**RESEARCH ARTICLE** 

**OPEN ACCESS** 

# **Tweet Based Bot Detection Using Big Data**

Mr D.Purushothaman MCA., M.E.<sup>[1]</sup>, K Pavan<sup>[2]</sup>

<sup>[1]</sup>Asst. Professor, Department of Computer Applications
 <sup>[2]</sup> Student, Department of Computer Applications
 <sup>[1], [2]</sup> Chadalawada Ramanamma Engineering College (Autonomous)

### ABSTRACT

Twitter is one of the most popular micro-blogging social media platforms that has millions of users. Due to its popularity, Twitter has been targeted by different attacks such as spreading rumors, phishing links, and malware. Tweet-based botnets represent a serious threat to users as they can launch large-scale attacks and manipulation campaigns. To deal with these threats, big data analytics techniques, particularly shallow and deep learning techniques have been leveraged in order to accurately distinguish between human accounts and tweet-based bot accounts. In this project, we discuss existing techniques, and provide a taxonomy that classifies the state-of-the-art of tweet-based bot detection techniques. We also describe the shallow and deep learning techniques for tweet-based bot detection, along with their performance results. Finally, we present and discuss the challenges and open issues in the area of tweet-based bot detection.

Keywords: - Big Data, Tweets, Bot Detection, Deep Learning, Social Media.

### I.INTRODUCTION

In recent research work in the field of Twitter social botnet detection. They provided an analytical review of each proposed method with its limitations and advantages. The techniques were classified into three main categories, namely graph-based, machine learning-based, and crowd sourcing based techniques. The crowd sourcing technique uses human intelligence to identify various patterns, which is stated to be the most error prone out of the three techniques. It was also shown that machine learning methods and, more specifically, random forest classifiers are the most commonly used for detecting social bots in Twitter users. In the existing system provided a short comparative survey of the research work in the field of Twitter spam detection within the year range of 2009-2015. They described different detection methods within four categories: account based, tweet-based, graph-based, and hybrid-based methods. The account-based methods were shown to leverage the user profile's metadata like followers and following count and other derived features such as age of the account. While in graph-based methods, features like distance and strength of connectivity between users were shown to be used for spam detection. However, in tweet-based methods, the survey mainly focused on detecting spam using URL and its derived features, such as length and domain name. To detect a spam user, posted URLs were analyzed and classified as malicious or benign. Besides this, the authors highlighted overlooked features that were argued to improve the spam detection.

Another comparative survey was presented in the field of multiplatform spam user detection. The authors recognized that different platforms, such as emails, blogs, or microblogs, require different techniques and features to achieve accurate detection. Therefore, proposed techniques within the year range of 2011-2015 were classified based on the platform that the dataset lies within. A qualitative comparison was conducted for each group of methods under the same platform. The existing system observed that the botnet used a URL network shortening services and redirections to obfuscate the actual landing pages. They disclosed that users clicked on these URLs, found the botmaster establishing the Bursty botnet, and registering landing pages on phishing websites. They confirmed that the botmaster is still successful in owning Twitter bot-related services. This study includes a review and insight into Twitter's cyberspace infrastructure, cybercrime operation, and the dark markets. The system doesn't have technique shallow learning-based detection methods. There is no technique deep neural networks are applied on twitter data to determine the relevant content for users, and hence improve their experience on the platform .

#### **II. RELATEDWORKS**

M Imran and S Asalam Khan proposed Toward an optimal solution against denial-of-service attacks in software defined networks. Software Defined Networking (SDN) separates the control logic from data forwarding and shifts the whole decision power to the controller, making the switch a dumb device. SDNs are becoming more and more important due to the key features like scalability, flexibility and monitoring. The centralized control of SDN makes it vulnerable to different attacks such as Flooding, Spoofing, Denial of Service (DoS), etc. These attacks can degrade the SDN performance by overwhelming its different components such as controller, switch and control channel. This project provides a comprehensive review of different mitigation approaches and categorizes them into three different classes on the basis of their methodology to handle the malicious traffic. In addition to that, we find out limitations in these mitigation approaches and propose the possible features of an optimal solution against DoS attacks.

E Karbab and A Darhab proposed MalDozer: Automatic framework for Android malware detection using deep learning. Android OS experiences a blazing popularity since the last few years. This predominant platform has established itself not only in the mobile world but also in the Internet of Things (IoT) devices. This popularity, however, comes at the expense of security, as it has become a tempting target of malicious apps. Hence, there is an increasing need for sophisticated, automatic, and portable malware detection solutions. In this project, propose MalDozer, an automatic Android malware detection and family attribution framework that relies on sequences classification using deep learning techniques. Starting from the raw sequence of the app's API method calls, MalDozer automatically extracts and learns the malicious and the benign patterns from the actual samples to detect Android malware. MalDozer can serve as a ubiquitous malware detection system that is not only deployed on servers, but also on mobile and even IoT devices.

K Yang and S Wang proposed Arming the public with artificial intelligence to counter social bots. The increased relevance of social media in our daily life has been accompanied by efforts to manipulate online conversations and opinions. Deceptive social bots automated or semi-automated accounts designed to impersonate humans — have been successfully exploited for these kinds of abuse. Researchers have responded by developing AI tools to arm the public in the fight against social bots. Here we review the literature on different types of bots, their impact, and detection methods. use the case study of Botometer, a popular bot detection tool developed at Indiana University, to illustrate how people interact with AI countermeasures. A user experience survey suggests that bot detection has become an integral part of the social media experience for many users. However, barriers in interpreting the output of AI tools can lead to fundamental misunderstandings. The arms race between machine learning methods to develop sophisticated bots and effective countermeasures makes it necessary to update the training data and features of detection tools.

#### III.PROPOSED SYSTEM ARCHITECTURE

The proposed model implemented the bidirectional strategy in which tweet sentences are processed both forward and backward for each layer enabling a better understanding of the overall text context. To train the model, a public dataset Cresci-2017 is used that consists of tweets from 3,474 human accounts and 1,455 bots, resulting in 11.4 million tweets in total. Before training, each tweet was preprocessed and tokenized to fit the word embedding model. A pre-trained GloVe model was used to convert text to numerical vectors that are acceptable by the network. The vectors were fed to a three-layer model with a decreasing dropout layer that was initially set to 0.5. Two subsets of testing datasets, composed of 1,982 and 928 accounts respectively, were used to evaluate their model resulting in precision and accuracy of 93% and 95%, respectively.

In the proposed system introduced a novel deep learning model to distinguish bots from humans using their retweet patterns termed RTbust. Before building the model, the authors analyzed the behavioral patterns of bots and humans alike. The analysis demonstrated a distinctive pattern of retweeting in terms of timing, and it was categorized into four patterns. The \_rst is the droplet pattern, which corresponds to normal users in which there exists a fair amount of time between the tweet being posted and the retweet operation. The three remaining patterns belonged to potential bots due to their suspicious and rapid retweeting pattern. The proposed system offered several innovative approaches that have vastly increased the efficiency of spam identification. The system is more effective due to presence of tweet-based bot detection.

In this proposed system, there are two modules. They are Tweet server and end user.

**1. Tweet Server:** In this Module, it Provides following functionalities like Login, data processing, Train and Test Data sets with Logistic Regression and Naïve Bayes, Generate Trained and Tested Accuracy, View Trained and Tested Accuracy Results, View Predicted Tweet Type Details, Generate tweet type ratio, log out.

**2. End User:** In this Module, it Provides following functionalities like register, login, post tweet dataset, Predict Tweet Type, view profile, Log out.

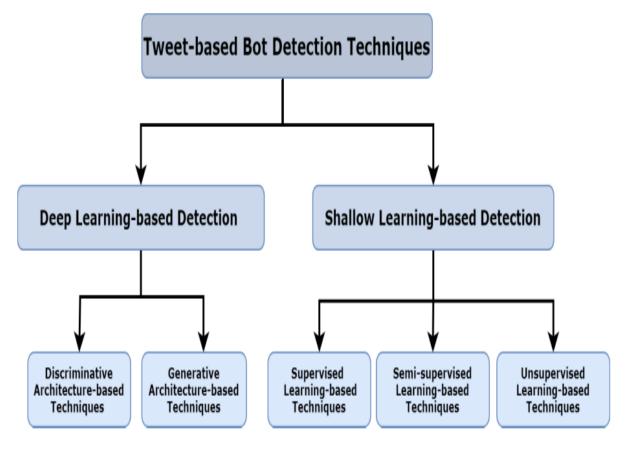


Fig. 1 Proposed system architecture

### **IV. RESULTS AND DISCUSSION**

The output screens obtained after executing and running the system are given from Fig. 2 to Fig.12

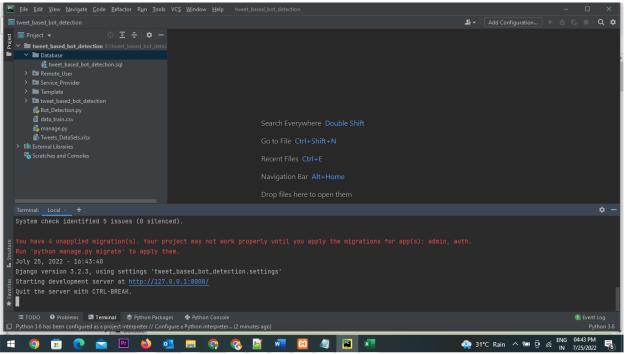


Fig. 2 Generating Link

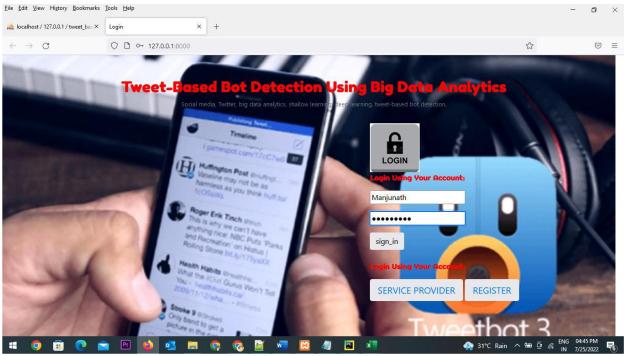


Fig.3 User Login

# International Journal of Computer Science Trends and Technology (IJCST) – Volume 10 Issue 5, Sep-Oct 2022

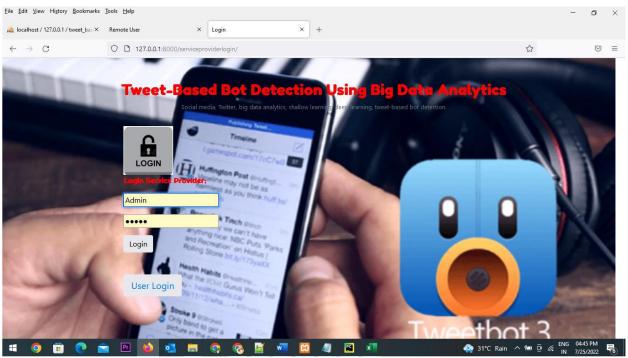


Fig.4 Admin login

									Tweets	s_DataSets.xls	x - Excel						Sign in	<b>a</b> –		×
File	Home	Insert	Page Layout	Formulas	Data	Review	View	Help Ç	) Tell me	e what you w	ant to do								R	Share
Paste V	א ₪ י [	Calibri B I U	•   🖽 •   🖄	A A .	= = =	≫ -   € →	ab Wrap ⊞ Merg	o Text je & Center		v % »	€.0 .00 .00 →.0	Condition Formatting			e Insert	Delete Format	∑ AutoSur ↓ Fill ~ ♦ Clear ~	n × AZT Sort & I Filter × S		
Clipbo	ard 🗔		Font	ß		Aligni	nent		rs I	Number	ß		Styles			Cells		Editing		~
	o ~ ♂ ~	÷																		
<b>B</b> 9	Ŧ	: ×	<i>√ f<sub>x</sub></i> T	agged by @	@CFOAlbaA	sturias @/	AlbaSwee	etCFpic.twit	ter.com/	/sTfpjwrFSC	:									٣
	Α														В					
1	10091 It's	s the everyt	hing else that's	complicat	ed. #PESun	nmit #PXpi	c.twitter	.com/Jsv6B	AFQMI											
2	10172 Ere	en sent a gla	are towards Mil	kasa then r	nodded and	stood up	to go hel	p his lovely	girlfrien	d @Sincere	Pyrrhic.	Once he arri	ived in the	kitchenâž	2-					
3	7012 l p	osted a new	v photo to Face	book http:	//fb.me/28	3e7LiyuJ														
4	3697 #ja	an Idiot Che	lsea Handler Di	iagnoses Tr	rump With	a Disease l	nttps://t.	co/k8PrqcV	/TRI http	s://t.co/dR	N35xtSJZ									
5	10740 Pe	edophile Ant	thony Weiner i	s TERRIFIED	D of Getting	g Beaten U	p in Priso	n https://t.	co/g3bU	9Q4gAg										
6	9572 EB	MUD ending	g penalties for	excessive	water user:	s https://t.	co/D5a1F	MVMHd												
7	10792 Big	g day. #WeT	TheNorth #yyz #	#thesix #su	inset #skyli	ne @ The	Six https:	//www.ins	tagram.c	om/p/BFgr/	A9gBZay/	/								
8	10963 Pa	araguayan m	an linked to He	zbollah fa	ces drug ch	arges in M	iami http	s://t.co/zBl	CgoqO2	I https://t.c	o/vZHbCl	Mx28S								
9			FOAlbaAsturia																	
10	5401 #N	NowPlaying:	De La Soul feat	t. Roc Marc	iano "Prop	erty of http	os://t.co/	KZmwHSA5	inn" #mu	isic https://	t.co/f3LF	KzB1Ct								
11	7636 Th	e Guardian:	Trump's Twee	ts Keep U.S	6. Manufact	urers on T	heir Toes	#MAGA #A	mericaFi	irst #Preside	entTrump	p https://t.o	co/V0sQsk	FolF https	://t.co/j	nxaxLDowy				
12	6908 Te	errorists. http	ps://t.co/RQkQ	XroDNB																
13	5265 Ea	stside Ivo (@	@eastside_ivo)	came thro	ugh Real T	alk on Sun	day. Sorr	we couldr	't be mo	re… https	://www.i	instagram.o	om/p/BFg	q0pRCzfS/						
14	10433 Bo	oston bombi	ing survivor s ch	narity gives	s 1st artifici	al limb #D	emnDeba	ite #DemDe	bate #to	ot										
15	1643 Lo	velyz - Book	mark https://t	twitter.con	n/shaytyen	/status/72	66714322	02117120/v	rideo/1Â	…										
16			uesday everyor																	
17	12562 [#!	MONSTA_X]	#TheCLAN #ë"	í ežœ #ejo	e스íŠ, #ALI	_IN #ì»´ë°:	± Here >>	https://yo	utu.be/w	/NxPGbk-gv	vAÂ http:	s://twitter.c	om/Officia	alMonstaX	/status/	732586501964	365824 …			
18	4941 Oh	h damn!! htt	tps://t.co/KgGP	YXogqf htt	:ps://t.co/a	LrDErjv2U														
19			b.me/7pXHVvK																	
20			he future of tra																	
21	12385 'it	is fear of go	oing into debt t	hat most p	uts people	from poor	er backgr	ounds off g	oing to u	iniversity.' (	G Osborn	ehttp://ww	w.mirror.	co.uk/new	/s/uk-ne	ws/george-os	borne-letter-	13-years-798	9350 â€	<b>_</b>
4	•	Sheet1	Sheet2 Shee	et3   (	Ð								4							▶
Ready	Access	sibility: Investig	gate													H				+ 100%
	9	1	) 💼 🛙	21 📫	•	<b>=</b> (	<b>•</b>	3 📝	w	8	🧊 📲					条 31°C R	ain 🔨 🗐 🤅	ENG IN	04:46 PM 7/25/2022	-
								F	ig.5 '	Tweet	data	aset								

# International Journal of Computer Science Trends and Technology (IJCST) – Volume 10 Issue 5, Sep-Oct 2022

localhost / 127.0.0.1 / tweet_bas ×	Remote User ×	Service Provider ×	+			
$\rightarrow$ C	O D or 127.0.0.1:8000/Add	_DataSet_Details/			۲۵ ۲	${igsidential}$
STATE AND A	Tweet-Base	d Bot Detectio	n Usin <b>g B</b>	ig Data Analy	ytics	
A TRUE FORM TH	The Division of the Owner, where the Own					
OST TWEETS DATA SE	TS P 🥹 File Upload			×		
	← → ~ ↑ <mark> </mark> « Er	nter > tweet_based_bot_det >	✓ Ö 🔎 Search tweet	based bot det	-	
						A TRANS
Browse Tweet Data	Organize 👻 New fold	^		· 🔟 🕜		
	This PC	Name	Date modified	Туре		En l
Upload	3D Objects	idea	7/25/2022 04:42 PM	File folder		
Chicag	Desktop	Database	7/25/2022 04:38 PM	File folder		
	Documents	Remote_User	7/25/2022 04:38 PM	File folder		
Idno Tweet follow	wing fot	Service_Provider	7/25/2022 04:38 PM 7/25/2022 04:38 PM	File folder		
	) Music	tweet based bot detection	7/25/2022 04:38 PM	File folder		
Contraction of the local division of the loc	Pictures	Bot_Detection.py	11/12/2021 05:22 PM	Python File		
	Videos	data train.csv	11/10/2021 06:17 PM	Microsoft Excel		
	Local Disk (C:)	🕞 manage.py	11/10/2021 04:57 PM	Python File		
and the second second	Entertainment (E	Tweets_DataSets.xlsx	11/11/2021 01:08 PM	Microsoft Excel	and the second second	
Contraction of the	MONIKA (F:)					
Contraction of the local distribution of the		<		, 600		
State of the local division of the local div	A Notwork		10.000 00.00			
	Filer	name: Tweets_DataSets.xlsx	✓ All Files (*.*)	~		
			Open	Cancel		
	The second		_			
	Cray barre	o pet a			1 1 2	
	y band a	and a second sec				

Fig.6 Browse dataset

<u>F</u> ile <u>E</u> dit <u>V</u> iew Higtory <u>B</u>	lookmarks <u>T</u> ools <u>H</u> elp	- 0	×
🍌 localhost / 127.0.0.1 / tw	eet_bas × Remote User × Service Provider × +		
$\leftarrow \  \  \rightarrow \  \  G$	○ D 127.0.0.1:8000/Add_DataSet_Details/	${\times}$	=
	Tweet-Based Bot Detection Using Big Data Analytics		Â
POST TWEETS D	ATA SETS PREDICT TWEET TYPE VIEW YOUR PROFILE LOGOUT		
Browse Twe	Pet Data Sets Browse No file selected.		
Upload			
Idno	Tweet		
10091	It's the everything else that's complicated. #PESummit #PXpic.twitter.com/Jsv6BAFQMl		
10172	Eren sent a glare towards Mikasa then nodded and stood up to go help his lovely girlfriend @SincerePyrrhic. Once he arrived in the	kitchenâŽ <sup>-</sup>	_
7012 1	l posted a new photo to Facebook http://fb.me/2Be7LiyuJ		
3697 4	#jan Idiot Chelsea Handler Diagnoses Trump With a Disease https://t.co/k8PrqcWTRI https://t.co/dRN35xt5JZ		
10740	Pedophile Anthony Weiner is TERRIFIED of Getting Beaten Up in Prison https://t.co/g3bU9Q4gAg		
9572	EBMUD ending penalties for excessive water users https://t.co/D5a1FMVMHd		
10792	Big day. #WeTheNorth #yyz #thesix #sunset #skyline @ The Six https://www.instagram.com/p/BFgrA9gBZay/		
10963	Paraguayan man linked to Hezbollah faces drug charges in Miami https://t.co/zBLCgoqO2I https://t.co/vZHbCMx28S		
10778	Tagged by @CFOAlbaAsturias @AlbaSweetCFpic.twitter.com/sTfpjwrFSC		
5401 4	#NowPlaying: De La Soul feat. Roc Marciano "Property of https://t.co/KZmwHSA5nn" #music https://t.co/f3LFKzB1Ct		
7636	The Guardian: Trump's Tweets Keep U.S. Manufacturers on Their Toes #MAGA #AmericaFirst #PresidentTrump https://t.co/VOsQskF	olF https://	t.co ~
4 0 🖬	💽 💼 🖭 🍓 🜉 🛤 😋 🧞 📓 🐙 🔞 🥒 📓 🖉 🕷 🔍	NG 04:49 PM IN 7/25/2022	5

Eile <u>E</u> dit <u>V</u> iew <u>N</u> a	vigate <u>C</u> ode	<u>R</u> efactor R <u>u</u> r	n <u>T</u> ools VC	<u>6 W</u> indow <u>H</u> elp	tweet_based_	bot_detection						-		×
tweet_based_bot_detect									<u>\$</u> -	Add Configura			=   Q	\$
च्च 🖻 Project 👻		9 E ¥	<b>*</b> –											
🖁 🗠 🖿 tweet_based_bo	ot_detection E		ot_detec											
Database			_											
tweet_bas > Image: Remote User	ed_bot_detectio	on.sql												
> Service_Provi														
> 🖿 Template														
> b tweet_based_	bot_detection													
Terminal: Local $ imes$														-
CONFUSION MATRIX														
[[1172 27] [ 113 1082]]														
Naive Bayes														
94.0685045948203	38													
[[1155 44]														
[ 98 1097]]														
pr			1-score	support										
0	0.92	0.96	0.94	1199										
. 1	0.96	0.92	0.94	1195										. 1
accuracy														
accuracy			0.94	2394										
macro avg	0.94 0.94	0.94 0.94	0.94 0.94	2394 2394										
• macro avg weighted avg	0.94	0.94	0.94	2094										- 6
Favo														- 1
*														
≣ TODO				Python Console										
Python 3.6 has been co	nfigured as a pr	oject interprete	r // Configure	a Python interpreter.	(9 minutes ag	o)	_		_				Python	3.6
🖬 🗿 📅	2	Pr 🗳	في ا	<b>E O</b>		w 🔀	🖺 🗴		🔕 31 <sup>o</sup>	C Rain 🔨	æ <u>ê</u> <i>«</i>	ENG 04:		-
· · · ·					<b>V</b>							IN 7/2	5/2022	

Fig.8 Train dataset with Naïve Bayes

<u>File Edit View</u> Higtory E	Bookmarks Tools Help		- 0	×
🙀 localhost / 127.0.0.1 / tw	veet_bas × Remote User	× Service Provider × +		
$\leftarrow \ \rightarrow \ \mathbf{C}$	○ □ 127.0.0.1:8000/Sea	rch_DataSets/	${\times}$	=
POST TWEETS D	IN THE OWNER.	ed Bot Detection Using Big Data Analytics		
PREI	DICTION OF TWEET TYPE !!!	Eren sent a glare towards Nikasg then nodded and stood up to go help his lovely girlfriend <u>SincersPurplis</u> . Once he arrived in the <u>Kirkens</u> ä		
		Enter Tweet Message Here		
_	P	edict		, ,
# 🧿 💼	💽 🖻 🖻 🧕 ।	🗖 🧐 🏀 🔛 🚨 🖉 🖼 🖬 📩 🧠 🖓 31°C Kain A Y 🕮 🙂 🌾 <sub>IN</sub>	04:51 PM 7/25/2022	6
		Fig.9 Enter details for prediction		

# International Journal of Computer Science Trends and Technology (IJCST) – Volume 10 Issue 5, Sep-Oct 2022

<u>File Edit V</u> iew Hi	<u>s</u> tory <u>B</u> ookmarks	<u>T</u> ools <u>H</u> elp						- 6	×
iocalhost / 127.0	0.0.1 / tweet_bas $ imes$	Remote User	× Service Provider	× +					
$\leftarrow \   \rightarrow \   G$		0 127.0.0.1:80	00/Search_DataSets/				ជ	1	⊠ ≡
			Timeline			2 1	1 1 1	NO.	
								-	
	PREDICTION	OF TWEET TYPE !!!	1						
								8	
200									
1			Enter Tweet Message	Here					
200						í.			81
Concernance of the second									
			Predict						
	-								
			Tw	eet Type	Quality				10
Participan	1000						and the second second	-	
1.1.1		A CONTRACT			LA M	100 M	Starte.	and -	
# 🧿 (	🗊 💽 📫	i 🖻 📦 🧕	i 📰 📀 🚱 🖬	🚾 🖾 🥥 📱	a x	条 31°C Rain	 	NG 04:51 PN N 7/25/202	1 2 <b>1</b>

Fig.10 Prediction status

	íools <u>H</u> elp		-	٥	×
localhost / 127.0.0.1 / tweet_bas ×	Remote User × Service Provider	× +			
$\leftrightarrow$ $\rightarrow$ G	O D 127.0.0.1:8000/Predict_Tweet_Type/		5	${igvar}$	≡
	Tweet-Based Bot Detect	tion Using Big Data Anal	ytics		
		ested Accuracy Results View Predicted Tweet Type Details	Find Tweet Type Ratio on Data Sets		
Download Trained Data Sets	View Tweet Type Ratio Results View All Remote Users Log	gout			
View Tweet Type P	diction Details !!!				
Idno Tweet					
10091 It's the e	erything else that's complicated. #PESummit #	#PXpic.twitter.com/Jsv6BAFQMI			
	a glare towards Mikasa then nodded and stop	d up to go belo his lovely girlfriend @Sincerel			
10172 Eren sen		or the for the prime in the start and the since real	Pyrrhic. Once he arrived in	the k	itch
	new photo to Facebook http://fb.me/2Be7Liyu		Pyrrhic. Once he arrived in	the k	itch
7012 I posted		Ŋ		the k	itch
7012 I posted 3697 #jan Idio	new photo to Facebook http://fb.me/2Be7Liyu	IJ sease https://t.co/k8PrqcWTRI https://t.co/dRN		the k	itch
7012 I posted 3697 #jan Idio 10740 Pedophil	new photo to Facebook http://fb.me/2Be7Liyu Chelsea Handler Diagnoses Trump With a Dis	J sease https://t.co/k8PrqcWTRI https://t.co/dRN aten Up in Prison https://t.co/g3bU9Q4gAg		the k	itch
7012 I posted 3697 #jan Idio 10740 Pedophil 9572 EBMUD	new photo to Facebook http://fb.me/2Be7Liyu Chelsea Handler Diagnoses Trump With a Dis Anthony Weiner is TERRIFIED of Getting Bea	J sease https://t.co/k8PrqcWTRI https://t.co/dRN aten Up in Prison https://t.co/g3bU9Q4gAg s://t.co/D5a1FMVMHd	35xtSJZ	the k	
7012         I posted           3697         #jan Idio           10740         Pedophil           9572         EBMUD           10792         Big day.           10963         Paragua	new photo to Facebook http://fb.me/2Be7Liyu Chelsea Handler Diagnoses Trump With a Dis e Anthony Weiner is TERRIFIED of Getting Bea ending penalties for excessive water users http WeTheNorth #yyz #thesix #sunset #skyline @ an man linked to Hezbollah faces drug charge	iJ sease https://t.co/k8PrqcWTRI https://t.co/dRN aten Up in Prison https://t.co/g3bU9Q4gAg is://t.co/D5a1FMVMHd p The Six https://www.instagram.com/p/BFgrAs s in Miami https://t.co/zBLCgoqO2I https://t.co	9gBZay/	the k	
7012         I posted           3697         #jan Idio           10740         Pedophil           9572         EBMUD           10792         Big day.           10963         Paragua	new photo to Facebook http://fb.me/2Be7Liyu Chelsea Handler Diagnoses Trump With a Dis e Anthony Weiner is TERRIFIED of Getting Bea ending penalties for excessive water users http WeTheNorth #yyz #thesix #sunset #skyline @	iJ sease https://t.co/k8PrqcWTRI https://t.co/dRN aten Up in Prison https://t.co/g3bU9Q4gAg is://t.co/D5a1FMVMHd p The Six https://www.instagram.com/p/BFgrAs s in Miami https://t.co/zBLCgoqO2I https://t.co	9gBZay/	the k	
7012         I posted           3697         #jan Idio           10740         Pedophil           9572         EBMUD           10792         Big day.           10963         Paragua	new photo to Facebook http://fb.me/2Be7Liyu Chelsea Handler Diagnoses Trump With a Dis e Anthony Weiner is TERRIFIED of Getting Bea ending penalties for excessive water users http WeTheNorth #yyz #thesix #sunset #skyline @ an man linked to Hezbollah faces drug charge	IJ sease https://t.co/k8PrqcWTRI https://t.co/dRN aten Up in Prison https://t.co/g3bU9Q4gAg is://t.co/D5a1FMVMHd ? The Six https://www.instagram.com/p/BFgrAs s in Miami https://t.co/zBLCgoqO2I https://t.co er.com/sTfpjwrFSC	9gBZay/	the k	

Fig.11 Predicted tweet type details

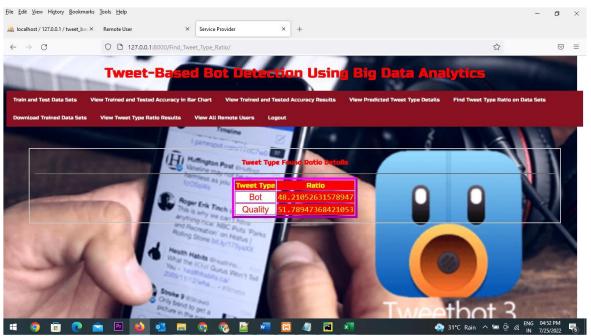


Fig.12 Tweet type ratio details

### V. FUTURE SCOPE AND CONCLUSION

Twitter is one of the most popular social media platforms that allows connecting people and helps organizations reaching out to customers. Tweet-based botnet can compromise Twitter and create malicious accounts to launch large-scale attacks and manipulation campaigns. In this review, we have focused on big data analytics, especially shallow and deep learning to fight against tweet-based botnets, and to accurately distinguish between human accounts and tweet-based bot accounts.

### REFERENCES

- M. Mohsin. (2020). 10 Social Media Statistics You Need to Know in 2021. [Online]. Available: https://www.oberlo.com/blog/social-mediamarketing-statistics
- [2] I. Arghire. (2020). Twitter Hack: 24 Hours From Phishing Employees to Hijacking Accounts. https://www.securityweek.com/twitter-hack- 24hours-phishing-employees-hijacking-accounts.
- [3] The Rise of Social Media Botnets. Accessed: Feb. 21, 2021. [Online].Available: https://www.darkreading.com/attacks-breaches/the-rise-of-social-media-botnets/a/d-id/1321177
- [4] M. Imran, M. H. Durad, F. A. Khan, and A. Derhab, "Toward an optimal solution against denial of service attacks in software defined

networks," Future Gener. Comput. Syst., vol. 92, pp. 444–453, Mar. 2019.

- [5] M. S. Savell. (2018). Protect Your Company's Reputation From Threats by Social Bots. [Online]. Available: https://zignallabs.com/blog/protect-yourcompanys-reputation-fromthreats- by-socialbots/
- [6] S. Aslam. (2021). Twitter by the Numbers: Stats, Demographics &Fun Facts. [Online]. Available: https://www.omnicoreagency.com/twitterstatistics/
- [7] A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and openissues," Knowl.-Based Syst., vol. 189, Feb. 2020, Art. no. 105124