IJARST
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

A FRAME WORK FOR REAL TIME SPAM DETECTIONIN TWITTER

[1]Dr. P Senthil Kumar, [2]Gali Anusha, [3]Mandava Aruna,[4]Gutha Venkata Sai Pavan, [5]Pallapati venkataramarao
[1,2,3,4,5]Assistant professors, Department of CSE in Narasaraopet Institute Of Technology

ABSTRACT

The popularity of social media sharing the information and communication each and every one. Twitter is one such popular network and to the little message communication (called tweets) has entice a more large number of users in network. Spammer is tweets fake news and illegal news both as advertisements, scams and help perpetrate phishing attacks or the spread of malware through the embedded to the users. To find these platforms without problems reachable to trap users in malicious activities by posting spam messages. In this work, to taken Twitter platform and performed spam tweets detection. Some of tools using tools can block malicious links of user they cannot protect the user in real-time as early as possible. Thus, industries and researchers have applied different approaches to make spam free social network platform. To comprehensive solution that can consolidate tweet's text information along with the user based features. To solve this issue, to propose to takes the user and tweet based features along with the tweet text feature to classify the tweets. The benefit of using tweet text feature is that to identify the spam tweets even if the spammer creates a new account which was not possible only with the user and tweet based features. To evaluated by solving the problems by different machine learning algorithms namely - Support Vector Machine, Neural Network, Random Forest and Gradient Boosting.The increased popularity of online social net works, spammers find these platforms easily accessible to trap users in malicious activities by posting spam messages. In this work, we have taken Twitter platform and performed spam tweets detection. To stop spammers, Google SafeBrowsing and Twitter's BotMaker tools detect and block spam tweets. These tools can block malicious links, h owever they cannot protect the user in real-time as early as possible. Thus, industries and researchers have applied different approaches to make spam free social network platform. Some of them are only based onuser-based features while others are based on tweet based features only. However, there is no comprehensive solution that can consolidate tweet's text information along with theuser based features. To solve this issue, we propose a framework which takes the user and tweet based features along with the tweet text feature to classify the tweets. The benefit of using tweet text feature is that we can identify the spam tweets even if the spammer creates a new account which was not possible only with the user and tweet based features. We have evaluated our solution with four different machine learning algorithms namely - Support Vector Machine, Neural Network, Random Forest and Gradient Boosting. With Neural Network.

1. INTRODUCTION

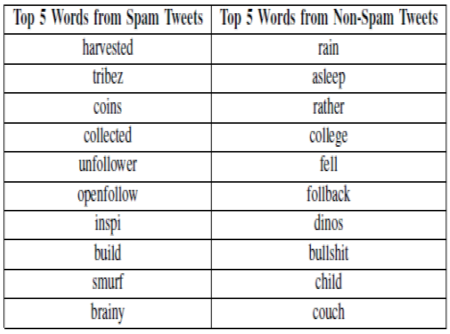
In the past few years, online social networks like Facebook and Twitter have become increasingly prevailing platforms which are integral part of people daily life. People spend lot of time in microblogging websites to post their messages, share their ideas and make friends around the world. Due to this growing trend, these platforms attract a large number of users as well as spammers to broadcast their messages to the world. Twitter is rated as the most popular social network among teenagers. Twitter also invites more unsolicited activities on this platform. Nowadays, 200 million users generate 400 million new tweets per day. This rapid expansion of Twitter platform influences more number of spammers to generate spam tweets which contain malicious links that direct a user to external sites containing malware downloads, phishing, drug sales, or scams. These types of attacks not only interfere with the user experience but also damage the whole internet which may also possibly cause temporary shutdown of internet services all over the world. This performed spam tweet detection based on deep learning. They used word vector to train their model, but they have not explored user or tweet based features to address the problem. On the other side, Chao Chen [1] used lightweight features (user's and tweet's specific feature) that are suitable for real-time spam tweet detection. As Twitter has increased their character limit to 280 characters, it is essential to scrutinize the tweet's text along with the user-specific features. Despite many existing solutions, there are very few comprehend sive solutions that can be used for blocking spam tweets in real-time. In this paper, we give a framework based on different machine learning approach that deals with various problems including accuracy shortage, time lag(BotMaker) and high processing time to handle thousands of tweets in 1 sec. Firstly, we have collected 400,000 tweets from HSpam14 [9] dataset. Then we further characterize the 150,000 spam tweets and 250,000 non-spam tweets. We also derived some lightweight features along with the Top-30 words that are providing high est information gain from Bag-of-Words model. This approach has been detailed in section III. This technique is proficient for spam detection in real-time. We also performed various experiments for detecting Twitter spam using our processed dataset. As a consequence, researchers as well as Twitter came up with spam detection solutions to make spam-free online social network platform. Twitter built BotMaker [6] to fight spam on Twitter platform. They have seen a 40% reduction in critical spam metrics launching BotMaker. But one of the weak aspects of BotMaker is that it fails to protect a victim from new spam, i.e. it efficient tool for real-time spam tweets detectionK. Thomas had observed that 90% users might visit a new spam link before it gets blocked the blacklist.

2. Literature Survey

C. Chen, J. Zhang, X. Chen, Y. Xiang, and W. Zhou, "6 million spam tweets: A large ground truth for timely twitter spam detection," in 2015 Twitter has changed the way of communication and getting news for people's daily life in recent years. Meanwhile, due to the popularity of Twitter, it also becomes a main target for spamming activities. In order to stop spammers, Twitter is using Google SafeBrowsing to detect and block spam links. Despite that blacklists can block malicious URLs embedded in tweets, their lagging time hinders the ability to protect users in real-time. Thus, researchers begin to apply different machine learning algorithms to detect Twitter spam. However, there is no comprehensive evaluation on each algorithms' performance for real-time Twitter spam detection due to the lack of large groundtruth. To carry out a thorough evaluation, we collected a large dataset of over 600 million public tweets. We further labelled around 6.5 million spam tweets and extracted 12 light-weight features, which can be used for online detection. In addition, we have conducted a number of experiments on six machine learning algorithms under various conditions to better understand their effectiveness and weakness for timely

Twitter spam detection. We will make our labelled dataset for researchers who are interested in validating or extending our work. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in In Collaboration, Electronic messaging, AntiAbuse and Spam Conference With millions of users tweeting around the world, real time search systems and different types of mining tools are emerging to allow people tracking the repercussion of events and news on Twitter. However, although appealing as mechanisms to ease the spread of news and allow users to discuss events and post their status, these services open opportunities for new forms of spam. Trending topics, the most talked about items on Twitter at a given point in time, have been seen as an opportunity to generate traffic and revenue. Spammers post tweets containing typical words of a trending topic and URLs, usually obfuscated by URL shorteners, that lead users to completely unrelated websites. This kind of spam can contribute to de-value real time search services unless mechanisms to fight and stop spammers can be found.

3. SYSYTEM ANALYSIS



Top 5 Words from Spam Tweets	Top 5 Words from Non-Spam Tweets
harvested	rain
tribez	asleep
coins	rather
collected	college
unfollower	fell
openfollow	follback
inspi	dinos
build	bullshit
smurf	child
brainy	couch

EXISTING SYSTEM:

The popularity of Twitter attracts more and more spammers. Spammers drive unnecessary tweets to twitter users to promote websites or services, which are harmful time to normal users. In order to stop spammers, researchers have proposed a number of mechanisms. The focus of recent works is on the application of machine learning techniques into Twitter spam detection is a problem throughout the Internet, and Twitter is not immune. In addition, Twitter spam is much more successful compared to email spam.

DISADVANTAGES:

- Personal Information can lead to potentially dangerous situations
- Unable to examine personally
- Due to that you loss.

PROPOSED SYSTEM:

This Project proposes consolidating the data from the given twitter dataset like account age, number of followers , no of following, user favorites, urls, tweets, retweets, chars, numbers, ascii... Then the dataset will go under preprocessing like removal of null records and columns. Filling the null vaues with relevant values. Removal of Stopwords, extra spaces, punctuations and other undesired symbols. Finally from the Tweet Texts and gained information from datasets we will classify the tweet messages as SPAM or HAM.

step - 1. Consolidating the data from the twitter dataset like

step-2

Prepocess the data like Stop word Removal, Extra space Removal, Punctuations Removal,etc...
Data Cleaning like: removing null records (record with most null values) or null columnsFilling the null values with relevant valuesFind the Frequency of the words , and Pick the top n words:

4. INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- ➢ What data should be given as input?
- ➢ How the data should be arranged or coded?
- ➢ The dialog to guide the operating personnel in providing input.
- ➢ Methods for preparing input validations and steps to follow when error occur.

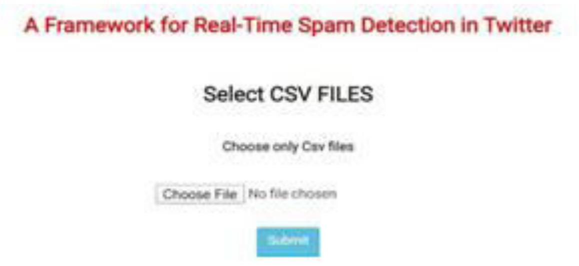
OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

5. Results:

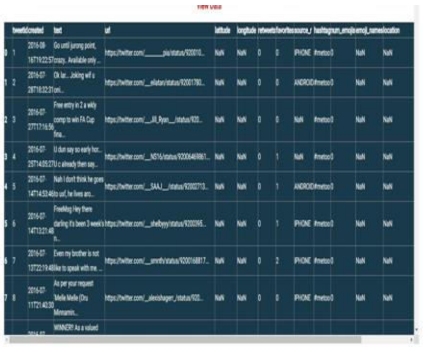


A Framework for Real-Time Spam Detection in Twitter

Select CSV FILES

Choose only Csv files

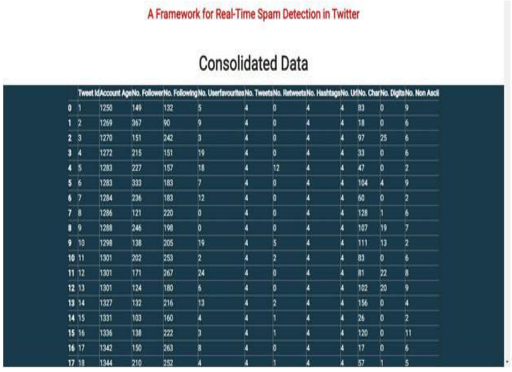
Fig 5.1 Successful CSV file upload



View Data

	tweetid	created	text	url	latitude	longitude	retweets	favorites	source_	hashtag	num_emoji	emoji_names	location
0	1	2016-08-16T19:22:57	Go until jurong point, crazy.. Available only ...	https://twitter.com/_______pia/status/920010...	NaN	NaN	0	0	IPHONE	#metoo	0	NaN	NaN
1	2	2016-07-28T18:32:31	Ok lar... Joking wif u oni...	https://twitter.com/__eliatan/status/92001780...	NaN	NaN	0	0	ANDROID	#metoo	0	NaN	NaN
2	3	2016-07-27T17:16:56	Free entry in 2 a wkly comp to win FA Cup fina...	https://twitter.com/__Jill_Ryan__/status/920...	NaN	NaN	0	0	NaN	#metoo	0	NaN	NaN
3	4	2016-07-25T14:05:27	U dun say so early hor... U c already then say...	https://twitter.com/__NS16/status/92006469861...	NaN	NaN	0	1	NaN	#metoo	0	NaN	NaN
4	5	2016-07-14T14:53:46	Nah I dont think he goes to usf, he lives aro...	https://twitter.com/__SAAJ__/status/92002713...	NaN	NaN	0	1	ANDROID	#metoo	0	NaN	NaN
5	6	2016-07-14T13:21:48	FreeMsg Hey there darling it's been 3 week's n...	https://twitter.com/__shelbyyy/status/9200395...	NaN	NaN	0	1	IPHONE	#metoo	0	NaN	NaN
6	7	2016-07-13T22:19:48	Even my brother is not like to speak with me. ...	https://twitter.com/__smnth/status/9200168817...	NaN	NaN	0	2	IPHONE	#metoo	0	NaN	NaN
7	8	2016-07-11T21:40:30	As per your request 'Melle Melle (Oru Minnamin...	https://twitter.com/__alexishagen_/status/920...	NaN	NaN	0	0	IPHONE	#metoo	0	NaN	NaN
		2016-07	WINNER!! As a valued										

Fig 5.2 Perproceesed data

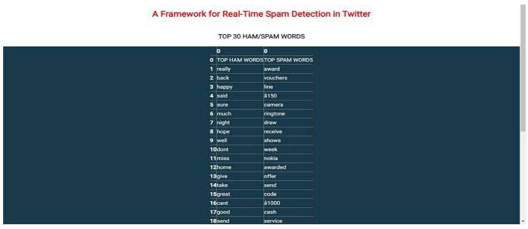


A Framework for Real-Time Spam Detection in Twitter

Consolidated Data

	Tweet Id	Account Age	No. Follower	No. Following	No. Userfavourites	No. Tweets	No. Retweets	No. Hashtags	No. Url	No. Char	No. Digits	No. Non Ascii
0	1	1250	149	132	5	4	0	4	4	83	0	9
1	2	1269	367	90	9	4	0	4	4	18	0	6
2	3	1270	151	242	3	4	0	4	4	97	25	6
3	4	1272	215	151	19	4	0	4	4	33	0	6
4	5	1283	227	157	18	4	12	4	4	47	0	2
5	6	1283	333	183	7	4	0	4	4	104	4	9
6	7	1284	236	183	12	4	0	4	4	60	0	2
7	8	1286	121	220	0	4	0	4	4	128	1	6
8	9	1288	246	198	0	4	0	4	4	107	19	7
9	10	1298	138	205	19	4	5	4	4	111	13	2
10	11	1301	202	253	2	4	2	4	4	83	0	6
11	12	1301	171	267	24	4	0	4	4	81	22	8
12	13	1301	124	180	6	4	0	4	4	102	20	9
13	14	1327	132	216	13	4	2	4	4	156	0	4
14	15	1331	103	160	4	4	1	4	4	26	0	2
15	16	1336	138	222	3	4	1	4	4	120	0	11
16	17	1342	150	263	8	4	0	4	4	17	0	6
17	18	1344	210	252	4	4	1	4	4	57	1	5

Fig 5.3 Top Ham/spam words



A Framework for Real-Time Spam Detection in Twitter

TOP 30 HAM/SPAM WORDS

	0	0
0	TOP HAM WORDS	TOP SPAM WORDS
1	really	award
2	back	vouchers
3	happy	line
4	said	â150
5	sure	camera
6	much	ringtone
7	night	draw
8	hope	receive
9	well	shows
10	dont	week
11	miss	nokia
12	home	awarded
13	give	offer
14	take	send
15	great	code
16	cant	â1000
17	good	cash
18	send	service

Fig 5.4 Model Performance of SVM with Kernal

6. CONCLUSION:

In this project, we present a novel framework for real-time spam detection in Twitter. We collected a large number of 400,000 public tweets. Based on tweet's text we extract top30wordswhich areableto givethehighest informationgain in order to classify the tweets. We have also tested our approach with real-time tweet detection that has outperformed existing approach by 18%. As Twitter API is available to all users, spammers may change their behavior overthetime.

REFERENCES:

1 C. Chen, J. Zhang, X. Chen, Y. Xiang, and W. Zhou, "6 million spam tweets: A large ground truth fortimely twitter spam detection," in 2015.

[2]A. Greig, "Twitter Overtakes Facebook as the Most Popular Social Network for Teens, According toStudy, Daily Mail, accessed on Aug. 1, 2015,Twitter-overtakes-Facebook-popular-socialnetwork-teens-according-study.

H. Tsukayama, "Twitter turns 7: Users send over 400 million tweets per day," Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in InCollaboration, Electronic messaging, AntiAbuse and Spam Conference

C. Pash., "The lure of Naked Hollywood Star Photos Sent the Internet into Meltdown in New Zealand, Bus. Insider, accessed