ISSN [Online]: 2583-2654

Detecting Fake Bots on Twitter using SVM and Neural Networks Algorithms

Nisha bai M[1], Priyadharshini G[2], Ramya R[3], Sonu Roshini N[4],*
Dept of Computer Science and Engineering,
Dr. T Thimmaiah Institute of Technology,
Kolar Gold Field, Karnataka, India

Abstract: In the last years the most used social networks are like facebook or Twitter acknowledges that on their network are counterfeit and duplicate accounts. These forged accounts can manipulate and lead the social media in wrong way for example, create hateful links within the posts/tweets. We can make use of several new features, For detecting Twitter accounts, which are more effective and robust than current features, so (We evaluated the proposed set of features by exploiting very popular machine learning classification algorithms, namely Support Vector Machine (SVM) and Neural Networks (NN). Now day's social networks have been part of many people's lives. Many activities such as communication, promotion, advertisement, news, agenda creation have started to be done through social networks. Some malicious accounts on Twitter are used for purposes such as misinformation and agenda creation. This is one of the basic problem in social networks. Therefore, detection of malicious account is significant. In this study, machine learning-based methods were used to detect fake accounts that could mislead people. For this purpose, the dataset generated was pre-processed and fake accounts were determined by machine algorithms are used for the detection of fake accounts. Classification performances of these methods are compared and the logistic regression proved to be more successful than the others.

Keywords: Twitter accounts, Support Vector Machine, Neural Network, Logistic Regression, Machine Learning, Fake Account.

I. INTRODUCTION

Twitter is an online news social networking services where users post and interact with messages. Tweets restricted to 140 characters. There are three 10m monthly active twitter users and a total of 1.3 billion accounts have been created. There are about 500 million tweets sent per day. Since twitter can be such an influential platform .There exists twitter bots these bots can be used in a variety of ways such as for increasing number of followers, spamming, re-tweeting etc. This practice is becoming so popular that it is estimated that there are about 48 million twitter bots already. It is sometimes important to identify when a twitter account is controlled by a bot. detecting non-human twitter users has been of interest to academics. One significant academic study estimated that up to 15% of twitter users were automated bot accounts. The prevalence of twitter bots coupled with the ability of some bots to give seemingly human responses has enabled these non-human accounts to garner widespread influence.
The main aim of this project is to distinguish twitter bots from real accounts:

Machine learning: Machine learning is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a model based on sample data known as training data to make predictions or decisions without being explicitly programmed to do so. A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers but not all machine learning is statistical learning.

ISSN [Online]: 2583-2654

Support vector machines: Support vector machines (SVM's) are powerful yet flexible supervised machine learning algorithms which are used both for classification and regression. But generally they are used in classification problems. SVM have their unique way of implementation as compared to other machine learning algorithms. It can solve linear and non linear problems and work well for many practical problems. The idea of SVM is simple the algorithm creates a line or a hyper plane which separates the data into classes.

Neural networks: A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way human brains operates. Neural networks refer to system of neurons either organic or artificial in nature. Neural networks can adapt to changing inputs so the network generates the best possible result without needing to redesign the output criteria.

Fake twitter bots: A twitter bot is a type of bot software that controls a twitter account via the twitter API. The bot software may autonomously perform actions such as tweeting, re-tweeting, liking, following, un-following, or direct messaging other accounts. Proper usage includes broadcasting helpful information, automatically generating interesting or creative content, and automatically replying to users via direct message. Improper usage includes circumventing API rate limits, violating user privacy, spamming.

