Fake Profile Detection on Twitter using SVM Classifier

V. Aravindh Pashwan, D. S. Ravi

***ABSTRACT:** Artificial intelligence or machine intelligence were process or actions performed by machine on its own by using the technique "problem solving" and "learning". Cyber threat is very serios problem in this modern and technology era which lead to huge lose and threat to the human asset and their personal information. The most of human information are gathered and misused from online social media networks by using fake profiles. The aim of the study is to detect the fake user on online social media network – TWITTER, utilizing machine learning algorithm (SVM – Support Vector Machine). This proposed work helps to identify the fake user using twitter user profile attributes, with the aim of improve the security related to OSN- Online Social Media Platforms and achieved the accuracy of 97.33%. In the near future a work may extended by considering many other attributes and implemented through various algorithm to improve in finding the fake users with more accuracy.*

***Key Words** – Fake user, Machine Learning, Online Social Network, Artificial Intelligence.*

I.INTRODUCTION

The integration of Artificial Intelligence into security systems can be of use in reducing the constantly increasing threats of cyber security that is faced by the global businesses. All over the industries the applications using Machine learning along with artificial intelligence (AI) are widely used for data collection, storage capabilities and computing power which are constantly increasing[1]. In this current technical world people are count on Social media networks such as Facebook , twitter and Instagram etc, they are used to communicate with each other ease and share information, point of views, their knowledges and their personal activities, these conditions were used by spammers to exploit the users and their information [2],The spammers prime target was a social media network because they can collect information about large amount of audience in a single platform, In the year 2019 around 3.2 billion people are using social media platforms (Facebook, twitter, Instagram, LinkedIn etc.,) were that equals 42% of world population. According to Facebook there are 2.3 billion monthly active users and it estimate that 5% Accounts were fake, Twitter is one of the largest online social media networks with 336 million monthly active users on the first quarter of 2018, From the normal person to worlds biggest celebrities are using twitter, where people find their interests and connect with them.Twitter had shut downed 70 million fake accounts according to Washington Post.The social media platform has a huge effect on what read and seen, The American gets more than two third of their news through social media according to pew research center. Most of the people go for search engines rather than the traditional media for information and the news.

Manuscript received on 13 January 2021 | Revised Manuscript received on 07 February 2021 | Manuscript Accepted on 15 February 2021 | Manuscript published on 28 February 2021.

* Correspondence Author
V. Aravindh Pashwan, M. Phil. Scholar, Dept. of Computer Science St. Joseph's College, Tiruchirappalli, Tamil Nadu, India.
D. S. Ravi, Associate Professor Dept. of Computer Science St. Joseph's College, Tiruchirappalli, Tamil Nadu, India.

There are several type of fake users are in online social media platforms, each of the fake users intention was vary from one another such as, The user only wants to spread fake news among the online social media networks, where another one wants to promote the product, person, or an organization with false identity, The person who creates the duplicate profile in the name of another person, The person who connects with random people and communicate with them using false identity.

II.RELATED WORKS

The proposed anovel technique to differentiate the fake accounts on online social networksfrom the real ones.. The fake accounts typical patterns are automatically characterize were the method utilize the knowledge automatically extracted from big data. The random forest algorithm were used,The method theoretically evaluated the proposed method onthe online social network Twitter, and scorednotable results interms of discrimination capacity [3]. Proposed an advanced framework to detect a tweet with fake news were content uses the method which includes analysis of twitter account, cross verification on of fake news origins, reverse image searching and the data mining technique, the tweets were collected using the twitter API. Factfinding results on a huge heterogeneous event dataset gives the practical exhibition of the potentof proposed work on finding fake tweets. The four major models experiment runner, feature set creator, tweet content fetcher, report generator and classifier were used [4]. The Presented an approach on detecting pathogenic online social media networks without network structure or content, cascade path information,and user's information. Themethodacquire higher accuracy (0.75) in detecting Pathogenic Social Media accounts in collation with random (accuracy of 0.11) andexisting bot detection (accuracy of 0.16) methods. Algorithm for Threshold-based problems and Label propagation algorithms were used [5].Fake user detection on online social networks using machine learning. The dataset used here were manually generated by CRESCI, The machine learning algorithm used to detect the account created by bots and the cyborgs for fake users. The model were able to differentiate the fake accounts that were created by humans and the bots [6].System to identify the fake user and fake news on twitter, The system used a NoSQL database and MongoDB for unstructured information of users and tweets. The architecture used here is service-oriented architecture,The four main modules used were collector module, database module, core module and analysing module, these modules were designed to interact with service-oriented architecture [7].

To find the Twitter fake accounts profile characteristics, Pattern matching algorithm were used to find the combinations among the screen-names and an analysis of the update times, were identified the fake user accounts. Profile-Based approach had been derived as a efficient way to identify the fake profiles in time-efficient manner [8].Finding fake retweets activity in twitter,The paper described that the organic tweets were behaviour were more variability then the fraudulent, were fraudulent is more synchronized than the organic tweets, 12 million real retweet dataset were used for the experiments that are crawled from the twitter and achieved accuracy of 97%. [9].

III.SUPPORT VECTOR MACHINE

SVM – Support Vector Machine perform the classification method by using the Hyperplane, Hyperplane as well as line which enlarges the margin between the two classes by separating. The hyperplane finds the closest points from the both classes, that points are said to be support vectors. The distance between the classes are called as a margin, SVM consider the dataset into two types called linearly separable and not linearly separable.The dataset we used here is a linearly separable data. The parameter C in SVM defines how enough we want to keep away the miscategorized training set. The maximum value of C will select smaller-margin hyperplane, if hyperplane provide a finer place of gaining every training point arrange admirably. The minimum value of C will go for larger-margin fractionation hyperplane, if the hyperplane miscategorized many points, even if the training date is linearly separable.

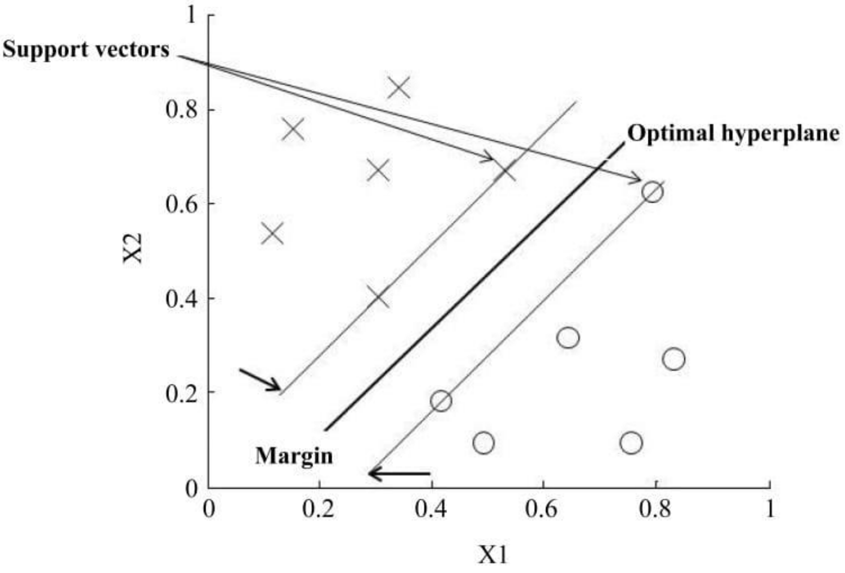


Figure 1: SVM example for proposed work.

