International Journal of Engineering & Technology, 7 (3.24) (2018) 536-540

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Machine Learning Algorithms for Spam Detetction in Social Networks

Yenuga Padma[1*],Dr.Y.K.Sundara Krishna[2]

[1]*Research Scholar in CSE , Krishna University,Machilipatnam, A.P*
[2]*Principal & Professor, Department of Computer Science & Engineering, Krishna University,Machilipatnam, A.P*
**Corresponding author E-mail: padmayenuga@gmail.com*

Abstract

Most of the web based social systems like Face book, twitter, other mailing systems and social networks are developed for users to share their information, to interact and engage with the community. Most of the times these social networks will give some troubles to the users by spam messages, threaten messages, hackers and so on.. Many of the researchers worked on this and gave several approaches to detect the spam, hackers and other trouble shoots. In this paper we are discussing some tools to detect the spam messages in social networks. Here we are using RF, SVM, KNN and MLP machine learning algorithms across rapid miner and WEKA. It gives the better results when compared with other tools.

Key words: *Machine learning, social networks, spam detection, WEKA and Rapid miner*.

1. Introduction

From the past 10 years, spam messages have been constant problem around the world. In this communication world, it was increasing in day to day life. Due to the rapid growth of internet and e-mailing system most of the information will be changing in terms of mails and other tools. On the other hand Spam was also increasing rapidly. Spam can originate anywhere from the world. In simple terms spam can be defined as unwanted or unrelated mails or messages it still existed in mailing systems, whenever a user is connected to his mail account apart from the inbox, user will find number of spam messages which were unrelated to him. Most of the times all these spam's were come from advertising companies, unofficial sites, by default clicking on the buttons or tabs, and sometimes due to the exchange of database all these will come to mails. Due to this, several times user will fail to identify the real one and fake one and they miss some important messages also. It became a tool for criminals and hackers to perform illegal activities on internet like stealing the information, selling fake goods, malware distribution etc. The huge amount of spam will be rendered for its manual analysis. Due to the spam bandwidth and storage space will be needed more. To prevent this, most of the researchers developed many methods, tools, and other alternatives. By using spam filters, gate ways, corporate email systems, and end user trains can resist the spams up to some extent.

In this paper we are dealing with the spam detection in social networks like Twitter, which is very popular now-a-days. Users will interact with short messages, is limited to 140 characters. According to the survey of two charts, almost 974 million twitter accounts were existing. The number seems good competitor to face book. Per second almost 6000 tweets are tweeting in twitter means per day almost 500 million tweets and on an average almost 200 billion tweets in a year. For better communication hash tag mention, shorten URL were introduced by Twitter. Most of the accounts were targeted by hackers and spammers. Spammers will do number of malicious things with the account which includes phishing, following too many users, random link connections, keeping the fake profiles, compromising the genuine accounts and malware distribution. So it is important to reduce the spam accounts in social networks.

There are some methods to detect the spam algorithms like Bayesian classifiers algorithm, black and white list algorithms, key word matching algorithms, neural networks, SVM, machine learning tools etc. In this article, will deal with how to identify and restrict the Spam in social networks will be discussed.

2. Literature Review

Many research scholars proposed different methods to identify the spams in mails as well as SMS in mobile applications. Most of the governments gave judgment against to this spam's also. If any sites, illegal advertisers or from any other source if a person receives unwanted mails or messages they will be penalized under the sections called cyber law. To restrict this sort of spam's some scholars gave their contribution in discovering the different types of methods and algorithms.

Kyumin Lee et.al [1] proposed honey pot based technique for uncovered spammers in social networks. In their research work they focused on 2 things. First one is, to install the social honey pots to gather the unreliable spam profiles from social networks to design the proposed work; secondly statistical analysis of spam profiles to create spam classifiers where they can filter the existed and new spammers in the system; they worked on twitter and my space networks by taking the features like pattern posting, friends information and content. By deploying the social honey pots they identified the spammers with low false rate, they developed machine learning based classifiers to identify the unknown hackers or spammers with high accuracy and low false rates.

