Identification of Fake profiles using Artificial Neural Network

Kompally Rekha¹, Malgi Aadarsh², Maggidi Srikar³, D. Sandhya Rani⁴

^{1,2,3}B.Tech Student, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad, Telangana, India

⁴Assistant Professor, Department of Computer Science and Engineering, CMR Technical Campus, Medchal, Hyderabad Telangana, India

Abstract- In day today life Social media became a part of everyone's life. Twitter and Facebook become the major platform for using social media. Nowadays social media is used for spreading real time information, and we have to make sure that the information published is real. But some people who tend to spread the fake news of a person or a product uses the fake Id's called as bots. So it is difficult for a common user to know which news is genuine and which is false news. These fake ID's if used in large scale can create a huge damage to the society. The main objective of our project is to identify these fake ID's. In this project to find the fake twitter account we presented a classification method. The paper determines the minimized set of attributes that influence for finding the fake Twitter account, and then the determined factors are applied using different classification techniques. The most accurate algorithm has been determined by comparing the techniques results. By using minimum set of attributes the amount of data to be analyzed is reduced and we can get faster results.

1.INTRODUCTION

Recently, Facebook had a data breach which affected about 50 million users. Facebook provides a set of clearly defined provisions that explain what they do with the user's data. The policy does very little to prevent the constant exploitation of security and privacy. Fake profiles seem to slip through Facebook's built-in security features. The other dangers of personal data being obtained for fraudulent purposes is the presence of bots and fake profiles. Bots are programs that can gather information about the user without the user even knowing. This process is known as web scraping. What is worse, is that this action is legal. Bots can be hidden or come in the form of a fake friend request on a social network site to gain access to private information.

The solution presented in this paper intends to focus on the dangers of a bot in the form of a fake profile on your social media. This solution would come in the form of an algorithm. The language that we chose to use is Python. 2

The algorithm would be able to determine if a current friend request that a user gets online is an actual person or if it is a bot or it is a fake friend request fishing for information. Our algorithm would work with the help of the social media companies, as we would need a training dataset from them to train our model and later verify if the profiles are fake or not. The algorithm could even work as a traditional layer on the user's web browser as a browser plug-in

2.PROPOSED SYSTEM

This concept is based on the confidence that humans usually behave differently than the fakes, therefore, detecting this behavior will lead to the revealing of the fake accounts. In this section, we will demonstrate some of the works that have been presented in this area. It has reached an accuracy of 84.5% to detect spammers by identifying 23 attributes, most of these attributes. (17 attributes) are demonstrated. However, in our research, we have reached more accuracy with smaller set of attributes as will be discussed. In the set of attributes has been minimized by identifying ten attributes for detected. However, in the previous research, the result was not promising for identifying fake accounts with more optimistic perspective that it is able to identify fake tweets with higher accuracy by using the support of graphical method.

Although has presented a minimized set of attributes which contained six attributes, however, it is mentioned that it could only detects determined types of spammers, they are bagger, and poster spammers. In our approach, we propose minimized set of attributes for detecting all types of false news. In addition, one of these attributes requires text analysis procedure for finding the similarities among messages which is not required for our proposed approach. Moreover, it is mentioned in that Random Forest algorithm is the best results for detection for Twitter

3. ARCHITECTURE



DESCRIPTION

Social life of everyone has become linked with online social networks in today's generation. These websites have drastically affected the way we do our social lives. Making new friends and staying in touch with them has become much easier. However, as a result of their rapid growth, plenty of new issues have emerged, including malicious users, fraudulent profiles, and online impersonation. There are no practical ways for resolving these issues. In this research paper, we use an artificial neural network to detect fake profiles automatically and effectively. We determine the risk that a Facebook friend request is genuine or not. This can be used by online social networks with millions of profiles that can't be manually reviewed

4.TESTING

Implementation and Testing:

Implementation is one of the most important tasks in project is the phase in which one has to be cautions because all the efforts undertaken during the project will be very interactive. Implementation is the most crucial stage in achieving successful system and giving the users confidence that the new system is workable and effective. Each program is tested individually at the time of development using the sample data and has verified that these programs link together in the way specified in the program specification. The computer system and its environment are tested to the satisfaction of the user.

4.1 Implementation

The implementation phase is less creative than system design. It is primarily concerned with user training, and file conversion. The system may be requiring extensive user training. The initial parameters of the system should be modifying as a result of a programming. A simple operating procedure is provided so that the user can understand the different functions clearly and quickly. The different reports can be obtained either on the inkjet or dot matrix printer, which is available at the disposal of the user. The proposed system is very easy to implement. In general implementation is used to mean the process of converting a new or revised system design into an operational one.

4.2 Testing

Testing is the process where the test data is prepared and is used for testing the modules individually and later the validation given for the fields. Then the system testing takes place which makes sure that all components of the system property functions as a unit. The test data should be chosen such that it passed through all possible condition. Actually testing is the state of implementation which aimed at ensuring that the system works accurately and 40 efficiently before the actual operation commence. The following is the description of the testing strategies, which were carried out during the testing period.