II. RELATED WORK

A literature review is brief summary of previous research on a topic [10,31,33]. The literature review surveys scholarly articles and other resources relevant to a particular area of research. The work should enumerate, summarize and objectively evaluate the previous research. The author R.Kaur and S.Singh et.al [1] reported, a review of number of data mining approaches used to detect anomalies. An uncommon directions made to the investigation of informal community z creator demonstrated the relationship between name worth inclination and the quantity of adherents. Accordingly, it is hard to gauge the validity of a client in these systems and to check his/her posts. As online interpersonal organizations have turned out to be progressively valuable for dispersing data to more extensive crowds, tending to the previously mentioned difficulties to decide the validity of clients in OSNs requires the advancement of hearty methods for estimating client and substance believe ability. They have investigate the nature of spam users on Twitter with the goal to improve existing spam detection mechanisms. For detecting Twitter spammers, they have used several new features, which are more effective and robust than existing used features. Based on the anticipated results from the machine learning models, it seems that existing features and machine learning models used to detect fake accounts are not enough to detect fake social networking accounts. We had made an analysis of the tergiversation idea that was used by Twitter spammers. They observed that Twitter spammer used to change their performance to evade spam detection techniques, so they suggested designing new features that would enhance detecting spammers and would be harder for them to evade. They had combined their new features in machine learning classifier algorithm and compared the implementation with other existing methods. Manuel Egele., Gianluca Stringhini., Christopher Krugel., [6] published "Towards detecting compromised accounts on social networks. Exchanges on reliable and security processing. Compromising social network accounts has become a profitable course of action for cybercriminals. By hijacking control of a popular media or business account, attackers can distribute their malicious messages or disseminate fake information to a large user base. The impacts of these incidents range from a tarnished reputation to multi-billion dollar monetary losses on financial markets. In our previous work, we demonstrated how we can detect large-scale compromises (i.e., so-called campaigns) of regular online social

ISSN [Online]: 2583-2654

network users. In this work, we show how we can use similar techniques to identify compromises of individual high-profile accounts. High-profile accounts frequently have one characteristic that makes this detection reliable-they show consistent behavior over time. We show that our system, were it deployed, would have been able to detect and prevent three real-world attacks against popular companies and news agencies. Furthermore, our system, in contrast to popular media, would not have fallen for a staged compromise instigated by a US restaurant chain for publicity reasons. Naman Singh., Tushar Sharma., Abha Thakral., Tanupriya Choudhary., [7] published "Detection of fake profile in online social networks using machine learning. " in international conference on advances in computing and communication engineering ". In today's world, the social media platforms are being used on daily basis and has become an important part of our lives. The number of peoples on social media platforms are incrementing at a greater level for malicious use. There are numerous cases where produced accounts have been effectively distinguished utilizing machine adapting techniques however the amount of research work is very low to recognize counterfeit characters made by people. For bots the ML models used various features to calculate the no. of followers to the no. of friends that an account has on social media platforms (SOCIAL MEDIA PLATFORMSs). The no. of friends to the no. of followers of any account are easily available in the account profiles and no rights are violated of any accounts. In order to accomplish the task of detecting, identifying and eliminate the fake accounts we establish a forged human account. ESTEE VAN DER WALF., and JAN ELOFF., [8] published "Using machine learning to detect fake identities: Bots vs Humans. (2017) There are a growing number of people who hold accounts on social media platforms (SMPs) but hide their identity for malicious purposes. Unfortunately, very little research has been done to date to detect fake identities created by humans, especially so on SMPs. In contrast, many examples exist of cases where fake accounts created by bots or computers have been detected successfully using machine learning models. In the case of bots these machine learning models were dependent on employing engineered features, such as the “friend-to-followers ratio these features were engineered from attributes, such as “friend-count” and “follower-count, which are directly available in the account profiles on SMPs. The research discussed in this paper applies these same engineered features to a set of fake human accounts in the hope of advancing the successful detection of fake identities created by humans on SMPs. Sarah Khaleed., Hoda .M. O . Mokhtar., Neamat El- Tazi .,[9] published “Detecting fake accounts on social media. “ in international conference on bigdata on international conference on IEEE.(2018) In the present generation, online social networks (OSNs) have become increasingly popular, people’s social lives has become more associated with these sites. They use OSNs to keep in touch with each others, share news, organize events, and even run their own e-business. The rapid growth of OSNs and the massive amount of personal data of its subscribers have attracted attackers, and imposters to steal personal data, share false news, and spread malicious activities. On the other hand researchers have started to investigate an efficient techniques to detect abnormal activities and fake accounts relying on accounts features, and classification algorithms. However, some of the account’s exploited features have negative contribution in the final results or have no impact, also using stand alone classification algorithms does not always reach satisfied results. In this paper, a new algorithm, SVM-NN, is proposed to provide efficient detection for fake Twitter accounts and bots, four feature selection and dimension reduction techniques were applied. Three machine learning classification algorithms were used to decide the target accounts identity real or fake, those algorithms were support vector machine (SVM), neural Network (NN), and the newly developed algorithm, SVM-NN, that uses less number of features, while still being able to correctly classify about 98% of the accounts of our

ISSN [Online]: 2583-2654

training dataset. Loredana Caruccio., Domenico Desiato., Giuseppe Polase., [10] published "Fake account identification in social media networks. In international conference on big data ,international conference in IEEE.(2018) Nowadays, the human influence often depends on the number of followers that an individual has in his/her own social media context. To this end, the presence of fake accounts is one of the most relevant problems and can potentially have a big impact on many real life and business activities. Fake followers are dangerous for social platforms, since they may alter concepts like popularity and influence, which might yield a strong impact on economy, politics, and society. Thus, it is necessary to devise new methodologies enabling the possibility to identify and characterize fake accounts. This work presents a novel technique to discriminate real accounts on social networks from fake ones. The technique exploits knowledge automatically extracted from big data to characterize typical patterns of fake accounts. We empirically evaluated the proposed technique on the Twitter social network, and achieved significant results in terms of discrimination capabilities K.Jino Abidha., Roshan Nilofer., A. Silviya., Dr.S.Raja ratna., [11] published "Detection of Twitter 's spam using machine learning algorithms. " (2019). Twitter platform is taken and spam tweets detection is performed. To stop spammers, semi supervised learning is used to detect spam tweets in twitter. Thus, industries and researchers have applied different approaches to make spam free social network platform. Some of them are only based on user-based features while others are based on tweet based features only. To solve this issue, a framework has been proposed which takes the user and tweet based features along with the tweet text feature to classify the tweets. The benefit of using tweet text feature is that the spam tweets can be identified even if the spammer creates a new account which was not possible only with the user and tweet based features. The work has been evaluated with three different machine learning algorithms namely - Support Vector Machine, Neural Network, Random Forest. With Naive Bayes classifier, about 80% of accuracy is obtained. Akshatha T.M., Dr.Veena.,[12] published "Machine learning framework for detecting spammer and fake users on Twitter. With the evolution of OSNs, the need to study and analyze users' behaviors in online social platforms has intensity Spammers can be identified based on: (i) fake content, (ii) URL based spam detection, (iii) spam in trending topics, and (iv) fake user identification. And with the help of machine learning algorithms we are going to identify the fake user and spammer in twitter. Gayatri.A., Radhika.S., Mrs.Jayalakshmi.S.L. [13] published Detecting fake accounts on media applications using machine learning. The social network, a crucial part of our life is plagued by online impersonation and fake accounts .Fake profiles are mostly used by the intruders to carry out malicious activities such as harming person , identity theft and privacy intrusion in Online Social Network(OSN).Hence identifying an account is genuine or fake is one of the critical problem in OSN .In this paper we proposed many classification algorithm like Support Vector Machine algorithm and deep neural network .It also studies the comparison of classification methods on Spam User dataset which is used to select the best. Ashraf Khalil., Hassan Hajdiab., and Nadeeb Al-Quirim.

III. PROPOSED KBDC APPROACH

In this following section, detail proceedings of the proposed In this paper, our aim is to use Machine learning classification algorithms to decide the accounts identity as real or fake, those algorithms were support vector machine, neural Network, and our newly developed algorithm, SVM-NN. The proposed algorithm (SVM-NN) uses less number of features, while its still being able to correctly classify about of the accounts of our training dataset shows the proposed system architecture of Twitter fake account detection. The input traffic data is used for twitter dataset with some features. The training dataset contains data preprocessing which includes two steps:

ISSN [Online]: 2583-2654

Feature Extraction and Machine learning technique. After uses these two are arranged in a model, which used for selecting number of features. After that apply the Support Vector machine for classifying our data and neural network use for training our model. After applying the algorithms it predicts whether our model account is fake or not.

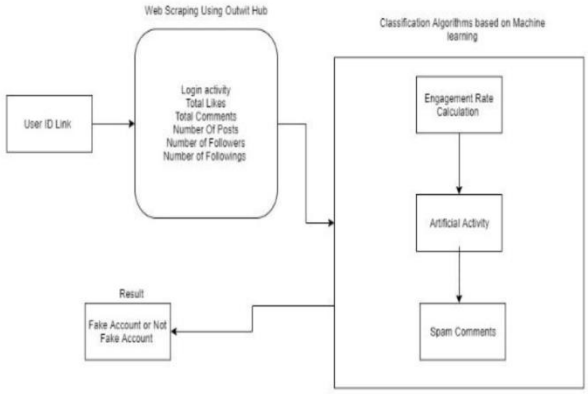

Fig 1: Architectural depiction of Proposed System.

