**Detection of a Fake Twitter accounts based on the minimally weighted range of Attributes**

**CMP7200 Individual Masters Project**

**Module Supervisor: Jagdev Bhogal**

**Module Leader: Mariam Adedoyin-Olowe**

**Logo

Description automatically generated**

***Submitted By:***

**Student Name: Naga Sai Ganesh Ronte**

**Student ID: 20110720**

**Date of Submission: 10 Jun 2021**

# Abstract:

Online Social Network (OSN) is an interconnected channel of networks for individuals with shared preferences or connections in the real world. If OSN's (Online Social Network) popularity grows, security and privacy concerns are also on the increase because the size of an audience, using that particular social media site is the basic reason of its fame and this has significant social and economic entanglement, which is complicated and exploited the true information about its population via the presence of unauthentic, unauthorized false and fake accounts. Fake and clone accounts cause social network users to complex security issues. Cloning user profiles is a significant risk if current user data are hacked to build double profiles and misused then to harm the original profile owner's identity. They may also threaten to phish, stalk & spam, etc. A fake or unauthorized account is the development of an account to conduct fraudulent acts on behalf of an individual or organization that are not actually present in social media. In this research, first the introduction of Twitter and fake Twitter accounts will be presented, In the 2nd section we will be defining the objective of our research, and then literature of fake profile detection will be reviewed and after that the need of it will be defined. We will also discuss some previous work of the researchers in this article, and in data methodology a tool for detecting fake and clone profiles will be suggested on Twitter Fake accounts are classified based on a sequence of guidelines that can distinguish false and true profiles efficiently. Two techniques will be used for profile identification—one with similarity and the other with the C.5 (five classification) algorithms for decision tree. in which we will used 5 classification algorithm that will be going to be applied on the minimized weighted attributes successfully. This is going to be held in three steps in total, in which we will be conducting experiments and forming the subsections of each as Subsection A, Subsection B and Subsection C. respectively. We will also be overviewing some other methodology and at last we will deduce the conclusion from our findings and analysis.

**Keywords**: Twitter, fake, identity theft, OSN (Online Social Network)

Acknowledgement

I would like to thank my supervisor Jagdev Bhogal for all her help and advice for sharing their expertise, providing invaluable feedbacks, and helping me to stay motivated during all the stages of dissertation. My gratitude extends to the faculty of Advanced Computer Science for their incredible lectures who helped me to complete my Master’s. My appreciation also goes out to my family and friends for their encouragement and support all through my studies.

Table of Contents

[Abstract: 2](#_Toc70512305)

[Introduction: 1](#_Toc70512306)

[The objective of the research 4](#_Toc70512307)

[Literature Review 5](#_Toc70512308)

[Why need of fake account detection? 9](#_Toc70512309)

[Previous work : 12](#_Toc70512310)

[Data Methodology: 13](#_Toc70512311)

[Dataset : 16](#_Toc70512312)

[Classification Algorithm selection step : 18](#_Toc70512313)

[Experimental Results : 19](#_Toc70512314)

[Subsection A : 19](#_Toc70512315)

[Subsection B: 19](#_Toc70512316)

[Subsection C : 19](#_Toc70512317)

[A: First Experiment 20](#_Toc70512318)

[B: 2nd Experiment : 21](#_Toc70512319)

[C: 3rd Experiment : 23](#_Toc70512320)

[An overview of other methodologies. 23](#_Toc70512321)

[Detection of Fake Profile (Honeypot/Machine Learning): 23](#_Toc70512322)

[Similarity measures for clone profile detection: 25](#_Toc70512323)

[Similarity of attributes 27](#_Toc70512324)

[The similarity of the network 27](#_Toc70512325)

[Finding Analysis and Discussion: 28](#_Toc70512326)

[Conclusion: 31](#_Toc70512327)

[References 32](#_Toc70512328)

Table of Figures:

Table 1: Accurate Detection of Unauthorised Twitter Accounts……………………………….…….….19

Table 2: Weight of Proposed Attributes is determined…………………………………………….…….….20

Table 3: 5-fold cross validation for attributes…………………………………………………..………………..23

Table 4: Classified parallel attributes……………………………………………………………….………………..24

Table 5: Cross Validation for Five Classification algorithms……………………………….….……………24

Table 6: Presents that these are some of the attributes…………………………………….……….….….25

Table 7: Validation Results for Cross Classification Algorithm……………………………….....……….26

Introduction:

Twitter is the most famous operating system network out of all the OSN’s in the world. The users in the twitter can get entertainment & can share information via various features such as, tweeting, hash tagging & mentioning other users. Internet networking services like Twitter and Facebook draw thousands of people around the world and have affected their lives in interaction with social networking. This success of the networking of social networks has led to numerous issues, including the risk of revealing false knowledge to their users by means of fake profiles. This will do enormous harm to civilization in the real world. We introduce a method of classifying false accounts on Twitter during our research. Millions of people worldwide are using online sites that are discovered for social networking such as Facebook, Twitter, LinkedIn, Instagram, etc., creating network connections. A new age of networking has been created by this system of interconnected social network; this is because of the simplicity of it in. Twitter is also a famous platform for news, jokes, moods, and news events. Users are supposed to send messages to their followers on Twitter immediately. Twitter also works as a search engine for users in this real time can also be used to get tweets retrieved. Several variables depend on the rating of the tweets in this search engine, including the number of users. It has become a popular place for spammer of all sorts due to Twitter's success. OSN users exchange an extensive range of network information such as pictures, photographs, school names, colleges, telephone numbers, e-mail addresses, family connections, bank records, job details etc. Now let us talk something about twitters’ usage and functionalities that it allows its users to avail. Twitter allows its users to tweet with a word limit of up to 140 characters per tweet, also it allows its user to share links and update their tweets as well, Users’ can use @ to directly mention the person they want and can address to him/her. Moreover, they can use hashtags to increase the reach of their tweets. Tweets that use any keyword with “#” character. Apart from this it also allows its users to follow others by just one click or tap on follow button, this has become way easier to search trending users by the suggestions twitter gives on the sidebar. So, from those suggestions you can easily scroll or swipe to follow your desired twitter accounts. However, twitter also suggests the accounts of your interests, to who you might know and want to follow. Also, the accounts that have followed you, you can follow them back by again taping the follow back button. Moreover, Twitter allows you to retweet the already posted tweets of other users’ and makes you to decide whether to like it or ignore it. When any user retweets, so actually he /she is posting that tweet to his/her own profile, so as a result all his/her followers, happen to view that tweet and can retweet further. Twitter also gives access of settings to its users, through which users’ have option to make his/her profile public or private, profile is being set to public by default, but when it comes to users’ choice, he/she can retain the default settings or can change the settings and can control the reach of people who are following him/her. In this situation, all the scenario depends upon the profile owner, and he/she decides that whom he/she wanted to be his/her follower and who’s follow request they want to decline. I can this is the best algorithm ever twitter has applied to gain the trust of its followers, that it has maintained an option of setting your profile to public or private, this makes a way more comfortable while being a twitter user.

This knowledge is severe when it is put in the hands of attackers. Most OSN apps do not know about and can easily be prey to security threats on social networks. When the victims are teenagers, the threats are more dangerous. Profile Cloning attacks stolen the private and public, every type of information from the account of current users to build duplicate accounts and misuses these profiles to distort the identification (Chatterjee, 2019, Kontaxis et al., 2011, Kolias et al., 2012, Bródka et al., 2014, Cresci et al., 2014, ElAzab, 2016, Devmane and Rana, 2014). Twitter plays dual microblogging role as an OSN and an online news warning site (OSN). The growth in Twitter social exchanges recently brought cyber criminals to the focus. Spammers spread aggressive messages over Twitter, links of publish phishing and flooding of the network by fake and spam accounts. The method of identifying the network of spammers involved in such operations is a censorious step in the detection of individual spam accounts. There are various methods for identification of spam and fake accounts, proposed by some legendary work of researchers.

Spammers have different purposes: selling advertisements, phishing or merely compromising the system's credibility. Falsifying, misleading and spam powering accounts could be possible. The Twitter rules violate fake accounts. They should not behave illegally. Automatic account interactions or attempts to mislead people such as post hurtful relationships, aggressive habits including mass monitoring, multiple account generation, posting subject items or notifications replacement, posting links to unrelated tweets, and misuse of replies and mention of functionality. To identify Twitter accounts as accurate or not, machine learning styles and harvesting approaches have been used. User profiles are explicitly generated in social honeypot strategies to attract spammers to gather information about their profile. These spam accounts are then studied using the techniques known as engine's learning, this is to understand and percept the behavior of Spammers.

A spammer aim can be expressly specified on the virtual social networks. Intruders may use the information shown on an OSN to prevent, intimidate, impersonate, or even damage a profile holder's identity. Cyber threats have seriously threatened the health and anonymity of users of social networking. Genuine users are accounts in compliance with Twitter guidelines. The same site and the cross-site cloning profile are two kinds of cloning profiles (Kiruthiga and Kannan, 2014, Erşahin et al., 2017). If login details are taken from an individual network, and a clone profile is generated within the same network, then the replication is termed the same site profile (Peled et al., 2013, Gupta and Kaushal, 2017, Fire et al., 2014). In the Cross Site Profile cloning, an attacker user details from a network is taken to build a double profile in another network where the user has no account (Khalil et al., 2017, Khayyambashi and Rizi, 2013, Conti et al., 2012).

As social networking has become very easy to register to draw increasing numbers of users, it is also disturbing those fake accounts are created. To link to a victim, the attacker creates a false identity to create malicious behaviors. Fake news and spam texts are also distributed. Furthermore, all Twitter users can find details such as picture, first and last names, sex, location, and friends' list. Choosing the same name is the easiest method. No specific words are wanted on Facebook. Facebook generates a unique profile identifier for any operation. For regular users, though, this id is not available. Photo profile is also straightforward to buy. The same picture as the survivor will be used by an abuser as all users on Facebook can see the main profile picture. The most challenging thing is to restore a whole list of friends, which means reinventing a user's social network. To identify Twitter accounts as authentic or not, machine teaching methods and harvesting approaches have been used. User profiles are explicitly generated in social honeypot strategies to attract spammers to gather information about their profile. These spam accounts are then studied using the techniques known as engine's learning, this is to understand and percept the behavior of Spammers. Other methods that are specific to detecting spammers and false accounts of twitter involve the following: identification based on tweet text, the use of twitter attributes such as the 'reputation ranking,' the 'amount of similar and clone messages.'

The Fake & Clone profiles in the social media platforms currently have become a significant threat. Therefore, a detection method is essential if these frauds are to be identified who use the faith of people to collect personal data and create duplication profiles. Many scholars in this field have researched and suggested ways of identifying certain social network accounts. Cloning of profiles is an operation to match the original profile. A profile owner's mate normally collects this detail. An intruder might, for instance, clone the profile of any user let say A, the husband of B. A will then send private messages to B, asking about a bank login or missing his smartphone and wallet. An asked B for money to be transferred to his "mate" account. B will assume that this is her husband and might reply or send money to him through a social network. It is also very risky to do this operation. Two nearly identical user accounts would emerge after the profile cloning. Several graphological strategies were proposed to classify Sybil accounts by their graphic social activities and characteristics to deal with spam in social networks. In response, "spammers," who build accounts with complete profiles and background information close to original users, merged Sybil’s into genuine user groups. Such techniques have made an arduous identification effort, and new spam-recognition methods need to be developed.

# Aims:

We aim to detect false and clone Twitter profiles. Because Fake and replica accounts in online social networks have been a terrible issue. In daily life, we encounter any of the other risks from these accounts. Thus, A variety of laws are used to identify fraudulent and authentic profiles as applied for fake profile identification.

# Objectives:

Our main objective is to devise a tool for the detection of fake accounts and for this purpose. Fake accounts are classified based on a sequence of guidelines that can distinguish false and true profiles efficiently. Two techniques will be used for profile identification—one with similarity and the other with the C.5 (five classification) algorithms for decision tree. in which we will use 5 classification algorithm that will be going to apply on the minimized weighted attributes successfully. This is going to be held in three steps in total, in which we will be conducting experiments and forming the subsections.

we will collect 22 attributes which will be then subjected to the series of experimentations, testified, and validated various time to collect the best ones and then minimized attributes were weighted. The weighted attributes will be sent forward to be applied 5 classification algorithms. The way to the application of C.5 (five classifications.) algorithm will be started when the 22 attributes will filter from twenty-two to seven attributes with high results of accuracy. These attributes will be then used for the detection of fake profiles on twitter.

# Literature Review:

The research is basically on subject matters with the inside experience about bots but isn't especially applicable otherwise. However, it does include some beneficial statistics. The authors of this paper intend to speak about the effect of computerized Twitter bills that unfold research information. Below is the critical evaluation of the work of some authors.

1. In May 2019, Dr. Ponnurangam Kumaraguru (Chair), Dr Arun Balaji Buduru, and Dr. Siddhartha Asthana authors paper mainly aim to search the attributes of fake profiles and devise a tool such as machine learning based system to identify and detect the fake accounts.

**Contribution**: An attempt to: qualitative analysis of reported and unreported users. Observe the contrasting and similar traits. Determine the inconsistence traits of a user to be declared as fake. Devise a deep learning technique.

**Data**: A large sum of genuine accounts found in comparison with the scam and reported profiles. Profiles have been reported various times. Reasons of reporting was taken into consideration. Spam profiles found to be more active in spamming people in large number.

**Methodology**: Methodology of Feature Engineering was used, for the detection of fake profile of twitters. Using this methodology, a behavior of user has been identified and then the summary of all the information along with the affinity of the users towards other categories. Moreover, the information was further computed with different time slots to check its validation.

**Conclusion**: In this paper, the research towards the detection of fake profiles was conducted and unique characters of spammers and scammers have been identified. A detailed analysis was carried out toon both reported and unreported users and because of detailed work a successful auto encoder model was trained to solve the problem of fake accounts on online matrimony.

**Critical Analysis**: Model could only learn a limited distribution of normal behavior and any profile that deviates from the standard of uniqueness is considered as fake. Access to private conversation is falsified so a spam through messages cannot be detected.

1. In March 2020, Nitika Kadam, Harish Patidar authors paper mainly aim to identify the fake accounts and establish a tool by the detailed study of data mining and machine learning.

**Contribution**: A Major portion on which this research was conducted are

* 1. Spam or unsolicited post identification
  2. Detection of bots
  3. Detection of profile.

**Data**: Data was obtained from client server architecture in which is a sort of centralized online social networks and from P2p architecture which uses connectivity at a local level and supports communication

**Methodology**: The dataset used in this research follows the following flow, Firstly, they headed from social media platform to relevant API, then to data extraction, after this the data was processed and profiled because of which the content and attributes were obtained, this led them to score the content and then decision making was upheld.

**Conclusion**: The research paper concludes that the fake profiles are classified accurately using a new model formed based on data mining and machine learning techniques and to detect the fake user’s data extraction technique was used.

**Critical Analysis**: Limitations to this paper was the usage of client server architecture which on one hand if provides easy implementation and management on the other hand also led to the network failure and blockage of communication if at a single point is affected by false management.

1. In Jan 2020, Dr. K. Sreenivasa Rao, Dr.G. Sreeram, and DR. B. Deevena Raju authors paper mainly aim to detect the malicious & fake accounts that leads to the misinformation and agenda creation for people.

**Contribution**: The major section on which this research paper is based are

1. Support vector Machine
2. Random forest
3. Neural Networks

**Data**: Sarah Khaled et al. introduced a novel algorithm, SVM-NN for detecting bogus Twitter accounts and bots, as well as feature selection and dimension reduction strategies. This suggested technique (SVM-NN) employs fewer characteristics while successfully classifying around 98 percent of the accounts in our training dataset.

**Methodology**: Following methodology is used, for the detection of fake profile of twitters. This methodology follows this sequence i.e., gather data from accounts and cleaning them Creation of accounts that are fictitious. Validation of data that is extracted Creation of modern and new feature Usage of machine algorithms that are supervised Evaluation of results

**Conclusion**: This paper has provided a Machine Learning pipeline for identifying bogus accounts in online social networks. Our examination of Support Vector Machine, Random Forest, and Neural Networks demonstrated good performance, with the accuracy of prediction appearing to be greater utilizing Support Vector Machine for the provided dataset. For a particular data set, the accuracy of detecting bogus accounts is shown to be greater when utilizing the Random Forest Algorithm, followed by the Neural Networks Algorithm.

**Critical Analysis:** No limitations has been observed. However Neural networks using recurrently for a better detection of fake accounts

A research conducted by Cresci (2015), analyzed the most important of the current features and guidelines for identifying bogus Twitter accounts. The rules and features have been used to train several system classifications. Then they have created a class A classification that can distinguish both original and false accounts effectively. To build a technique to detect profiles effectively, you must know how to clone profiles. By Wang (2019), one of the first techniques was introduced. The solution for cloning profiles in Twitter is identified. Three principal components were developed for this solution: information distiller, hunter profile and profile verifier. First, the core data from the present user profile must be found in the process for each specific user to be unique in its attributes. Users will limit the query used to identify identical profiles, copied using only particular features and not typical such as home city or nationality. The second part is used in various social networks to identify related profiles. The n profile list is then moved on to the last detail, verifying the similarities and displaying all the profiles to the user to verify which are real and which are clones. On all social networks, this approach is universal. Only a principal example on LinkedIn was presented, however, by the authors. In other social networking sites and other search engines, the same approach may also locate more details. Research by Grier et al. (2017) is the second solution. This paper presents a close detection mechanism to the last one (Wang, 2019). Firstly, the profiles with a name relative to the base profile would be looked for. The list of candidates includes all accounts of the same name. It measures the similarity using its algorithm for all profiles in the applicant list. The profile will be added to the suspect identify catalogue, a list of all profiles that can be clone, if the measured resemblance is above the predetermined limit.

The identification of false accounts on Twitter was suggested by ELazab and Mahmoud (2018). They gathered some successful characteristics from different researchers for the detection process, then filtered it and weighed it in the first step. Additional trials are carried out to obtain minimal attributes that yield correct outcomes. Just seven characteristics have been chosen from the 22 features that can successfully identify fraudulent accounts and apply them to classification techniques. The results-based classification methods are compared, and the most accurate results are chosen. In this research, we further explored the technique of easily identifying fake profiles based on users’ information, to be able to early classify the classified unauthentic Twitter accounts. We use a consumer profile-sample detection primarily based on the method, which uses the inclusion of consumers’ old information, to expand a brand-new method for detection of false & fake profiles with high score of reliability.

# Why need of fake account detection?

Social networking has grown a lot in past 2 decades and is still growing very fast, during this rise different sites for social activities have been created, which is immensely attracting the people to join and avail these sites and flood it with population. Besides the over flooding of sites, online social networks (OSNs) are also facing the issue of increasing number of fake accounts. Fake accounts are those accounts whose information doesn’t match with the real account or human account and in this way, they are compromising the credibility of that particular social site. Fake accounts are intending to spread fallacies and flood the site with bots, also some fake and unauthorized accounts are there for stealing the personal information of the real user to make clone accounts and then use that information as their own, no fake accounts are acceptable to exist in any social site this is because fake accounts not only misuse the information of authentic user but also compromises the reputation of him, and spreads fake information and fake rating of web. . Another thing that is needed to be introduced is Twitter API, Twitter API is something that lets a user to access and get assistance from twitter data, which is used to compose tweets, get insight of users’ profile, and can have way to get the data of other users for having a more and more tweets at a specific location. Twitter’s API is not for the leisure person, it is only available for the authentic certified users, However, these certified users are allowed to make a search on desired information via different ways. Initially these are four main ways via which users can search the required information, which includes Tweets, Users, Objects/Entities, Places. Moreover, it also includes the ID of the twitter user and attributes + characteristics of the places. People mostly makes a use oath for sending the delegation of having access to twitters’ API. This delegation is accepted by determining the person’s signature and identity of his/her request, in addition to the identity of the users’ access identity and the access tokens’ identity that is the representing the interface, identity is granted by the end user. For reaching the information, the whole world uses the keyword of API in their tweets; posts and information of other users can be searched by targeting only one city and one place mainly, using the location, can search all the tweets, retweets and all the replies related to one user, by just using follows and following that user. The API of twitter is not free as we have said earlier that it only gives information access to certified users, so basically it doesn’t give free hand to its certified users’ also, and they cannot do anything they desire. It has set some restrictions and limitations in its API, to protect the bandwidth damage from the attackers and spam bots. It has limited the users’ request to 180 each 15 minutes. It means only 180 minimum requests can be accessed for each 15 minutes. If the limit is exceeded, the document is produced by the name “REST CALL” which informs the other delegators that the API is at rest right now and no more requests can be utilized right now. So, no matter what, twitter never allows its API to reach more than 180 request per fifteen minutes. Also, there is another limitation, in which the page count and parameters are considered, in which it only returns up to 3200 states, ignoring the pages above them. Although the twitter has laid down such restrictions, but it doesn’t really own other restrictions. For instance: It is recommended to cache the results locally to avoid retrying requests with the same status. generally, there are various sorts of the HTTP request, POST, and GET. This paperwork additionally invokes Twitter API. Simply put, it forwards the POST and GET requests from customers to the unique API address and returns the HTTP header and the contents to the client, that fulfills all the capabilities of the unique Twitter API. For the client, in addition to offering a choice of opportunity configuration API address, they now no longer want to change any of the code. For the subsequent scenarios, commonly the most typical approach researchers use to look a few facts is to access twitter.com at once to look the Friends list, in fact, it calls the GET request. Also, there are personnel of twitter, who've get right of entry to the total statistics of every account. They use these statistics in a 3-step process: first constructing a cluster of accounts, then a ‘profile texturizer,’ and finally a decision-making step. They rent system getting to know strategies as many other researchers have, however those authors have gotten right of entry to statistics that most do not, such as IP addresses. A worth quote addressing the effect of bots:

