prediction is negative.

False Positive (FP): The Observed label is negative, but is prediction is positive.

Accuracy: is the number of total correct prediction divided by the total prediction made. And multiplied by 100 in order to get in percentage.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Recall: is the number of total correct prediction divide by the total number of true positive and false positive. Highest Recall indicates the class is correctly identified

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Precision: is the total number of true positive divided by the total number of true positive and false positive.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

F-measure: To calculate F-measure both precision and recall are considered. F-measure will always be nearer to the smaller value of Precision and Recall.

$$F\text{-}measure = \frac{2*Recall*Precision}{Recall + Precision} \quad (4)$$

		predicted	
		ham	spam
actual	ham	1203	12
	spam	11	167

Fig 4

Various combinations of different pre-processing and data transformations were applied with the classifiers on the datasets, and a numerous number of results were observed. The best

results for each classifier were collected. The various evaluation metrics like precision, recall, f-measure, score etc. were calculated

V.CONCLUSION

In this final phase, we will test our classification model on our prepared dataset and also measure the SMS spam detection performance on our dataset. To evaluate the performance of our created classification and make it comparable to current approaches, we use Accuracy to measure the effectiveness of classifiers. The Experiment was performed on various classifier such as decision tree, KNN classifier, Naïve Bayes for SMS spam detection. Naïve Bayes classifier showed the highest accuracy among others classifier.

Future work must rehearse a few ways to deal with raise the part of the feature plot. Including progressively important features like certain limits for the length and learning curves can add to the improvement in results. An application can be used for mobile phones utilizing this techniques in future for protecting our mobile phones from spam message.

REFERENCE

[1] Roy, P. K., Singh, J. P., & Banerjee, S. (2020). Deep learning to filter SMS Spam. *Future Generation Computer Systems*, *102*, 524-533.

[2] Nagwani, N. K. (2017). A Bi-Level Text Classification Approach for SMS Spam Filtering and Identifying Priority Messages. *International Arab Journal of Information Technology (IAJIT)*, *14*(4).

[3] Mokri, M. A. E. S., Hamou, R. M., & Amine, A. (2019). A new bio inspired technique based on octopods for spam filtering. *Applied Intelligence*, *49*(9), 3425-3435.

[4] Shafi'I, M. A., Latiff, M. S. A., Chiroma, H., Osho, O., Abdul-Salaam, G., Abubakar, A. I., & Herawan, T. (2017). A review on mobile SMS spam filtering techniques. *IEEE Access*, *5*, 15650-15666.

[5] Choudhary, N., & Jain, A. K. (2017, March). Towards filtering of SMS spam messages using machine learning based technique. In *International Conference on Advanced Informatics for Computing Research* (pp. 18-30). Springer, Singapore.

[6] Suleiman, D., & Al-Naymat, G. (2017). SMS spam detection using H2O framework. *Procedia computer science*, *113*, 154-161.

[7] Chen, H. (2018, September). Spam message filtering recognition system based on TensorFlow. In *2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE)* (pp. 564-567). IEEE.

[8] Sjarif, N. N. A., Azmi, N. F. M., Chuprat, S., Sarkan, H. M., Yahya, Y., & Sam, S. M. (2019). SMS Spam Message Detection using Term Frequency-Inverse Document Frequency and Random Forest Algorithm. *Procedia Computer Science*, *161*, 509-515.

[9] https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection