IDENTIFYING TWITTER BOTS BASED ON TRENDING HASHTAGS USING MACHINE LEARNING ALGORITHMS

Jegan Amarnath.J[1],Nivetha.K.J[2],Pavithra.J[3],Geetha.P[4] [1]Associate Professor,Department Of Computer Science and Engineering
[2,3,4]Department Of Computer Science and Engineering ,
Sri Sairam Engineering College

ABSTRACT:

A bot is a piece of software that completes automated task over the internet .On social media the prevalence of bot is ubiquitous. Nearly 48 million twitter accounts are automated using bots.
Detecting bots is necessary in order to identify bad actors in “Twitter verse” and protect genuine users from misinformation and malicious intents. According to a study released by Pew Research Center, these bots contribute to approximately 66% of total tweets in twitter.Bots are becoming smarter, they mimic humans to avoid being detected and suspended, and increase throughput by creating many accounts.The idea is to develop a bot detection algorithm to identify twitter bot accounts by using attributes like followers _ count, no of tweets or likes, status _ count etc. Thus evaluate these features using k-nearest neighbour algorithm and Random Forest. Finally the result of two models are combined using suitable ensemble method to produce more accurate solution. The Test dataset is prepared specifically for any hashtag. The dataset consists details of the accounts which uses that hashtag. This dataset is cleaned and used as a test dataset.

Keywords: Twitter Bot detection, Hashtag, KNN, Random Forest, Evaluation module.

INTRODUCTION:

Twitter is a social media network, it is most commonly used network among social media platforms. It is launched in the year 2006 as a microblogging site [1]. Communicating via tweets, which are limited in size to only 280 characters, users relay messages to each other. These messages can be in the forms of tweeting, authoring messages; replying, responding to another person’s message; and direct messaging, tweeting a message to another user that is not available for view to the public. User accounts converse with each other by tagging each other with the “@” symbol preceding the target account’s name. Additionally, users have the ability to interact with other accounts on specific topics by using the hashtag symbol “#”. Any tweet containing the hashtag symbol is grouped on a timeline of all tweets that contain that same hashtag. A bot is an automated software application. Twitter bots is a bot that operate from twitter users. Many types of twitter bots are available such as to increase the number of followers, some bots used to promote specific content. It has been determined up to 66% of the activity in twitter is carried out by bots. They are found to be pervasive on Social media. Detecting the bot accounts on Twitter is very difficult and problematic one. At the same time it is necessary to spot out the bots in Twitter for information exchanging system and to protect the users from pernicious intent & misinformation. In this paper, we present supervised and unsupervised machine learning algorithm to identify the twitter bot accounts and to find the intent of the bots. Each of these algorithm is performed independently & the best accuracy result is taken for discovering the bots.

OBJECTIVE:

The objective of our project is to detect the Bots to protect the genuine user in Twitter

- To obtain the account details associated with trending hashtags
- Datasets are collected from Twitter API.
- Perform Machine learning algorithm to the datasets and compare the results.
- To predict whether the account is bot or not.

RELATED WORK:

In 2011, Indiana University team renamed Bot-or-Not to Botometer.They increased their features from 1,000 to 1,150 account related sets [2]. They used Random Forests, AdaBoost, Logistic Regression and Decision Tree classifiers for comparison to predict the final results. In 2011, Texas A&M team became the first one that used honey pots to detect bots [2]. This honey pots bots are used to generate content which is nonsensical, designed only to attract other bots. Their team bots attracted thousands of bots, and generated a labelled data set that has been used on many later research efforts. This method was repeated by others to create similar data sets in other parts of the world. In 2015, Cresci et al. [3] proposed a model based on supervised Machine learning describing their features to the identity of an account only, to detect bots on Social Media Platforms. On the same way.Gupta et al. [4] suggested that behaviour, such as the time of the day and frequency of messages, provides enough information to detect bots successfully through supervised machine learning models. These models require a label included in the corpus to predict the expected outcome [4].In 2015, Yang et al detected spam bots, where this paper survey attempt to mimic human behaviour, aim primarily to spread malware and unwelcome advertising. Later discovered, these classic feature sets lacks to detect bots and suggest relationship based technique [5]. In the same year, A.Wang detected bots in Online networking sites. These bots serve a variety of purposes ranging from simple tasks such as following a user to more complex tasks like engaging in discussion with other users. Social bots are a type of bot that interacts with users and whose purpose is to generate content that promotes a particular viewpoint This paper consider accounts independent of each other(i.e. per-user detection), and are mostly supervised. New classes may need to be introduced in the taxonomy after new defense mechanisms are proposed. Chu et al. classifies Twitter accounts as human, bot, or cyborg accounts [6]. The distinction between these three classifiers is the level of automation placed on the account. An account that Chu et al. classified as human had no activity that is automated. An account where all of its activity is automated is considered a bot. An account that is a mix of automated and non-automated tweets is considered a cyborg. An account that is classified as a cyborg can be run two different ways. The account could be run in a way that would be classified as a human, but also have some automated messages. An account that is classified as a cyborg could also be automated for all of its activity, but it’s controller may sometimes send other, nonscheduled tweets. Tacchini et al. for Fake News detection. On Facebook as well as by Lee et al. for bot detection on Twitter is the use of community-related features, which take advantage of the fact that bots are often created at the same time. To increase the verisimilitude of these accounts, these bots will also follow each other and like each other’s’ posts. The main drawback of this approach is that gathering follower information on Twitter is time-consuming, as the information is not immediately available in the API .In 2016, Nikan Chavoshi, Hossein

Hamooni, Abdullah Mueen proposed a Debot to detect Twitter Bots using warped correlation. This is unsupervised method to detect bots in social media. Develop a novel lag-sensitive hashing technique to quickly group correlated users based on their warping correlations. This allows us to cross-match millions of activity series under time warping, a self-governing framework since no restorative move is made for the issue.

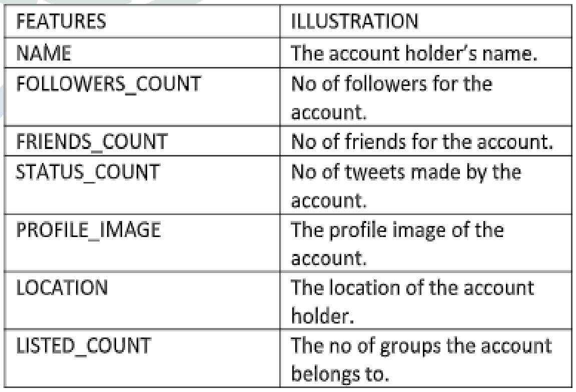
Ebrahimi et al., compared a one-class support vector machine model to a Naïve Bayes machine learning model and showed how the one-class SVM outperforms the binary classification model when one of the classes is the minority. The norm is to train a one-class SVM on the minority class. In SMPs it is not practical to mine the minority class consisting of fake accounts or ben certain that an account is indeed deceptive. Semi-supervised machine learning models require a clear boundary between classes.

PROPOSED WORK:

In the proposed task, In Twitter, information can be gained about a user from their personal account information, tweets, likes, retweets, and direct messages. Users' direct messages are not accessible for privacy reasons. To identify bots, we use supervised and unsupervised machine learning algorithm namely: K-nearest neighbour and Random Forest. These algorithms are used based on their previous accuracy results. On social media platforms, most users place some personal information indicate their individuality. In unsupervised machine learning like K-nearest neighbor the data is Unlabelled, and data are grouped based on similarity framework by means of Clustering works are made to detect bots as they usually share similar characteristics and has the same purpose. In supervised machine learning like Random forest it requires a dataset of features with a label classifying each row or outcome.

INPUT

The features can be the attributes which found via twitter API's which describes the information about twitter account. Features used by machine learning models are mostly referred to as ‘‘engineered features’’ as they are a combination of attributes and engineered features. They can be given in the following table:



FEATURES	ILLUSTRATION
NAME	The account holder's name.
FOLLOWERS_COUNT	No of followers for the account.
FRIENDS_COUNT	No of friends for the account.
STATUS_COUNT	No of tweets made by the account.
PROFILE_IMAGE	The profile image of the account.
LOCATION	The location of the account holder.
LISTED_COUNT	The no of groups the account belongs to.

DATASETS AND FEATURES:

In this the given datasets and features, we train the datasets with the given algorithm.

We determine the most frequently used words i.e. in trending, by searching them in form of the hashtags used in tweets which is a time-consuming method rather than checking the accounts separately. This increases the efficiency of the bot detector making our proposed model better than other models. And we finally test the datasets associated with the hashtag accounts with training datasets and predict whether it is bot or not. Moreover, in the past research of detecting the bots models like SVM & Linear regression are not a good option. Typical features used in these methods need a long duration of activities which makes the detection process useless. Thereby, bots are becoming smarter. They mimic humans to avoid being detected and suspended and increase throughput by creating many accounts. From the above methods like K- nearest neighbor and Random Forest, we train each of these models with the datasets separately to find out the results. And finally, we will compare these results with each other to find the one with the highest accuracy. All models should be created and fitted using algorithms implemented in the “sci-kit-- learn” python library. List of modules for collecting datasets using Twitter’s rest API Tweepy: Feature selection, partitioning the data, using the machine learning models to evaluate the result. The data consists of different attributes related to corresponding twitter accounts. The trending hashtags are collected using the Twitter rest API.The accounts which correspond to the trending hashtags are searched using the search API feature, and the features of those accounts are evaluated. Types of Attributes used are: Integers: Id, Id String, Friends, Followers Count, Listed Count, Favourites and Statuses Count.

Strings: Screen_Name, Location, Description, URL, Name and Language.

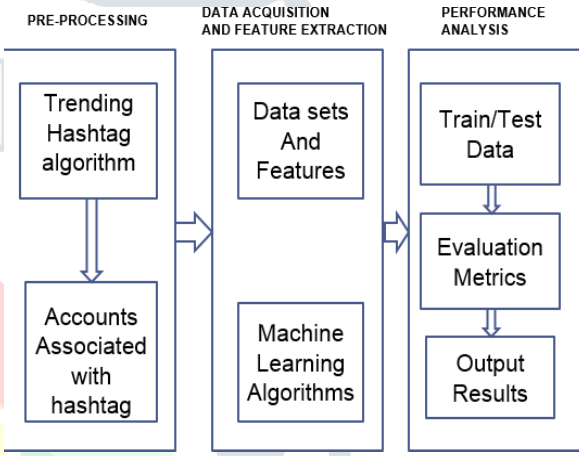
Boolean: Verified, Default_Profile, Image and Has_Extended_Profile.

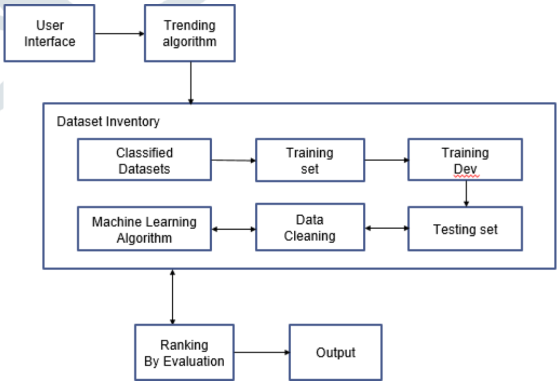
Date: Created_At

Json: Status.

Thus our model train the datasets with the algorithm and evaluate the performance by analysing the results obtained from the KNN and Random Forest algorithms.

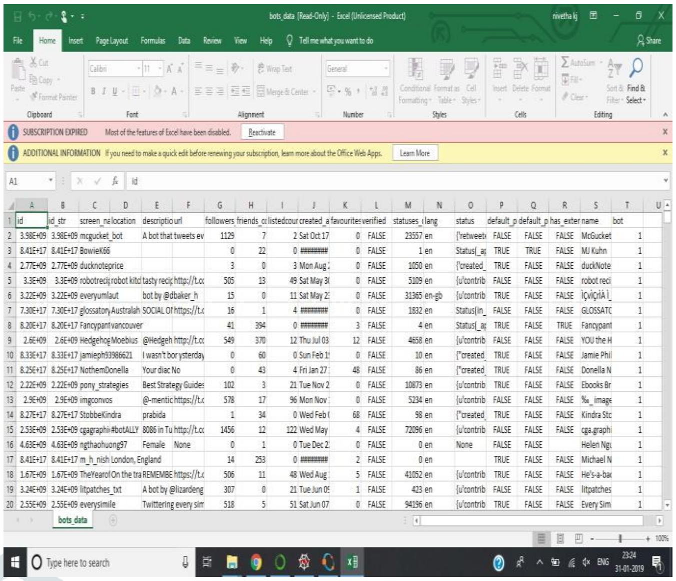
Best accuracy and results will be compared and chosen to detect the bots in twitter. The main architecture of our proposed model:


PRE-PROCESSING
DATA ACQUISITION AND FEATURE EXTRACTION
PERFORMANCE ANALYSIS
Trending Hashtag algorithm
Accounts Associated with hashtag
Data sets And Features
Machine Learning Algorithms
Train/Test Data
Evaluation Metrics
Output Results



OUTPUT:

After processing all training datasets along with its feature, our model produce results in excel file. In this file, it predicts whether the given account is Bot or not. It is represented in binary format in which 1 represented as Bot and 0 as Non Bot. These results displayed in a user interface which looks like web interface. Our Model produce more accurate than the existing models available.



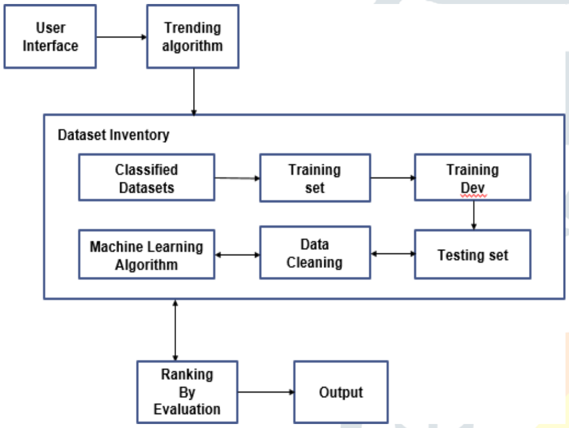
SUBSCRIPTION EXPIRED Most of the features of Excel have been disabled. Reactivate

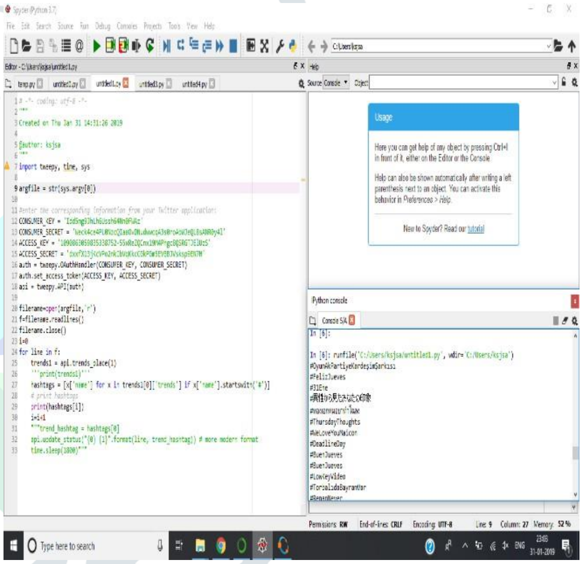
ADDITIONAL INFORMATION If you need to make a quick edit before renewing your subscription, learn more about the Office Web Apps. Learn More

A1 id

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	id	id_str	screen_na	location	descriptio	url	followers_	friends_cc	listedcour	created_a	favourites	verified	statuses_(	lang	status	default_p	default_p	has_exter	name	bot
2	3.98E+09	3.98E+09	mcgucket_bot		A bot that tweets ev		1129	7	2	Sat Oct 17	0	FALSE	23557	en	{'retweet	FALSE	FALSE	FALSE	McGucket	1
3	8.41E+17	8.41E+17	BowieK66				0	22	0	########	0	FALSE	1	en	Status(_ap	TRUE	TRUE	FALSE	MJ Kuhn	1
4	2.77E+09	2.77E+09	ducknoteprice				3	0	3	Mon Aug 2	0	FALSE	1050	en	{'created_	TRUE	FALSE	FALSE	duckNote	1
5	3.3E+09	3.3E+09	robotrecip	robot kitc	tasty recip	http://t.co	505	13	49	Sat May 3(	0	FALSE	5109	en	{u'contrib	FALSE	FALSE	FALSE	robot reci	1
6	3.22E+09	3.22E+09	everyumlaut		bot by @dbaker_h		15	0	11	Sat May 2	0	FALSE	31365	en-gb	{u'contrib	TRUE	FALSE	FALSE	ÌÇvìÇrìÀ ì_	1
7	7.30E+17	7.30E+17	glossatory	Australah	SOCIAL OI	https://t.c	16	1	4	########	0	FALSE	1832	en	Status(in_	FALSE	FALSE	FALSE	GLOSSATC	1
8	8.20E+17	8.20E+17	Fancypant	vancouver			41	394	0	########	3	FALSE	4	en	Status(_ap	TRUE	FALSE	TRUE	Fancypant	1
9	2.6E+09	2.6E+09	Hedgehog	Moebius	@Hedgeh	http://t.co	549	370	12	Thu Jul 03	12	FALSE	4658	en	{u'contrib	FALSE	FALSE	FALSE	YOU the H	1
10	8.33E+17	8.33E+17	jamieph93986621		I wasn't bor ysterday		0	60	0	Sun Feb 1	0	FALSE	10	en	{"created_	TRUE	FALSE	FALSE	Jamie Phil	1
11	8.25E+17	8.25E+17	NothemDonella		Your diac	No	0	43	4	Fri Jan 27	48	FALSE	86	en	{"created_	TRUE	FALSE	FALSE	Donella N	1
12	2.22E+09	2.22E+09	pony_strategies		Best Strategy Guides		102	3	21	Tue Nov 2	0	FALSE	10873	en	{u'contrib	TRUE	FALSE	FALSE	Ebooks Br	1
13	2.9E+09	2.9E+09	imgconvos		@-mentic	https://t.c	578	17	96	Mon Nov	0	FALSE	5234	en	{u'contrib	FALSE	FALSE	FALSE	‰_image	1
14	8.27E+17	8.27E+17	StobbeKindra		prabida		1	34	0	Wed Feb (	68	FALSE	98	en	{"created_	TRUE	FALSE	FALSE	Kindra Stc	1
15	2.53E+09	2.53E+09	cgagraphi	#botALLY	8086 in Tu	http://t.co	1456	12	122	Wed May	4	FALSE	72096	en	{u'contrib	FALSE	FALSE	FALSE	cga.graphi	1
16	4.63E+09	4.63E+09	ngthaohuong97		Female	None	0	1	0	Tue Dec 2	0	FALSE	0	en	None	FALSE	FALSE		Helen Ngu	1
17	8.41E+17	8.41E+17	m_h_nish	London, England			14	253	0	########	2	FALSE	0	en		TRUE	FALSE	FALSE	Michael N	1
18	1.67E+09	1.67E+09	TheYearof	On the tra	REMEMBE	https://t.c	506	11	48	Wed Aug	5	FALSE	41052	en	{u'contrib	TRUE	FALSE	FALSE	He's-a-bac	1
19	3.24E+09	3.24E+09	litpatches_txt		A bot by @lizardeng		307	0	21	Tue Jun 0	1	FALSE	423	en	{u'contrib	FALSE	FALSE	FALSE	litpatches	1
20	2.55E+09	2.55E+09	everysimile		Twittering every sim		518	5	51	Sat Jun 07	0	FALSE	94196	en	{u'contrib	TRUE	FALSE	FALSE	Every Sim	1

23:24 31-01-2019





```
# -*- coding: utf-8 -*-
"""
Created on Thu Jan 31 14:31:26 2019

@author: ksjsa
"""
import tweepy, time, sys

argfile = str(sys.argv[0])

#enter the corresponding information from your Twitter application:
CONSUMER_KEY = 'Idd5mg9JhLh6Ussh64Wn0FWAz'
CONSUMER_SECRET = 'Weck4ce4PL0NzcQIae0xDN.dwwcqA3s0roAoWJeQL0sANR0y4l'
ACCESS_KEY = '1090063059835338752-55xReZQCmx19MAPrgc0QSRGTJE1UzS'
ACCESS_SECRET = 'dxxfX13jXcVPo2nkIbVdKkcCSkPGm5EVEBJVskspEEN7M'
auth = tweepy.OAuthHandler(CONSUMER_KEY, CONSUMER_SECRET)
auth.set_access_token(ACCESS_KEY, ACCESS_SECRET)
api = tweepy.API(auth)

filename=open(argfile,'r')
f=filename.readlines()
filename.close()
i=0
for line in f:
    trends1 = api.trends_place(1)
    '''print(trends1)'''
    hashtags = [x['name'] for x in trends1[0]['trends'] if x['name'].startswith('#')]
    # print hashtags
    print(hashtags[i])
    i=i+1
    """trend_hashtag = hashtags[0]
    api.update_status("{0} {1}".format(line, trend_hashtag)) # more modern format
    time.sleep(1800)"""
```

Usage

Here you can get help of any object by pressing Ctrl+I in front of it, either on the Editor or the Console.

Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in Preferences > Help.

New to Spyder? Read our tutorial

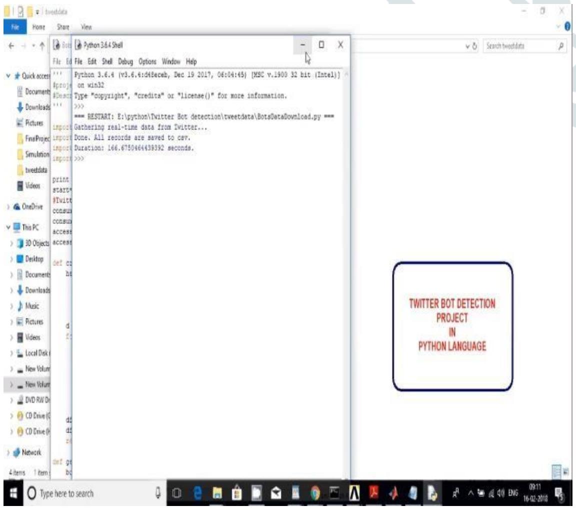
```
In [6]:

In [6]: runfile('C:/Users/ksjsa/untitled1.py', wdir='C:/Users/ksjsa')
#OyunAkPartiyeKardeşimŞarkısı
#FelizJueves
#31Ene
#異性から見たあなたの印象
#[illegible]
#ThursdayThoughts
#WeLoveYouMalcon
#DeadlineDay
#BuenJueves
#BuenJueves
#LowKeyVideo
#TorsalidaBayramVar
#RenanNever
```

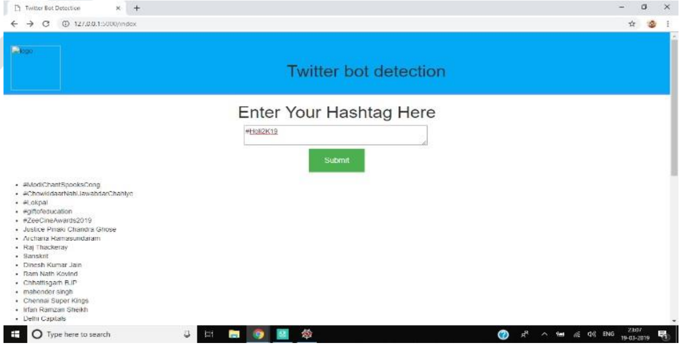
Permissions: RW End-of-lines: CRLF Encoding: UTF-8 Line: 9 Column: 27 Memory: 52 %

23:03 31-01-2019

SCREENSHOTS:



```
Python 3.6.4 (v3.6.4:d48eceb, Dec 19 2017, 06:04:45) [MSC v.1900 32 bit (Intel)]
 on win32
Type "copyright", "credits" or "license()" for more information.
>>>
=== RESTART: E:\python\Twitter Bot detection\tweetdata\BotsDataDownload.py ===
Gathering real-time data from Twitter...
Done. All records are saved to csv.
Duration: 166.6750464439392 seconds.
>>>
```

