International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9, Issue-7, May 2020

Fake User Detection in Twitter using Random forest algorithm with Python

Sai Poojitha Bommadevara, A. Jitendra, S. Babu

Abstract: ***Millions of users are engaged with social networking sites around the world. Social sites like twitter, Facebook have a large impact on rare unwanted consequences caused in our regular life in user's interactions. In order to disperse a large amount of inappropriate and harmful data protruding social networking sites are made as a target platform for the spammers. Twitter is main example that has become one of the important platforms for unreasonable amount of spam in all the tomes for fake users to tweet and promote websites or services that crates a major effect for legitimate users and also it disturbs resource consumption. By resulting the opening for unusual and harmful information there is an increase of fake identities that expands invalid data. Research on current online social networks (OSN) is quite natural for detection of fake users on twitter. In this paper using random forest classifier and ROC curve to detect fake users.***

Keywords: Classifier, Detection, Fake user, Twitter.

I. INTRODUCTION

Twitter has become a mainstream media center point where individuals can share news, jokes and discussion about their states of mind and examine news occasions. In Twitter clients can send Tweets in a split second to his/her devotees. Likewise, Tweets can be recovered utilizing Twitter's ongoing internet searcher. The positioning of tweets right now relies upon numerous variables, one of which is the client's number of supporters. Twitter's fame has made it an alluring spot for spam and spammers of numerous types. Spammers have different objectives: spreading promoting to produce deals, phishing or essentially simply trading off the framework's notoriety. Given that spammers are progressively showing up on twitter, the achievement of continuous pursuit administrations and mining apparatuses lies in the capacity to recognize important tweets from the spam storm .There are different approaches to battle spam and spammers, for example, URL boycotts, aloof long range informal communication spam traps, manual order to produce datasets used to prepare a classifier that later will be utilized to identify spam and spammers.

Revised Manuscript Received on May 30, 2020.
***** Correspondence Author
Sai Poojitha Bommadevara*, CSE, V R Siddhartha Engineering College, Vijayawada, India. Email: spoojitha24@gmail.com
A.Jitendra, CSE, V R Siddhartha Engineering College, Vijayawada, India. Email: jitendra@vrsiddhartha.ac.in
S.Babu CSE, V R Siddhartha Engineering College, Vijayawada, India. Email: babunaidu.504@gmail.com

II. RELATED WORK

There has been late enthusiasm for the recognition of malignant as well as phony clients from both the online interpersonal organizations and computer organizing networks. For example, Wang [1] looks at diagram put together highlights to distinguish bots with respect to Twitter, while Yang, Harkreader, and Gu [2] consolidate comparable chart-based highlights with syntactic measurements to manufacture their classifiers. Thomas et al. [3] utilize a comparable arrangement of highlights to give a review examination of a huge arrangement of as of late suspended Twitter accounts. Boshmaf et al. [4] rather make bots (instead of recognizing them), guaranteeing that 80% of bots are imperceptible and that Facebook's Immune framework [5] couldn't identify their bots. Lee, Caver lee, and Webb [6] make "honeypot" records to draw the two people and spammers away from any confining influence, at that point give a measurable examination of the vindictive records they distinguished. In PC systems inquire about, the discovery of Sybil accounts in PC systems has been applied to interpersonal organization information; these procedures will in general depend on the "quick blending" property of a system—which may not exist in informal communities [7] and don't scale to the size of present day social networks. Most important is ongoing work by (Twitter worker and hostile to spam engineer) Chu and partners [8], [9], which uses diagram theoretic, syntactic, and some semantic highlights to arrange people, bots, and cyborgs (human-helped bots) in a Twitter dataset. Twitter has been broadly utilized since 2006, and there is some related writing in twittering [9], [10], [11]. To more readily comprehend microblogging utilization and networks, Java et al. [9] concentrated more than 70,000 Twitter clients and classified their posts into four primary gatherings: every day gab (e.g., "going out for supper"), discussions, sharing data or URLs, and detailing news. Their work likewise examined 1) the development of Twitter, indicating a straight development rate; 2) its system properties, demonstrating the proof that the system is sans scale like other informal organizations [12]; and 3) the geological dispersion of its clients, demonstrating that most Twitter clients are from the United States, Europe, and Japan.