Ali Shafiguhaski, Navid k Sourati [2] proposed an efficient algorithm to filter the spam by using machine learning tools. They

worked with many mailing systems and other search h engines. They discussed about three algorithms of machine learning to filter the spam from the valid emails with high accuracy and low error rate by using MLP (Multi Layer perceptron), Naïve Bayes Classifier(NBC), and C4.5 decision tree classifier used to train the data whether they were valid mails or spam's.

Hanif, Mohammed Hazim et.al [3] performed an evaluation to detect the spam in large social networks like twitter. In their article by taking the set of features to identify the spammers, they added some extra features for the classifiers. They performance done on four algorithms of machine learning SVM,RF, MLP and KNN across two machine tools rapid miner and WEKA. Taking the 32 features data set RF and MLP on both the learning tools gave 95% .

Scott Clayton [4] proposed a new method to detect the spam by using Azure machine learning. Here Author trained classifiers in Azure to identify whether that message was spam or not. He used 16 bit hash for 65,536 features and selected best of the 1000. Author explored direct word frequency approach to get the accuracy; surprisingly author got 99 % accuracy in his article.

Paras sethi et.al [5] dealed with SMS spam detection and comaping of different machine learning algorithms in their article. In every spam detection Bayesian filters will play a major role. Here authors compared different algorithms on spam detection by taking a public survey in mobile applications. They took two data sets for validation and testing purpose. The results gave different feature classification of spam messages under different algorithms.

Son dinh et.al [6] proposed a soft ware frame work for spam campaign detection, analysation and investigation. The frame work gives law enforcement administrators a platform to perform the investigation on the cyber crimes. By combining the spam mails into campaigns it minimizes the investigation efforts. To handle the huge number of spam mails they kept feature-rich and scalable database. The proposed frame work recognizes spam operations on fly. Adding to this it labels gathers the information and scores the campaigns.

Victor.M.Prieto et.al [7] proposed a content based web spam analyzer and detector in their article. They concentrated on www; means websites. Web spam is the major problem in today's world. This paper deals with study of different types of web spam pages and detects the new elements in it to describe the heuristics capable to detect them. They proposed a new method called SAAD means spam analyzer and detector works based on C4.5 classifier improved by boosting and bagging methods. This one is also very effective in finding spam data sets.

Y.Padma et.al [8] proposed a novel frame work on machine learning to detect the key words to progress the detection work by deploying context exposure approaches. They proposed automatic detector of spam of fake mails or messages in the web sites. When compared with other algorithms like SVM, Bayesian and Naive they got 100 % accuracy in finding the spam's.

Grier et.al [9] proposed a black list method to detect the spammers in twitter accounts. They analyzed click through data was generated which was posted through URL in twitter. Phishing attacks were successfully used in twitter.

3. Methodolgy

In this article, for comparative analysis four classifiers were selected. The algorithms were K-nearest neighbour (KNN), Random forest (RF), Multi layerperceptrons (MLP), Support vector machine (SVM). Most of the researchers used these four for the accurate results. On two working machine tools WEKA and Rapid miner, these four classifiers were trained with 32 features of data set. Here discuss about the four classifiers briefly.

KNN algorithm, which computes the new instance class like its most K-nearest neighbours. It is illustration based learning algorithm having linear computational complexity, it's been used in many applications, when a new instance to be classified, to compute the closest KNN it uses Euclidean distance.

Random forest is set of decision tree algorithms depends on ensemble approach; by using the tree structure the decision tree algorithms will categorize the instances. Test of attributed value will be denoted by node and test results will be denoted by its branches. RF creates classifiers of ensemble by constructing distinct decision trees by using random feature selection and approach of bagging at training level. Decision Tree generates 2 nodes one is class is labeled with leaf node and the other one is feature associated with interior node. All these will be trained [10,11].