3.1. *Web Scraping Using Outwit Hub Stage*

Web Scraping: You suppose want some information from a website? Let's say a paragraph on Donald Trump! What do you do? Well, you can copy and paste the information from Wikipedia to your own file. But what if you want to obtain large amounts of information from a website as quickly as possible? Such as large amounts of data from a website to train a Machine Learning algorithm? In such a situation, copy and paste is not going to work! And that's when you'll need to use Web Scraping. Unlike the long and mind-numbing process of manually obtaining data, uses intelligence automation methods to get thousands or even millions of data sets in a smaller amount of time. So let's understand what Web Scraping is in detail and how to use it to obtain other websites.

WHY Web Scraping: Web Scripting is an automatic method to obtain large amounts of data from websites. Most of this data is unstructured data in an HTML format which is then converted into structured data in a spreadsheet or a database so that it can be used in various applications. There are many different ways to perform web scraping to obtain data from websites. These include using online services, particular API's or even creating your code for web scraping from scratch. Many large websites like Google, Twitter, Facebook, Stack Overflow, etc. have API's that allow you to access their data in a structured format. This is the best option but there are other sites that don't allow users to access large amounts of data in a structured form or they are simply not that technologically advanced. In that situation, it's best to use Web Scraping to scrape the website for data. Web scraping requires two parts namely the **crawler** and the **scraper**. The crawler is an artificial intelligence algorithm that browses the web to search the particular data required by following the links across the internet. The scraper, on the other hand, is a specific tool created to extract the data from the website. The design of the scraper can vary greatly according to the complexity and scope of the project so that it can quickly and accurately extract the data.

How Web Scrapers Work: Web Scrapers can extract all the data on particular sites or the specific data that a user wants. Ideally, its best if you specify the data you want so that the web scraper only extracts that data quickly. For example, you might want to scrape an Amazon page for the types of juicers available, but you might only want the data about the models of different juicers and not the customer reviews. So when a web scraper needs to scrape a site, first it is provided the URL's of the required sites. Then it loads all the HTML code for those sites and a more advanced scraper might even extract all the CSS and JavaScript elements as well. Then the scraper obtains the required data from this HTML code and outputs this data in the format specified by the user. Mostly, this is in the form of an Excel spreadsheet or a CSV file but the data can also be saved in other formats such as a JSON file.

ISSN [Online]: 2583-2654

Different Types of Web Scrapers: Web Scrapers can be divided on the basis of many different criteria including Self-built or Pre-built Web Scrapers, Browser extension or Software Web Scrapers, and Cloud or Local Web Scrapers. You can have **Self-built Web Scrapers** but that requires advanced knowledge of programming. And if you want more features in your Web Scraper, then you need even more knowledge. On the other hand, **Pre-built Web Scrapers** are previously created scrapers that you can download and run easily. These also have more advanced options that you can customize. **Browser extension Web Scrapers** are extensions that can be added to your browser. These are easy to run as they are integrated with your browser but at the same time, they are also limited because of this. Any advanced features that are outside the scope of your browser are impossible to run on Browser extension Web Scrapers. But **Software Web Scrapers** don't have these limitations as they can be downloaded and installed on your computer. These are more complex than Browser extension Web Scrapers but they also have advanced features that are not limited by the scope of your browser. **Cloud Web Scrapers** run on the cloud which is an off-site server mostly provided by the company that you buy the scraper from. These allow your computer to focus on other tasks as the computer resources are not required to scrape data from websites. **Local Web Scrapers**, on the other hand, run on your computer using local resources. So if the Web Scrapers require more CPU or RAM, then your computer will become slow and not be able to perform other tasks. **What is Web Scraping used for:** Price Monitoring Web Scraping can be used by companies to scrap the product data for their products and competing products as well to see how it impacts their pricing strategies. Companies can use this data to fix the optimal pricing for their products so that they can obtain maximum revenue. Market Research Web scraping can be used for market research by companies. High-quality web scraped data obtained in large volumes can be very helpful for companies in analyzing consumer trends and understand which direction the company should move in the future.