III. METHODOLOGY

The process of our methodology consists of following phases they are,

- Twitter Data
- Data Analysis
- Feature Extraction
- Model selection
- Validation

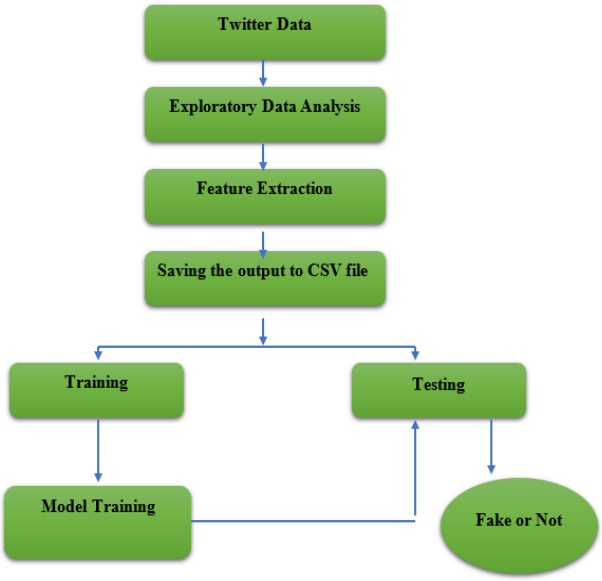

Fig 1. Proposed methodology

A. Twitter Data

The was taken from Tweepy and partitioned it into preparing set and test set, using machine learning algorithm and preparing information we've prepared our classifier model. So as to utilize machine learning model to recognize counterfeit twitter accounts, we required a marked assortment of clients, pre named phony or certifiable. We get the constant information from tweepy API which comprise of 2798 preparing set and 578 test set. The dataset is separated into 70% (training set) and 30% (test set) on which information exploratory examination has been done just as it is additionally investigated to highlight extraction and highlight designing, both preparing and test set are spared in .CSV format.

B. Units

Our Dataset contains 20 qualities out of which we have chosen 8 traits dependent on the spearman connection.

- ID: These IDs are extraordinary 64-piece unsigned whole numbers, which depend on schedule, rather than being successive. The full ID is made out of a timestamp, a specialist number, and a succession number. While devouring the API utilizing JSON, it is critical to consistently utilize the field id_str rather than id.
- Friends_count: It delineates that the friends_count ought to be in appropriate proportion with the follower_count.
- Follower_count: As the friends_count is subject to devotee count; these two properties are emphatically corresponded.
- Listed_count: A Twitter list is actually a rundown of individuals on Twitter that are by one way or another associated. They may have a place with a specific classification (for example news associations), or they may be associated through their substance (for example identified with planting), or they may even be associated through an occasion that they're all joining in (for example the Oscars).
- Favourite_count: It's the quantity of tweets that given client has set apart as top choice.
- Checked: A record might be confirmed in the event that it is resolved to be a record of open intrigue.
- Statuses_count: Returns the latest Tweets composed by the confirming client that have been retweeted by others.
- Default_profile: Profile that who has not given enough data on the profile are in all likelihood has a place with bot or spammers.
- Default_profile_image: spammers and individuals who pester others regularly use accounts without a profile picture

C. Feature Extraction

In this phase, we extracted the features used to detect fake user or not, for that the required are

- Screen name
- Name
- Description
- Status
- Verified
- Followers
- Count of friends
- Count of Statuses
- Bot

These are the features data set have and these are used in training or testing phase also respectively.

D. Model Selection

We actualize different machine learning models and demonstrate and consequently discover their exactness on test and preparing set. We will likewise plot the ROC bend which is a graphical plot made by plotting the genuine positive rate against the bogus positive rate at the different edge which portrays the execution of models

Random forest classification

Random Forest is an adaptable, easy-to-use machine learning algorithm that produces an amazing result more often than not, even without hyper-parameter tuning. Furthermore, it is one of the most used equations because it is effortless and the way it can be used for characterization and relapse errands quite well.

Proposed algorithm:

Step 1: Creating copy of dataframe in train_set`
Step 2: Converting id to int
Step 3: Replacing Null values with 0 in Friends_count column
Step 4: Replacing Null values with 0 in Followers_count column

Step 5: Preparing bag of words for bot accounts
Step 6: Converting verified account into vectors (True->1 False->0)
Step 7: If the name, description, screen_name, status columns contains bot, then Store the data set in predicted_set_1
7.1: Assign bot column as 1 (Fake)
7.2: Store the rest data set in verified_set for next step
Step 8: For all verified account, Store the data set in predicted_set_2
8.1: Assign bot column as 0 (NON_Fake)
8.2: Store the rest data set in followers_following_set for next step
8.3: predicted_set_1=: concatenate (predicted_set_1, predicted_set_2)
Step 9: If followers_count is less than 50 and statuses_count greater than 1000, then then Store the data set in predicted_set_2
9.1: Assign bot column as 1 (Fake)
9.2: Store the rest data set in followers_retweet_set for next step 9.3: predicted_set_1=: concatenate (predicted_set_1, predicted_set_2)
Step 10: If followers_count is less than 150 and statuses_count greater than 10000 then then Store the data set in predicted_set_2
10.1: Assign bot column as 1 (FAKE)
10.2: predicted_set_1=: concatenate (predicted_set_1, predicted_set_2)
Step 11: Store the rest data set in predicted_set_2 and assign bot column as 0 (NON_Fake)
11.1: predicted_set_1=: concatenate (predicted_set_1, predicted_set_2)
Step 12: Return predicted_set_1

E. Evaluation

The validation results are as shown in results and discussion section. In this phase, for validation purpose we use twitter data from tweepy API respectively. We identify the required results through ROC curve and as well as observes the training and testing accuracy respectively

IV. RESULTS AND DISCUSSION

The classifier's final output will be of the form as seen below-the class mark ' Fake' reflects whether a given user is bot or a real user.

