ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 5, May 2019

Twitter Spam Detection on Real Time Data using Machine Learning Algorithms

Sumit S. Garud[1], Piyush Shewale[2], Akshay Khedkar[3], D.D.Sapkal[4]

BE Students, Department of Computer Engineering, Pune Vidyarthi Griha's COET, Pune, India[1,2,3]

Professor, Department of Computer Engineering , Pune Vidyarthi Griha's COET, Pune, India[4]

ABSTRACT: The online social networks are a very large growth in the world today, but the attacks are more common, including one of the attacks is the attack of Twitter in this spammer spreading several malicious tweets that can take the form of links or hash tags in the website and online services, which are too harmful for real users. To prevent these attacks, training tweets are added and, moreover, these problems are solved by extracting 12 lightweight functions, like the age of the account, no. of followers, no. to follow, no. of tweets, no. of re-tweets, etc. For the transmission of spam detection from tweets, the discretization of a function is important for the performance of spam detection. There is a great truth in the system that includes a total of 600 public tweets based on the URL-based security tool. Spam detection primarily creates the classification model that includes binary classification and can also be solved using the automatic learning algorithm. Machine learning algorithms such as the Naïve Bayesian classifier or the vector support machine classifier have informed the behavior of the models. The system reported the impact of data-related factors, such as the relationship between spam and non-spam, the size of training data and data sampling, and detection performance. The implemented system function is the detection of simple and variable tweets of spam over time. The system shows how spam detection is a major challenge and bridges the gap between performance appraisals and focuses primarily on data, features and patterns to identify the real user and inform the user of spam when providing the valuable response binary. The contribution work is to detect the tweets of spam in real time, because the new tweets come in the form of sequences and use the updated training data set.

KEYWORDS: Distributed Computing, Spam Detection, Scalability, Twitter, Mining.

I. INTRODUCTION

Online social networking sites like Twitter, Facebook, Instagram and some online social networking companies have become extremely popular in recent years. People spend a lot of time in OSN making friends with people they are familiar with or interested in. Twitter, founded in 2006, has become one of the most popular microblogging service sites. Around 200 million users create around 400 million new tweets a day for spam growth. Twitter spam, known as unsolicited tweets containing malicious links that the non-stop victims to external sites containing the spread of malware, spreading malicious links, etc., hit not only more legitimate users, but also the whole platform Consider the example because during the election of the Australian Prime Minister in 2013, a notice confirming that his Twitter account had been hacked. Many of his followers have received direct spam messages containing malicious links.

The ability to order useful information is essential for the academic and industrial world to discover hidden ideas and predict trends on Twitter. However, spam generates a lot of noise on Twitter. To detect spam automatically, researchers applied machine learning algorithms to make spam detection a classification problem. Ordering a tweet broadcast instead of a Twitter user as spam or non-spam is more realistic in the real world.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 5, May 2019

II. RELATED WORK

Nathan Aston, Jacob Liddle and Wei Hu*[1] describe the "Twitter Sentiment in Data Streams with Perceptron" in this system the implementation feature reduction we were able to make our Perceptron and Voted Perceptron algorithms more viable in a stream environment. In this paper, develop methods by which twitter sentiment can be determined both quickly and accurately on such a large scale.

Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro [2] describe the "Aiding the detection of fake accounts in large scale social online services".in this paper, SybilRank, an effective and efficient fake account inference scheme, which allows OSNs to rank accounts according to their perceived likelihood of being fake. It works on the extracted knowledge from the network so it detects, verify and remove the fake accounts.

G. Stringhini, C. Kruegel, and G. Vigna [3] describe the "Detecting spammers on social networks" in this paper, Help to detect spam Profiles even when they do not contact a honey-profile.The irregular behavior of user profile is detected and based on that the profile is developed to identify the spammer.

J. Song, S. Lee, and J. Kim [4] describe the "Spam filtering in Twitter using sender receiver relationship" in this paper a spam filtering method for social networks using relation information between users and System use distance and connectivity as the features which are hard to manipulate by spammers and effective to classify spammers.

K. Lee, J. Caverlee, and S. Webb [5] describe the "Uncovering social spammers: social honeypots and machine learning" in this System analyzes how spammers who target social networking sites operate to collect the data about spamming activity, system created a large set of "honey-profiles" on three large social networking sites.