Support vector machine algorithm analyzes the data and identifies the patterns by using label samples. This algorithm was developed by Vapnik and others.SVM used for regression and classification tasks; by using hyper plane user can divide the boundary among different classes in the data set. [12]. Hyper plane will separate the classes by enlarging the boundaries between the closes points is called as support vectors.

Multi layerperceptrons is set of feed forward artificial neural network having activation units generally called as artificial neurons and weights. Standard linear perceptrons was modified by MLP by insisting multiple layers like hidden, inputs, and outcome layers to resolve the both non linear and linear classification troubles. MLP maps input data for accurate results. [13] In training level, to adjust the weights MLP used learning algorithm, mostly back propagation. By doing this network obtain adequate knowledge to classify the unknown data.

In Twitter to classify the profiles whether they were related to spam or non spam, programmers developed a crawler used by Twitter REST (Representational State Transfer) API which allows user to retrieve the tweets and other related information. By collecting huge amount of data, crawler was incorporated with black lists which uses the Phish Tank (Anti phishing site) and Google safe browsing APIs. When tweets come through the URL, the crawler query along with Google safe browsing and phish tank it checks the URL to know whether it is real one or fake one. All collected outcomes from every API are in JSON format. When research was conducted around 7000 profiles, almost 2500 were spam and rest is non spam.

Items Description	**Total Number of items**
Tweets	4,35,658
Unique Profiles	43,546
Hash tags	298, 789
URL's	231,638
Features	32
Profiles	7000
Spam	2500
Non Spam	4500

Table 1: By using REST API, collected Twitter data set

Selected Classifier	Evaluated Metrics	Rapid Miner	WEKA	Configured Algorithm
Random Forest	Accuracy	95%	94%	Default
	Rate of Error	4.36%	5.68%	
	Kappa-Statistic	0.903	0.8874	
	RMSE	0.242	0.235	
	MAE	0.158	0.0994	
	Accuracy	92.98%	95.0594%	K=10
	Rate of Error	6.32%	4.458%	

KNN	Kappa-Statistic	0.845	0.8902	
	RMSE	0.218	0.2214	
	MAE	0.100	0.0798	
Support Vector Machine	Accuracy	85.64%	95.01%	Default
	Rate of Error	10.25%	4.028 %	
	Kappa-Statistic	0.682	0.8905	
	RMSE	0.286	0.2013	
	MAE	0.112	0.0398	
MLP Multi Layer Perceptron	Accuracy	95.02%	95.32%	Momentum =0.9 Hidden Layers=1 Seed=100
	Rate of Error	4.02%	4.3267%	
	Kappa-Statistic	0.886	0.889	
	RMSE	0.198	0.219	
	MAE	0.058	0.038	

The features which were applied are already proposed for spam profiles. To improve the classifiers performance while detecting the spam or non spam new features were introduced. During the crawling period, huge data was collected to derive these new features. When it comes for manual analysis, in selected profile a link was established from crawler to home page of twitter. Researchers found that selected user is interacting with his friends and tweeting when important messages or some discussions were raised. Active user will show this type of behavior where as spammers will not do this type of things. Generally they used links in their tweets with hash tags and other mentioned tags. Secondly they never change their profile pictures. Thirdly, to mislead the victims they use pretty or eye catching profile images. Finally spammers use same picture for different profiles.

4. Results and Discustions

To evaluate and compute the performance of selected classifiers on popular machine learning tools Rapid miner version studio basic 7.0.001 and (Waikato Environment for Knowledge Analysis) WEKA 3.6.13 softwares.

Both will be used to implement the different machine learning algorithms. They gave best results of Error rate, MAE (Mean absolute error), Kappa Static and RMSE. The results of this experiment was conducted with 32 features, when random Forest performed on both the softwares RM Gives around 95 % and WEKA gives 94 percentage. WEKA gave best results when SVM, MLP and KNN were performed. MLP gave almost 95% on WEKA software.