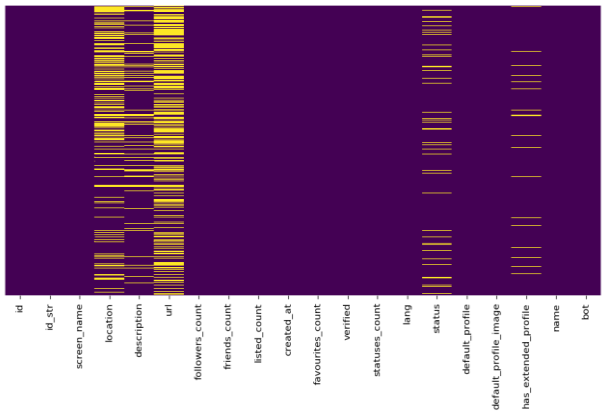

Fig 2. Structure of twitter data set

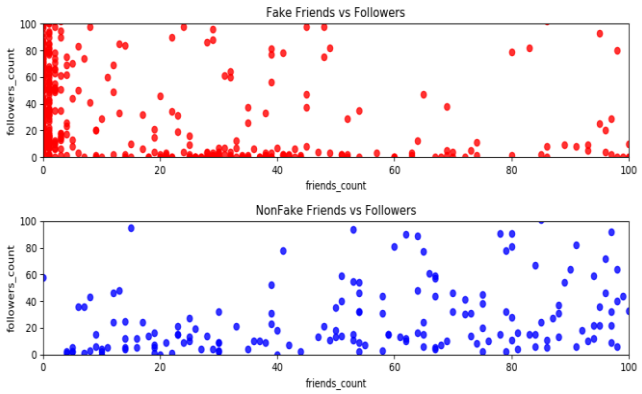

Fig 3. Identifying the Amusingness in the data

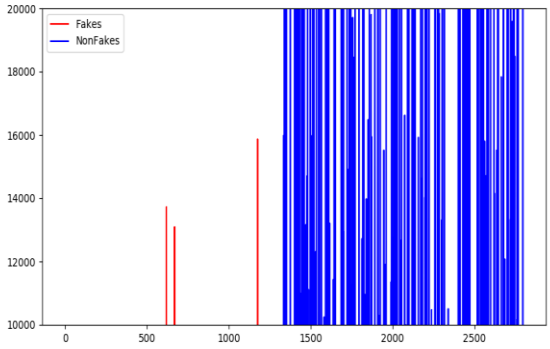

Fig 4. Identifying the Imbalance in the data

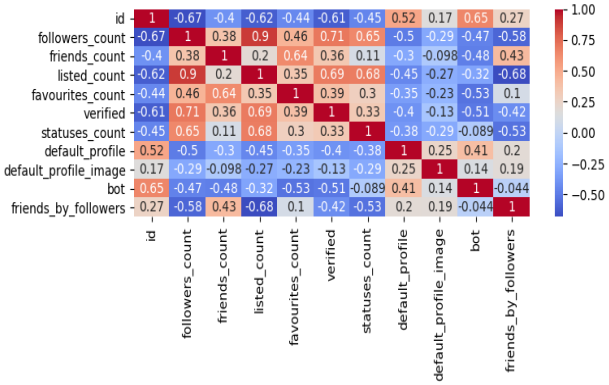

Fig 5. Spearman Correlation Result

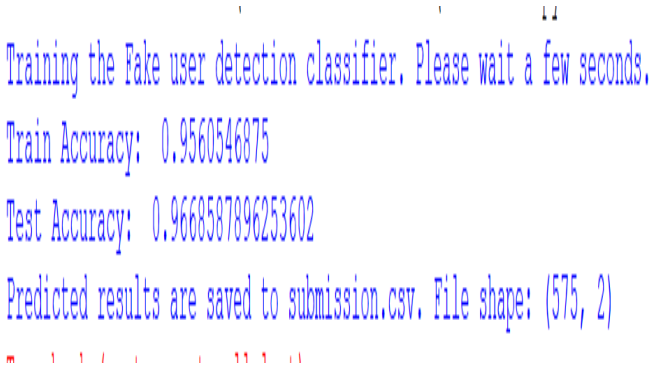

Fig 6. Performance analysis of training and testing data

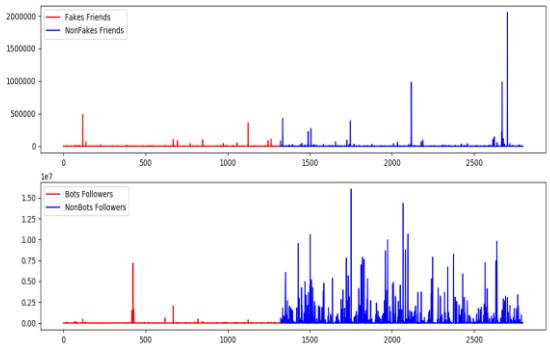

Fig 7. Fake user detection

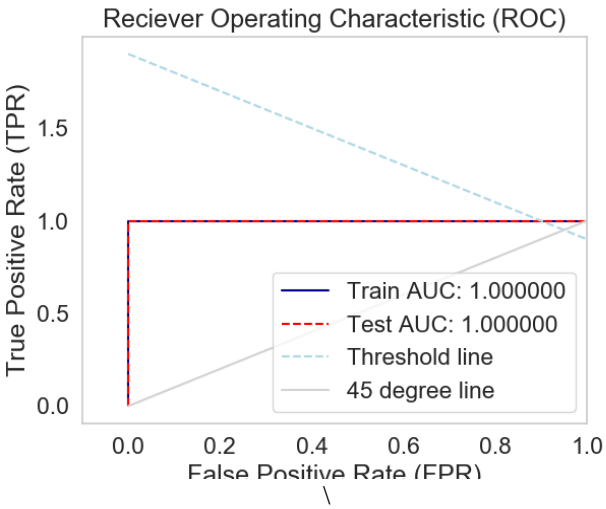

Fig 8. ROC curve

V. CONCLUSION

Fake profile is a reality over twitter. Fake profiles are growing rapidly, as people embrace them as a common communication medium. More and more personal data are now being displayed on social network pages, resulting in enormous content for exploitation of the information. We used Machine Learning techniques in this paper to predict the weather a Twitter account is a Fake or a real person. We have done a tremendous amount of product development, along with extraction of the functionality. Our system can classify whether or not a twitter user is a fake. We will expand our research to include other social media sites, such as Twitter.

REFERENCES

1. J. Sutton, L. Palen, and I. Shlovski, "Back-Channels on the Front Lines: Emerging Use of Social Media in the 2007 Southern California Wildfires," Proc. Int'l ISCRAM Conf., May 2008.
2. A.L. Hughes and L. Palen, "Twitter Adoption and Use in Mass Convergence and Emergency Events," Proc. Sixth Int'l ISCRAM Conf., May 2009.
3. S. Gianvecchio, M. Xie, Z. Wu, and H. Wang, "Measurement and Classification of Humans and Bots in Internet Chat," Proc. 17th USENIX Security Symp, 2008.