News Monitoring Web scraping the news sites can provide detailed reports on the current news to a company. This is even more essential for companies that are frequently in the news or that depend on daily news for its day to day functioning. After all, news reports can make or break a company in a single day. Sentiment Analysis if companies want to understand the general sentiment for their products among their consumers, then Sentiment Analysis is a must. Companies can use web scraping to collect data from social media websites such as Facebook and Twitter as to what the general sentiment about their products is. This will help them in creating products that people desire and moving ahead of their competition. Email Marketing Companies can also use Web scraping for Email marketing. They can collect Email ID's from various sites using web scraping and then send bulk promotional and marketing Emails to all the people owning these Email ID's.

3.2. Classification Of Algorithims Based Machine Learning Stage:

Calculation of Engagement Rate: Engagement rate is a **formula that measures the amount of interaction social content earns relative to reach or other audience figures**. Think reactions, comments, and shares. There are multiple ways to measure this engagement, and different calculations may better suit your social media objectives

So, it's time to do the math. Add these formulas to your social media toolkit so you can be sure you're using the right equation in the correct context.

Engagement rate by reach (ERR)

This formula is the most common way to calculate engagement with content.

ERR measures the percentage of people who chose to interact with your content after seeing it. Use the

ISSN [Online]: 2583-2654

first formula for a single post, and the second one to calculate the average rate across multiple posts.

- **ERR = total engagements per post / reach per post * 100**
  To determine the average, add up the all the ERRs from the posts you want to average, and divide by number of posts:
- **Average ERR = Total ERR / Total posts**
  *In other words: Post 1 (3.4%) + Post 2 (3.5%) / 2 = 3.45***Pros**: Reach can be a more accurate measurement than follower count, since not all your followers will see all your content. And non-followers may have been exposed to your posts through shares, hash tags, and other means.

Cons: Reach can fluctuate for a variety of reasons, making it a different variable to control. A very low reach can lead to a disproportionately high engagement rate, and vice versa, so be sure to keep this in mind.

Machine learning Fake account detection: In this step, we combine all the data we extracted from the website. In this paper we mainly focus on engagement rate, artificial activity and spam comments. The data collected using web scraper is used to compute the values for the factors mentioned above. Using these factors different decision tress is formed. Using gradient boosting algorithm and with the formed decision trees fake accounts are detected.
Decision Trees: Decision trees are made seeing the success rate i.e., in our case taking the value which contains more fake accounts. The first tree is made using an engagement rate as the root node and artificial activity as its child node along with spam comments as another node. The second tree is made keeping artificial activity as the root node, engagement rate and spam comments as subsequent nodes. The third tree is formed using spam comments as the root node, artificial activity and engagement rate as subsequent nodes.

3.2.1. Branching out data points using the residual values

We use a limit of two leaves here to simplify our example, but in reality, Gradient Boost has a range between **8 leaves to 32 leaves**. Because of the limit on leaves, one leaf can have multiple values. Predictions are in terms of log(odds) but these leaves are derived from probability which cause disparity. So, we can't just add the single leaf we got earlier and this tree to get new predictions because they're derived from different sources. We have to use some kind of transformation. The most common form of transformation used in Gradient Boost for Classification is :

$$\frac{\sum Residual}{\sum[PreviousProb * (1 - PreviousProb)]}$$

The numerator in this equation is sum of residuals in that particular leaf.
The denominator is sum of (previous prediction probability for each residual) * (1 - same previous prediction probability).
The derivation of this formula shall be explained in the Mathematical section of this article.
For now, let us put the formula into practice:
The first leaf has only one residual value that is 0.3, and since this is the first tree, the previous probability will be the value from the initial leaf, thus, same for all residuals. Hence,

$$\frac{0.3}{[0.7 * (1 - 0.7)]} = 1.43$$

For the second leaf,

$$\frac{0.3 + 0.3}{[0.7 * (1 - 0.7)] + [0.7 * (1 - 0.7)]} = 4.29$$

Similarly, for the last leaf:

$$\frac{-0.7 + 0.3 + 0.3}{[0.7 * (1 - 0.7)] + [0.7 * (1 - 0.7)] + [0.7 * (1 - 0.7)]} = -0.16$$

Now the transformed tree looks like:

ISSN [Online]: 2583-2654

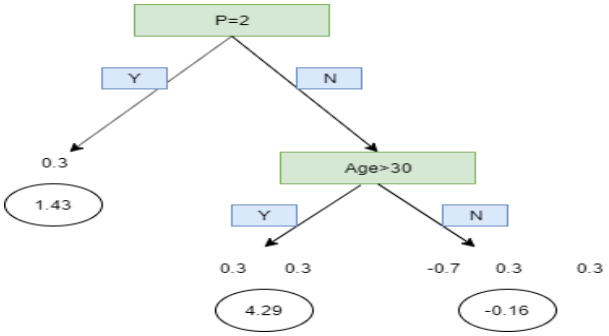

FIG 2: TRANSFORMED TREE.