IV.METHODOLOGY

The Proposed model used to find the fake users

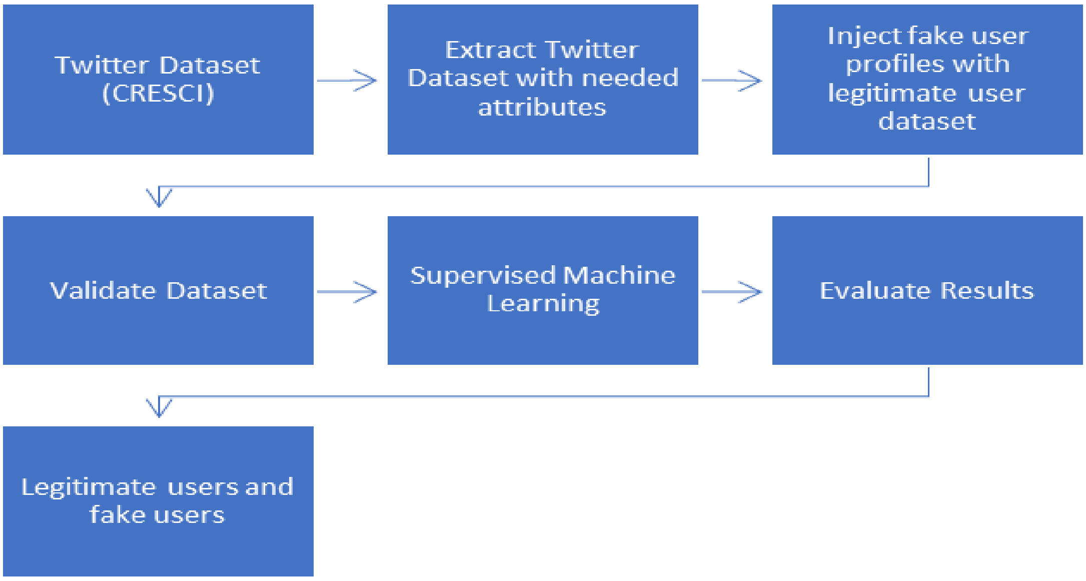

Twitter Dataset (CRESCI)
Extract Twitter Dataset with needed attributes
Inject fake user profiles with legitimate user dataset
Validate Dataset
Supervised Machine Learning
Evaluate Results
Legitimate users and fake users

Figure 2: Work flow of proposed methodology

Steps involved in Methodology to find the fake users.

Step 1: The twitter dataset used here was manually generated by Cresci et al [10]. A new dataset cannot be manually generated from any online social networks for identifying fake user on twitter.

Step 2: Form the dataset, the necessary attributes such as id_name, screen_name, date_of_joining, tweet frequency, followers_count (followers), friends_count (following), favurites_count (liked posts), language, geo_location(location), verified and tweetswere extracted.

Step 3: In the legitimate user dataset, the fake user dataset was injected in which the attributes are similar to the legitimate users but the values on the attributes are different from the legitimate users.

Step 4: The next step was validating the dataset, the data set with the same attributes from the real user and fake user dataset was extracted ,since the R studio did not support characters or string it only accepts the numerical values, so the attributes which are having only the numerical values are extracted from the above dataset such as followers_count, friends_count, favorites, tweet, tweet_frequency, geo_enabled and verified.

Step 5: After validating the dataset, the supervised machine learning technique was used to classify the users based on SVM (Support Vector Machine) classification algorithm.

Step 6: Evaluating the results based on the given attributes and the condition, the fake users and legitimate users are classified separately.

Attributes used in proposed work with descriptions

Table 1: attributes used in proposed work

Attribute	Description
FOLLOWERS_COUNT	No of people follows the account holder
FRIENDS_COUNT	Mutual users (Friends who follows back the account holder)
FAVURITES_COUNT	The number tweets liked/favorited by the user
TWEET	Total number of tweets made by a user
TWEET_FREQUENCY	The no of tweets by a user from the date of creation to till date
GEO_ENABLED	Whether the account holder provided their location
VERIFIED	Whether the account is verified or not

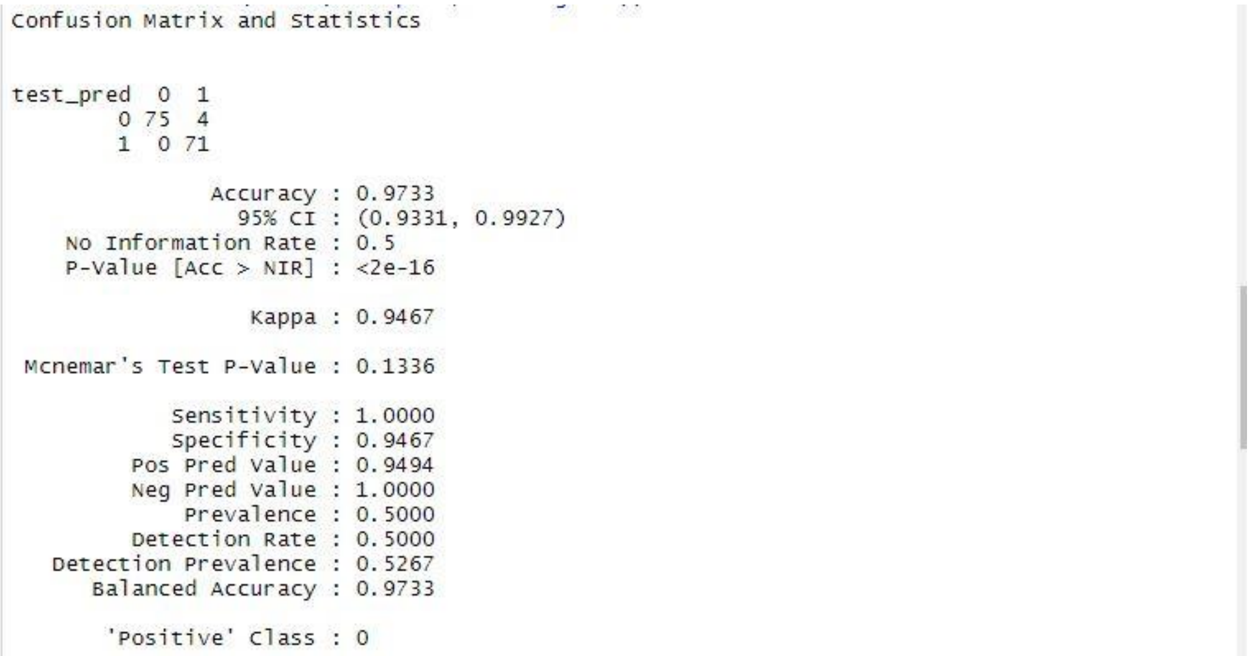
The work progress for fake user identification in online social media networks are given in the fig.1. In the first step the data collection is performed were we used the cresci et.al. dataset for the work, In the initial step the data collection is performed and it extracted the needed attributes from the dataset, where the R Studio only supports numerical values and we extracted the attributes which only contains the numerical values such asfollowers_count, friends_count, favorites, tweet, tweet_frequency, geo_location and verified. Then the fake user dataset is injected into the legitimate user dataset, and validate the dataset. For the validation the injected dataset is import into R studio, here the 7 attributes were trained and tested with R studio using the machine learning algorithm SVM (Support Vector Machine) for classification, and evaluate the results.

V.PRE- PROCESSING

Data Pre-Processing is important step in Data mining process. Data from Real world are habitually inconsistent or incomplete may have missing and null values, and out of range values, etc. Data preprocessing is used to resolve these kinds of errors. Data cleaning, Data Integration, Data Transformation, Data Reduction and Data Discretization are the steps involved in Data Preprocessing. Non-Removal of these kind of values may lead to the inaccuracy in resulting. The result after the data pre-processing would generate complete result with accuracy.

VI.FINDING FAKE USERS

From the analysis made from the various papers there are several machine learning classification algorithms were used to classify fake users from legitimate users, and some the study shows that classify the legitimate users form the accounts that were generated by bots, some of the techniques were classify the users using clustering method (Random Forest), were most of the accurate results were generated using machine learning algorithms. In this work Support vector Machine is used and the collected data attributes were applied in RStudio to identify the fake users and legitimate users based on followers_count, friends_count, favorites, tweet, tweet_frequency, geo_location and verified. The target value had pre-defined as res in Dataset with the class variable 0 and 1. The end result of the work is represented in class variable 1 or 0, where 0 denotes legitimate users and 1 denotes the fake users, After the Repeated Cross-Validation Cost c denotes the hyperplane, the maximum hyperplane of 5 is selected andresulted in accuracy of 97.33%.



```
Confusion Matrix and Statistics

test_pred  0  1
        0 75  4
        1  0 71

               Accuracy : 0.9733
                 95% CI : (0.9331, 0.9927)
    No Information Rate : 0.5
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.9467

 Mcnemar's Test P-Value : 0.1336

            Sensitivity : 1.0000
            Specificity : 0.9467
         Pos Pred Value : 0.9494
         Neg Pred Value : 1.0000
             Prevalence : 0.5000
         Detection Rate : 0.5000
   Detection Prevalence : 0.5267
      Balanced Accuracy : 0.9733

       'Positive' Class : 0
```