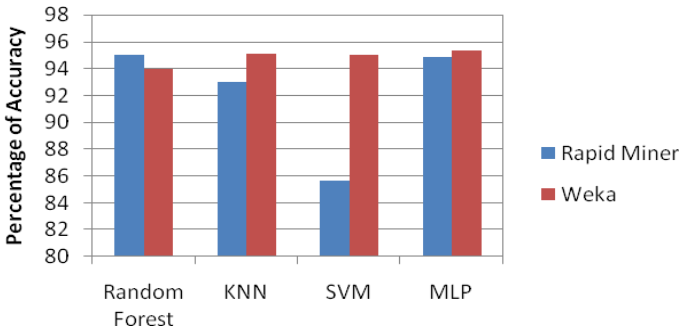

Figure 1: Accuracy rate analysis of Classifiers with RM and WEKA

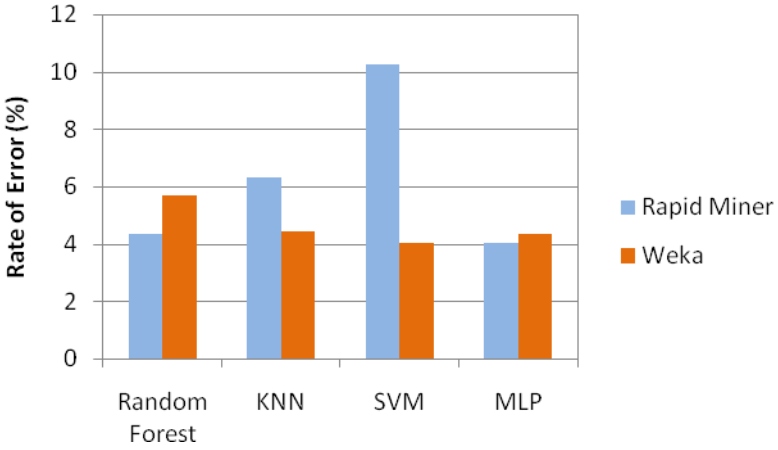

Figure 2: Rate of Error (%) analysis of Classifiers with RM and WEKA

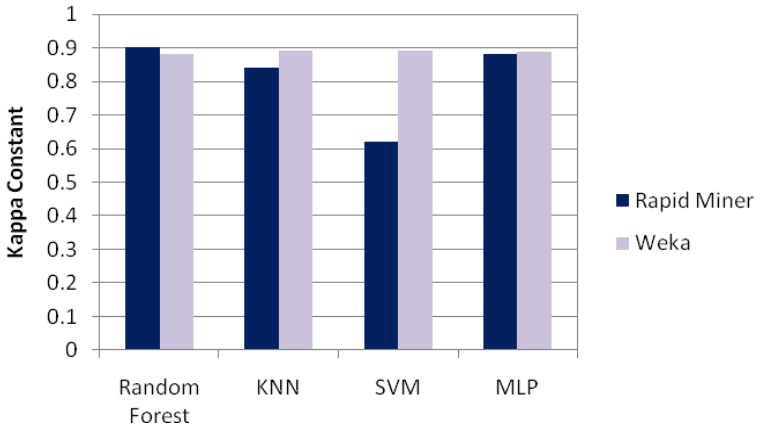

Figure 3: Kappa Constant analysis of Classifiers with RM and WEKA

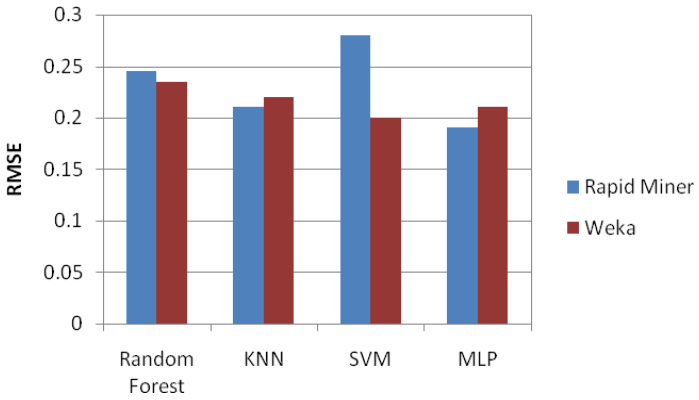

Figure 4: RMSE analysis of Classifiers with RM and WEKA

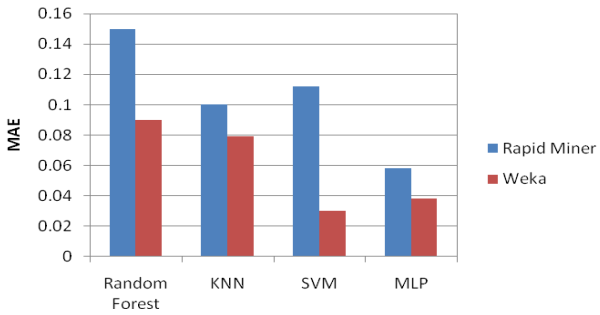

Figure 5: MAE analysis of Classifiers with RM and WEKA

5. Conclusion

This article deals with the spam profiles detection in large social network Twitter. Selected four classifiers like Support vector machine, Random Forest, KNN and Multi Layer Perceptron; and analyzed their performances on two popular learning software's WEKA and Rapid miner. By using the existing features researchers identify the spam profiles in Twitter and also added some new features to get the best results. Accuracy, Rate of Error, Kappa statistic, RMSE and MAE were taken as evaluation metrics to get the better results. Rapid forest gave its best in accuracy on Rapid miner than WEKA. All these findings will be useful to other researchers to develop the new tools for detection spam in social networks.

References

[1] Uncovering social spammers: social honey pots and machine learning by kyumin lee, James caver lee and Steve webb, https://dl.acm.org/citation.cfm?id=1835522

[2] Proposed efficient algorithm to filter spam using machine learning techniques by Ali Shafiguhaski, Navid k Sourati Pacific Science Review A: Natural Science and Engineering Volume 18, Issue 2, July 2016, Pages 145-149

[3] Performance Evaluation of Machine Learning Algorithms for Spam Profile Detection on Twitter Using WEKA and RapidMiner by Hanif, MohamadHazimMd; Adewole, KayodeSakariyah; Anuar, Nor Badrul; Kamsin, Amirrudin Source: Advanced Science Letters, Volume 24, Number 2, February 2018, pp. 1043-1046(4)

[4] Detecting Spam with Azure Machine Learning by Scott Clayton, 12 Feb 2018.

[5] SMS spam detection and comparison of various machine learning algorithms by Paras sethi et.al https://ieeexplore.ieee.org/document/8284445/

[6] Spam campaign detection, analysis, and investigation by son dinh et.al

https://www.sciencedirect.com/science/article/pii/S1742287615000079

[7] SAAD, content based Web Spam Analyzer and Detector by Victor.M. Prieto et.al https://www.sciencedirect.com/science/article/pii/S0164121213001684.

[8] An automated frame work for document spam detection using enhanced context feature matching by Y.Padma et.al in www.ijarcs.info in volume number 9 number 1 jan-feb 2018.

[9] Grier, C., Thomas, K., Paxson, V., & Zhang, M. (2010). @spam: the underground on 140 characters or less. Proceedings of the 17th ACM conference on Computer and communications security, 27-37.

[10] Chu, Z., Gianvecchio, S., Wang, H., &Jajodia, S. (2012). Detecting automation of twitter accounts: Are you a human, bot, or cyborg? IEEE Transactions on Dependable and Secure Computing, 9(6), 811-824. doi:10.1109/TDSC.2012.75

[11] Narudin, F. A., Feizollah, A., Anuar, N. B., &Gani, A. (2014). Evaluation of machine learning classifiers for mobile malware detection. Soft Computing, 1-15.

[12] Smola, A. J., &Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, 14(3), 199-222.

[13] Noriega, L. (2005). Multilayer perceptron tutorial. School of Computing. Staffordshire University