Transformed tree now that we have transformed it, we can add our initial lead with our new tree with a learning rate.

$$OldTree + LearningRate * NewTree$$

3.2.2. Algorithm

Input: training set $\{i, \ \}=1n$ a differentiable loss function L(y,F(x)), number of iterations *M*.

Algorithm:

I. Initialize model with a constant value:

$F0\ x = argminLyi, .n\ i=1$

II. For m = 1 to M:

1.Compute so-called *pseudo-residuals*:

$rim=-[\partial(L(yi,Fxi\ )\partial(F(xi)]Fx = Fm-1(x)$

For i =1, . . . , n.

2.Fit a base learner (e.g.tree) $h(x)$ to pseudo-residuals, i.e. train it using the training set $\{xi,yi\}$ $i=1n$.

3.Compute multiplier γm by solving the following one-dimensional optimization problem:

$\gamma m = argminLyi, m-1\ xi + \gamma h(xi))\ .ni=1$

4.Update the model:

$Fmx = Fm-1\ x + \gamma mhm.$

III. Output Fm.

IV CONCLUSION

In this paper, we have maintained the highest accuracy in detecting fake accounts by different classifying algorithms. The results shows the increase of the accuracy results of two of the classification algorithms after using the suggested attributes with their corresponding heaviness. The classification algorithms are proposed to improve detecting fake accounts on social networks, where the SVM trained model decision values were used to train a NN model, and SVM testing decision values were used to test the NN model.

REFERENCES

[1] R.Kaur and S.Singh, "A survey of data mining and social network analysis based anomaly detection techniques", Egyptian in formatics diary,vol.17, no.2, pp.1992-216, 2016.

[2] Yubao Zhang, Xin Ruan, Haining Wang, Hui Wang, and Su He "Twitter Trends Manipulation: A First Look Inside the Security of Twitter Trending" IEEE Exchanges on Data Crime scene investigation and Security (Vol. 12 ,Issue, 1 ,Jan. 2017)

[3] ManuelEgele, Gianluca Stringhini, Christopher Kruegel, and Giovanni Vigna "Towards Detecting Compromised Accounts on Social Networks "IEEE Exchanges on Reliable and Secure Processing (Vol.14, Issue: 4, July-Aug. 1 2017)

[4] MajedAlrubaian, Muhammad Al-Qurishi, Mohammad Mehedi Hassan, and Atif Alamri, "A Credibility Analysis System for Assessing Information on Twitter", IEEE Exchanges on Reliable and Secure Processing (Vol.15, Issue. 4-July-Aug.1.2018)

[5] ESTEE VAN DER WALT and JAN ELOFF "Using Machine Learning to Detect Fake Identities: Bots vs Humans" Received December 5, 2017,acknowledged January 12, 2018, date of production January 23, 2018, date of current rendition Walk 9, 2018.

[6] MyoMyoSwe and Nyein NyeinMyo, "Fake accounts on twitter using Blacklist," in International Conference on Information System (ICIS), 2018 International Conference on. IEEE, 2018, pp. 562–566.

[7] Naman Singh, Tushar Sharma, Abha Thakral, Tanupriya Chaudhury, "Detection of Fake profile in Online Social Network using Machine Learning," in International Conference On Advances in Computing and Communication Engineering (ICACCE), 2018 International Conference on. IEEE, 2018, pp. 231-234.

[8] Ilham Aydin, Mehmet SEVİ, Mehmet Umut SALUR, "Detection of Fake Twitter Account with Machine Learning Algorithms," in International Conference on Artificial Intelligence and Data Processing (IDAP), 2018 International Conference on. IEEE, 2019.

[9] Sarah Khaleed., Hoda M.O.Mokhtar., Neamat El- Tazi "Detecting fake accounts on Social media." In International conference on the bigdata, bigdata,pp. 3672-3681,2018.

10] LoredanaCaruccio., Domenico Desiato., Giuseppe Polase., "Fake account identification in social media networks. In international conference on big data, in international conference on IEEE, pp. 5078-5085, 2018.