4.2.1 System Testing

Testing has become an integral part of any system or project especially in the field of information technology. The importance of testing is a method of justifying, if one is ready to move further, be it to be check if one is capable to with stand the rigors of a particular situation cannot be underplayed and that is why testing before development is so critical. When the software is developed before it is given to user to user the software must be tested whether it is solving the purpose for which it is developed. This testing involves various types through which one can ensure the software is reliable. The program was tested logically and pattern of execution of the program for a set of data are repeated. Thus the code was exhaustively checked for all possible correct data and the outcomes were also checked.

4.2.2 Integration Testing

After the module testing, the integration testing is applied. When linking the modules there may be chance for errors to occur, these errors are corrected by using this testing. In this 41 system all modules are connected and tested. The testing results are very correct. Thus the mapping of jobs with resources is done correctly by the system.

4.2.3 Acceptance Testing

When that user fined no major problems with its accuracy, the system passers through a final acceptance test. This test confirms that the system needs the original goals, objectives and requirements established during analysis without actual execution which elimination wastage of time and money acceptance tests on the shoulders of users and management, it is finally acceptable and ready for the operation.

Test case	Test casename	Test casedesc	Test steps			Test case stats	Test
id			step	Expected	Actual		priority
01	Admin login	Test whether the adminis login or not into the system	If the admin cannot enter the valid details	Admin cannot login the system	Admin canlogin to thesystem	High	High
02	Generate ANN train model	Verify theANN train model generateor not	Without login tothe system	Cannot generate the ANN trainmodel	ANN train model generated successfully	High	High
03	View ANN Train dataset	Verify the ANN train dataset displayed or not	Without generating the training model	the ANN train dataset details may not displayed	The ANN train dataset details displayed successfully	High	High
04	Checking user accounts	Verify either the user account is fake or genuine	Without entering the user account details	We cannot find that account is fake or genuine	Account predicted details are displayed successfully	High	High





Screenshot 1: Home page

© April 2023 | IJIRT | Volume 9 Issue 11 | ISSN: 2349-6002

Ģ	0 location	e500/outwo						• • •	6
	e Profi								
45	te Oliestere	Pressons 1							
				Data Colle	ction Step - 1				
	8	ala.	screen_name	lecation	description	ut	followers_count	Tiends_count	lister;
1	8.550000+17	101574535693441715	2 NootoPoeNar	"Haution, 7.8"	Mare and strong Fragmen in Headen UK, See nove Fragmen at Highurt contents/CarPer	https:/t.com/w2tr/Her	1211	#:)	τα.
							Activati So se So	Windows Ingels and an	10.10
_								a ~ 💵 🖬	1/4/10
0	type here to sea	eta:					····		
01	lype here to ass	icii	S	creenshot 2:	Data Collecti	on			
1	lype here to an	terior X +	S	creenshot 2:	Data Collecti	on		-	0
) 1at	e Account Public Ce	tector X +	S	creenshot 2	Data Collecti	on		- 2000	0 6
) fair ÷	e Account Profile Co C © local ake Pro	ntion × + Next5000/preprocess	s	creenshot 2:	Data Collecti	on		- * 0 •	5
→ F	s Account Profile Co C © local ake Pro	eetion × + Reat5000/preproces ofile Detection	S	creenshot 2:	Data Collecti	on		- 2 0 8	5 6 0
→ F(e Account Husfie Co C © toca a k.e. Pro	tedion X + Heat5000/preprocess offile Detection	ion	creenshot 2:	Data Collecti	on		- * 0 e	0 6 0
→ F	s Account Profile Co C @ Toose a ke Pro	edion X + Peet5000/preprocess ofile Detection	S	Cereenshot 2:	Data Collecti	on		- 1 0 8	- 7 6 ()
→ F(e Account Huffle Co C © Ioca a k.e. P.r.o Data Contecta	tation x + Rest5000/preprocess offile Detection on Preprocess 1	ion totowers, court	Preproce	Data Collecti ess Step - 2	on Issee_count	141, 1209	- 	0 6 0
→ F(e Account Profile Ce C @ Toose a ke Pro Data Collecta	tedion X + PrevESODO/preprocess ofile Detection on Preprocess	ion televers_count t201	Preproc	Data Collecti ess Step - 2	on I I I I I I I I I I I I I I I I I I I	MK,0000 0	- 	5
→ F	e lessent Polite De C © tour a ke Pro	tedion x + Heat5000/preprocess ofile Detect on Preprocess 1 tates, count 1	S	Preproc	ess Step - 2	On I	Бек,code 0 0	- \$ 0 0 1	5
→ F(e Account Puelle Ce C © local a ke Pro Duta Collecti 0 765 1 21 2 713	tation × + heat5000/preprocess offile Detection on Preprocess 1 heats_count 1 54	ion biowers, court table 1 1 105	Preproce	EData Collection	ON	88%,code 0 0 0	+ 0 e a	5
→ F2	e laceneri Hudie Ce C © tora ake Pro Dota Collecti 1 21 3 675	tection x + heat5000/preprocess offile Detection on Preprocess 1 hases_count 1	ion toloesta_count toloesta_count toloesta_count toloesta_count toloesta_count toloesta_count	Preproc	Data Collecti Step - 2 fivoutini, court 0 35 0 0	on Ballet_court 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	вяк,софя 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000	+ 0 6 1	σ 6 0
÷ F(tedion x + Heat5000/preprocess	sion toinees, count taitees, count taitees, count taitees, count taitees, count taitees, count	Preproc	E Data Collection	ON Solution So	bex_code 0	+ O C	0 6 Q
F	e Account Puelle Ce C © local a ke Pro Duta Collecti 1 2 713 5 675 6 11	tation × + heat5000/preprocess offile Detect on Preprocess	ion biowes, court talees 1 1065 10 11	Preproce	Every Step - 2	On Son Son Son Son Son Son Son So	bex_ccde 0	→ → → → → → → → → → → → →	5
		testion x + heat5000/preprocess offile Detection on Preprocess 1 tases_count 1 54	tolovers_count tolovers_count 1 1006 101 11 11 1006 1007 1007 1007 10	Preproce sass 0 5 5 5 5 5 5 5 5 5 5 5 5 5	Evaluation Collection	ON Biblied_roounti S0 S0 S0 S0 S0 S0 S0 S0 S0 S0	 век,софе о 	→ O O O O O O O O O O O O O O O O O O O	σ 6 ()
F	Igen here to and Market C O Ioca a ke P r o a ke 713 a co 713 a co 713 a co 605 a co 258	tedion x + heat5000/preprocess of Preprocess tates, count 5 54	S Con Con Con Con Con Con Con Co	Preproc	E Data Collection	On Son Son Son Son Son Son Son So	ssc.cods 0		0 6 Q