Figure 3: confusion matrix of proposed work

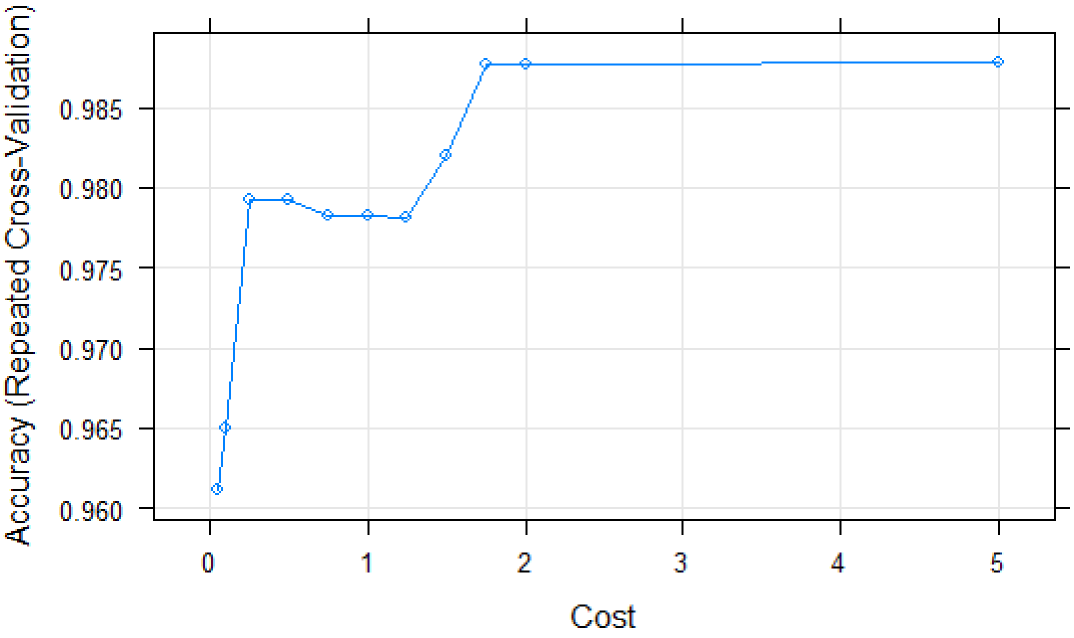


Figure 4: Accuracy plot of proposed system

VII.CONCLUSION

In artificial intelligence, machine learning methods are applied in order to nataurally improve and learn from the experience without being obviously programmed. This work provides the view about the fake user identification on twitter using classification algorithm. It helps to identify the fake users on twitter by using machine learning techniques. Fake users are increasing day by day, huge amount of private and personal data is shared using online social networks which are dangerous if disclosed to unknown people may cause potential risk. The collision of machine learning techniques had been used as effective mechanism to counter fake profiles.

REFERENCES

1. Aravindh Pashwan V and Ravi D S, “A Survey On Fake Account Detection In Social Media Using AI”, IJRAR- International Journal of Research and Analytical Reviews (IJRAR), Vol. 6, Issue 2, 2019, pp. 74-78.
2. Akash Kumbhar, Meghana Wable, Supriya Nigade and Komal Darekar, "A survey on: Malicious Application and Fake user Detection in Facebook using Data Mining", International Journal of Engineering Science 15768, 2017.
3. Loredana Caruccio, Domenico Desiato, and Giuseppe Polese, "Fake Account Identification in Social Networks", In 2018 IEEE International Conference on Big Data (Big Data), IEEE, 2018, pp. 5078-5085.

4. Saranya Krishnan and Min Chen, "Identifying Tweets with Fake News", In 2018 IEEE International Conference on Information Reuse and Integration (IRI), IEEE, 2018, pp. 460-464.
5. Elham Shaabani, Ruocheng Guo and Paulo Shakarian, "Detecting pathogenic social media accounts without content or network structure", In 2018 1st International Conference on Data Intelligence and Security (ICDIS), IEEE, 2018, pp. 57-64.
6. Naman Singh, Tushar Sharma, Abha Thakral and Tanupriya Choudhury, "Detection of Fake Profile in Online Social Networks Using Machine Learning", In 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), IEEE, 2018, pp. 231-234.
7. Costel-Sergiu Atodersei, Alexandru Tanaselea and Adrian Iftene, , "Identifying Fake News and Fake Users on Twitter", In 2018 International Conference on Knowledge Based and Intelligent Information and Engineering Systems, ELSEVIER, 2018, pp. 451-461.
8. Supraja Gurajala, Joshua S White, Brian Hudson, Brian R Voter and Jeanna N Matthews, "Profile characteristics of fake Twitter accounts", In 2016 Big Data and Society, 2016, pp. 1-13.
9. Maria Giatsoglou, Despoina Chatzakou, Neil Shahy, Alex Beutely, Christos Faloutsosy and Athena Vakali , "Spotting Fake Retweeting Activity in Twitter".
10. Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi and Maurizio Tesconi, "Fame for sale: Efficient detection of fake Twitter followers", In 2015 Decision support systems, ELSEVIER, 2015, pp. 56-71.

.