4. B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your Botnet Is My Botnet: Analysis of a Botnet Takeover," Proc. 16th ACM Conf. Computer and Comm. Security, 2009.
5. S. Gianvecchio, Z. Wu, M. Xie, and H. Wang, "Battle of Botcraft: Fighting Bots in Online Games with Human Observational Proofs," Proc. 16th ACM Conf. Computer and Comm. Security, 2009.
6. A. Java, X. Song, T. Finin, and B. Tseng, "Why We Twitter: Understanding Microblogging Usage and Communities," Proc. Ninth WebKDD and First SNA-KDD Workshop Web Mining and Social Network Analysis, 2007.
7. B. Krishnamurthy, P. Gill, and M. Arlitt, "A Few Chirps about Twitter," Proc. First Workshop Online Social Networks, 2008.
8. A. Mislove, M. Marcon, K.P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and Analysis of Online Social Networks," Proc. Seventh ACM SIGCOMM Conf. Internet Measurement, 2007..
9. S. Yardi, D. Romero, G. Schoenebeck, and D. Boyd, "Detecting Spam in a Twitter Network," First Monday, vol. 15, no. 1, Jan. 2010.
10. "Barack Obama Uses Twitter in 2008 Presidential Campaign,"http://twitter.com/Obama/, Dec. 2009.

AUTHORS PROFILE

Ms. Sai Poojitha Bommadevara, Contemplating M.Tech, Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijayawada.

Mr. A.Jitendra is currently working as an Assistant Professor, Department of Computer Science & Engineering, VRSEC (Autonomous), Vijayawada, Andhra Pradesh. He has 12 years of teaching experience. His areas of interest include Object Oriented Programming.

Mr. S.Babu is currently working as an Assistant Professor, Department of Computer Science & Engineering, VRSEC (Autonomous), Vijayawada, Andhra Pradesh. He has 11 years of teaching experience. His areas of interest include Database Security and Big Data Analytics.