Screenshot 3: Preprocessing

© April 2023 | IJIRT | Volume 9 Issue 11 | ISSN: 2349-6002

🕲 File Account Profile Detection 🗴 🛨	- 0)
€ → C © loahes500(tain	÷ 0 • & 0
Training Step - 3	
Trained Accuracy : 77.10196779964222 %	
Status Count:	
1994	
Fallower Count:	
Friends Count:	
- 0. 	
Faviourate Count.	
•	
Liste:	Activate Windows
	Go to Settings to activate Windows.
atorg for locatha 17 O Type here to search 🛛 🖓 🖆 🍠 🥅 🦓 🚱 🧃 🖉 🖉 🖉	· · 영문 ······· 문
	314330
Screenshot 4: Training	
	- 5 - 5
€ + Ø 10 hatter billigen bereiten in der Hilligkeiter bereiten die Bereiter dieser Berg-14	0 0 0 0 0 1
Fake Account Profile Detection	
The survey Descent Descent Press	
Prediction Step - 4	
Prediction : Normal user Account	
Hedicion , Normai user Account	
🖬 🔘 Type have to search 💦 📮 📮 💭 🛄 🗶 🗐 🔯 🚳 🚳 🚳 🚳 🚳 🚺 🖉 🛃 🖉	A - 20 D 1000

Screenshot 5: Testing

6.CONCLUSION

We have given a framework using which we can identify fake profiles in any online social network by using Neural Network with a very high efficiency as high as around 77%. The model presented in this project demonstrates Neural Network (NN) is an elegant and robust method for classification in a large dataset. Regardless of the non-linearity of the decision boundary, NN is able to classify between fake and genuine profiles with a reasonable

© April 2023 | IJIRT | Volume 9 Issue 11 | ISSN: 2349-6002

degree of accuracy. This method can be extended on any platform that needs classification to be deployed on public profiles for various purposes. This project uses only publicly available information which makes it convenient for users.

REFERENCES

[1] Qiang Cao, Michael Sirivianos, Xiaowei Yang, and Tiago Pregueiro, "Aiding the detection offake accounts in large scale social online services," in Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation, 2012.

[2] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete, "Information credibility on twitter," in Proceedings of the 20th international conference on Worldwide web, 2011.

[3] B. Y.E. Ferrara, O.Varol, C. Davis, F.Menezer, and A. Flammini, "The rise of social bots," Commum.ACM, vol.59 No.7,pp.96-104,2016.

[4] Turing, A.M. Computing machinery and intelligence. Mind 49, 236 (1950), 433-460

[5] Weizenbaum, J. ELIZA—A computer program for the study of natural language communication between man and machine. Commun. ACM 9, 1 (Sept. 1966), 36–45

[6] M. Camisani-Calzolari. (2012, August) Analysis of Twitter followers of the US Presidential Election candidates: Barack Obama and Mitt Romney. (Online). http://digitalevaluations.com/