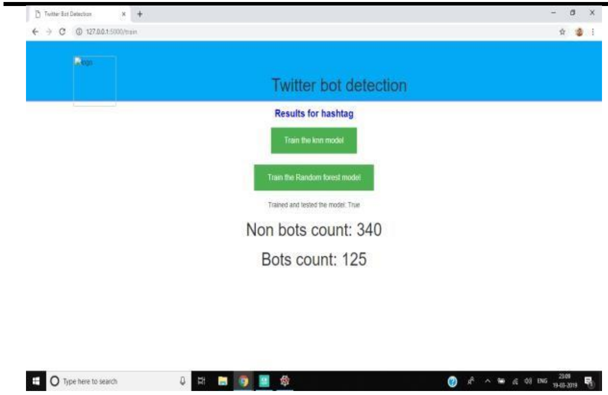


Twitter bot detection

Enter Your Hashtag Here

#Holi2K19

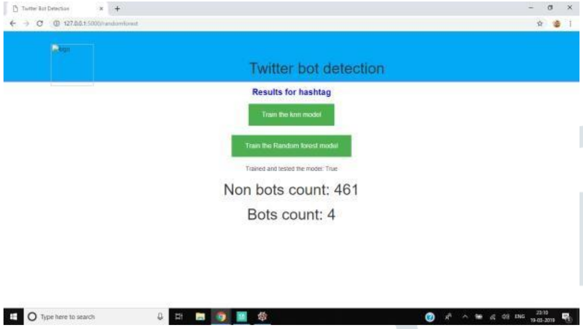
- #ModiChantSpooksCong
- #ChowkidaarNahiJawabdarChahiye
- #Lokpal
- #giftofeducation
- #ZeeCineAwards2019
- Justice Pinaki Chandra Ghose
- Archana Ramasundaram
- Raj Thackeray
- Sanskrit
- Dinesh Kumar Jain
- Ram Nath Kovind
- Chhattisgarh BJP
- mahender singh
- Chennai Super Kings
- Irfan Ramzan Sheikh
- Delhi Capitals



Trained and tested the model: True

Non bots count: 340

Bots count: 125



Trained and tested the model: True

Non bots count: 461

Bots count: 4

CONCLUSION:

In conclusion, after comparing the performance of two different models (KNN and Random Forest) on the problem of classifying a given Twitter user as “bot” or “Not”, we found that the Random Forest had more accuracy on highly confident predictions. Though all two models performed well on the Training and Train-Dev. datasets, the maximum accuracy attained on the Test dataset with a **86.2%.** Looking ahead, we would like to collect our own dataset for further experimentation. Gathering more training examples could improve our distribution mismatch by making the Training distribution more like the real-world.

REFERENCES:

[1]. https://en.wikipedia.org/wiki/Twitter.

[2]. Kyumin Lee, Brian David Eoff, and James Caverlee. Seven months with the devils: A long-term study of content polluters on twitter. In ICWSM, 2011.

[3]. J. P. Dickerson, V. Kagan, and V. S. Subrahmanian, ‘‘Using sentiment to detect bots on Twitter: Are humans more opinionated than bots?’’ in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), Aug. 2014, pp. 620–627.

[4]. K. Lee, B. D. Eoff, and J. Caverlee, “Seven months with the devils: a long-term study of content polluters on twitter,” in In AAAI Intl Conference on Weblogs and Social Media (ICWSM, 2011

[5]. E. Tacchini, G. Ballarin, M. L. D. Vedova, S. Moret, and L. de Alfaro, “Some like it hoax: Automated fake news detection in socialnetworks,” CoRR, vol. abs/1704.07506, 2017.[Online].Available:http://arxiv.org/abs/1704.07506.

[6]. Kyumin Lee, Brian David Eoff, and James Caverlee. Seven months with the devils: A long-term study of content polluters on twitter. In ICWSM, 2011.