K. Thomas, C. Grier, D. Song, and V. Paxson [6] describe the "Suspended accounts in retrospect: An analysis of Twitter spam" in this paper the behaviors of spammers on Twitter by analyzing the tweets sent by suspended users in retrospect. An emerging spam-as-a-service market that includes reputable and not-so-reputable affiliate programs, ad-based shorteners, and Twitter account sellers.

K.Thomas, C.Grier, J.Ma, V.Paxson, and D.Song [7] describe the "Design and evaluation of a real-time URL spam filtering" in this paper, service Monarch is a real-time system for filtering scam, phishing, and malware URLs as they are submitted to web services.Monarch's architecture generalizes to many web services being targeted by URL spam, accurate classification hinges on having an intimate understanding of the Spam campaigns abusing a service.

X. Jin, C. X. Lin, J. Luo, and J. Han [8] describe the "Social spam guard: A data mining based spam detection system for social media networks" in this paper ,Automatically harvesting spam activities in social network by monitoring social sensors with popular user bases.Introducing both image and text content features and social network features to indicate spam activities. Integrating with our GAD clustering algorithm to handle large scale data. Introducing a scalable active learning approach to identify existing spams with limited human efforts, and Perform online active learning to detect spams in real-time.

S. Ghosh et al [9] describe the "Understanding and combating link farming in the Twitter social network" in this paper Search engines rank websites/webpages based on graph metrics such as PageRank High in-degree helps to get high PageRank. Link farming in Twitter Spammers follow other users and attempt to get them to follow back.

H. Costa, F. Benevenuto, and L. H. C. Merschmann **[10]** describe the "Detecting tip spam in location-based social networks" in this paper identifying tip spam on a popular Brazilian LBSN system, namely Apontador.Based on a labelled collection of tips provided by Apontador as well as crawled information about users and locations, we

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 5, May 2019

identified a number of attributes able to distinguish spam from non-spam tips

III. PROPOSED SYSTEM

In proposed system, the process of Twitter spam detection by using machine learning algorithms. Before classification, a classifier that contains the knowledge structure should be trained with the prelabeled tweets. After the classification model gains the knowledge structure of the training data, it can be used to predict a new incoming tweet. The whole process consists of two steps: learning and classifying. Features of tweets will be extracted and formatted as a vector. The class labels i.e. spam and non-spam could be get via some other approaches. Features and class label will be combined as one instance for training. One training tweet can then be represented by a pair containing one feature vector, which represents a tweet, and the expected result, and the training set is the vector. The training set is the input of machine learning algorithm, the classification model will be built after training process. In the classifying process, timely captured tweets will be labelled by the trained classification model.

Advantages Of proposed system:

1. The system implements a method that will use the ML mechanism to detect if the post is spam or not.
2. Implementation of system can also be hosted online for use and data will be archived and retrieved from the server.
3. The user with the maximum amount of spam can be blocked by the system.

System Architecture:

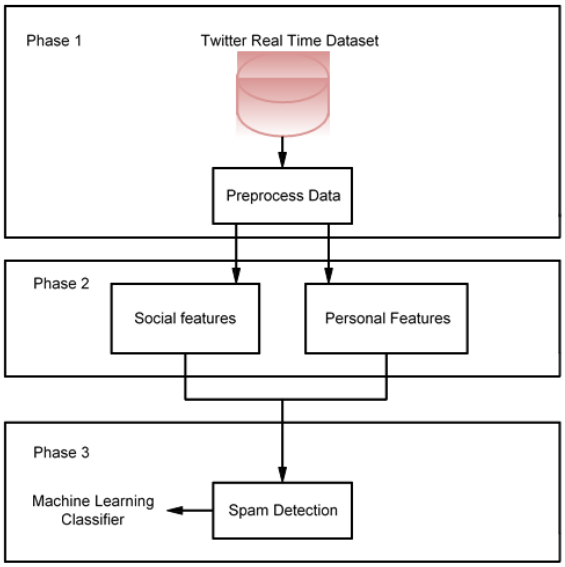

Fig 1.System Architecture

Algorithms:

Naive Bayes

- Given training dataset D which consists of documents belonging to different class say Class A and Class B

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 5, May 2019

- Calculate the prior probability of class A=number of objects of class A/total number of objects
  - Calculate the prior probability of class B=number of objects of class B/total number of objects
- Find NI, the total no of frequency of each class
  - Na=the total no of frequency of class A
  - Nb=the total no of frequency of class B
- Find conditional probability of keyword occurrence given a class:
  - P (value 1/Class A) =count/ni (A)
  - P (value 1/Class B) =count/ni (B)
  - P (value 2/Class A) =count/ni (A)
  - P (value 2/Class B) =count/ni (B)
  - ……..
  - ……..
  - ……..
  - P (value n/Class B) =count/ni (B)
- Avoid zero frequency problems by applying uniform distribution
- Classify Document C based on the probability p(C/W)
- Find P (A/W) =P (A)*P (value 1/Class A)* P (value 2/Class A)……. P(value n /Class A)
- Find P (B/W) =P (B)*P (value 1/Class B)* P(value 2/Class B)……. P(value n /Class B)
- Assign document to class that has higher probability.

IV. RESULT AND DISCUSSION

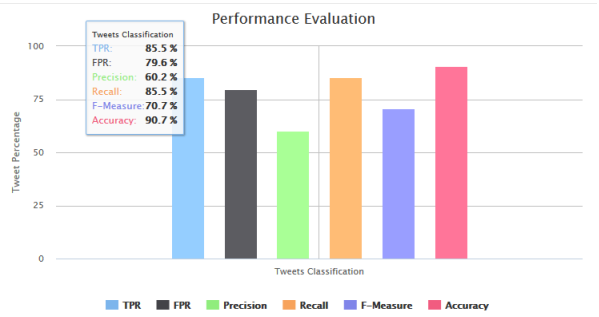

Fig 2. Accuracy Graph

TPR	FPR	Precision	Recall	F-Measure	Accuracy
77%	72%	62%	79%	68%	89.5%

Table 1.Metrics Table

V.CONCLUSION

In this Project, System found that classifiers ability to detect Twitter spam reduced when in a near real-world scenario since the imbalanced data brings bias. System also identified that Feature discretization was an important pre-process to ML-based spam detection. Second, increasing training data only cannot bring more benefits to detect Twitter spam after

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 5, May 2019

a certain number of training samples. System should try to bring more discriminative features or better model to further improve spam detection rate.

REFERENCES

[1] Nathan Aston, Jacob Liddle and Wei Hu*, "Twitter Sentiment in Data Streams with Perceptron," in Journal of Computer and Communications, 2014, Vol-2 No-11.

[2] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," in Proc. Symp. Netw. Syst. Des. Implement. (NSDI), 2012, pp. 197–210.
[3] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in Proc. 26th Annu. Comput. Sec. Appl. Conf., 2010, pp. 1–9.
[4] J. Song, S. Lee, and J. Kim, "Spam filtering in Twitter using sender receiver relationship," in Proc. 14th Int. Conf. Recent Adv. Intrusion Detection, 2011, pp. 301–317.
[5] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: social honeypots + machine learning," in Proc. 33rd Int. ACM SIGIR Conf. Res.Develop. Inf. Retrieval, 2010, pp. 435–442.
[6] K. Thomas, C. Grier, D. Song, and V. Paxson, "Suspended accounts in retrospect: An analysis of Twitter spam," in Proc. ACM SIGCOMM Conf. Internet Meas., 2011, pp. 243–258.
[7] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time URL spam filtering service," in Proc. IEEE Symp. Sec. Privacy, 2011, pp. 447–462.
[8] X. Jin, C. X. Lin, J. Luo, and J. Han, "Socialspamguard: A data mining based spam detection system for social media networks," PVLDB, vol. 4, no. 12, pp. 1458–1461, 2011.
[9] S. Ghosh et al., "Understanding and combating link farming in the Twitter social network," in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 61–70.
[10] H. Costa, F. Benevenuto, and L. H. C. Merschmann, "Detecting tip spam in location-based social networks," in Proc. 28th Annu. ACM Symp. Appl. Comput., 2013, pp. 724–729.