“For example, the credibility of the community, can be weakened, if customers begin to doubt the authenticity of profile information. Also, it impacts negatively on the networks’ advertisement revenue, considering advertisers would possibly query the rates they pay to attain a sure wide variety of customers if a lot of them are bots instead of real person. this paper presents the authenticity of this topic, through talking about basic approximate brand degradation, that's results from distrust in an OSN. The paper accomplishes its goal, which ends up in a high score of accuracy cluster detection, claimed at 95 % accurate matches. Overall, this paper offers perception to different metrics that can be brought to the toolbox, which includes pattern detection in naming schemes. If it had been to try to study it again, then definitely, it would again be done in the same way. Either in a skeptical way that such a high score of accuracy rate may be detected for complete clusters, or the worker has been done in a simplest way seeking for huge bot networks which are unsophisticated such that they're detected with ease due to: the use of the same IP, comparable names, registered very carefully in time, etc. and now no longer seeking to discover some other types of bots. Operation system networks operators are continuously working on managing these bots, and various authentic resources to detect and confirm the human and bot accounts or spammers’ account. Twitter is used by a large amount of community. A wide range of its users includes the users that are using from mobiles, they contribute up to 46% of the total, tweets are getting published via the medium of text messages, sending emails or SMS, and exchanging the information containing up to the limit of 140 characters as discussed earlier. These exchanges can be done directly from the mobiles, and it has made it an easy access to the users to be used twitter at their smartphones, in groups point of view people who are regarded as active users; share information. There two main classified events that are regarded as “Seemingly incredibly” and “Clearly incredible”, the type named as seemingly incredible is actually representing the events, that seems to be fake such as the contradiction between the sayings and tweets and do not containing any authenticity of reference, however the 2nd credibility is clearly fake, for instance; fake rumors and spreading of the fake information such as unauthentic news announced by any celebrity or politician, and propagandas that are going viral among the public for no reason.

The main issue evolving today on social networks is the prevailing fake accounts, that are using their powers to target the account of verified user, then stealing and using it for their own evil purposes. One of the problems that has been addressed due to this targeting is the ruining of reputation of a specific company that is affecting their business approach. Or also these are affecting the society at a larger scale. There is a situation occurred in 2013, that during the time of people suffering from the aftershocks of Boston bomb blasting, a spammer taking the benefit of this situation released the fake tweet for donation of 1$ (1 dollar) for each retweet, to help those who has been suffered from the blast.

Thus, from this case scenario it was being felt that social media is having a huge impact on the society, and it should be protected against all such scammers and spammers, else they would be going to destroy the society and can be proved a way devastating. As the social media has a great impact on the society and there is a need of an authority which should kept an eye and keep a check on each social site, along with it the dire need for the detection of fake accounts was also felt, and then some legends decided to work on it devise different and particulars solution. That not only detect but defy all the fake accounts from the twitter. Although we have other sites such as Facebook, WhatsApp, and Instagram, which have impact on people too, yet Twitter is considered the most authentic source of news and any announcements or incident. Researchers start to keep on working day and night and intended to devise a technique for detecting fake accounts from twitters’ OSN (online social network) as a step ahead. The section of Data methodology presents a detailed description of how the technique of detecting fake profile was devised, what were the algorithms applied and how the classification of algorithms was obtained.

# Previous work:

Various research has been conducted to detect fallacies of the twitter accounts with various approaches. In this research the feature-based strategy of detection will be used, this strategy of detection will be monitored based on the behavior of the twitter user, or we can say the owner of a particular twitter profile who is running the twitter account. The behaviors of the twitter account user can be taken into consideration such as number of tweets he does per day, no of retweets he does per day, no of friends he has on his twitter account etc. This program was based on the concept that there’s a difference between the human activities and bots’ activities, Human usually behaves differently in this case scenario, therefore observing these behavioral activities and the deep study of this approach lead us to the discovery of fake accounts and give us a way to differential between authentic users and spam bots. In this section, some of the pre-working of the researchers has been presented and is keenly objected to the detection of fake accounts. The work has been already reached the high score of accuracy and the accuracy rate is almost up to 84.5%. In this work, the approach was initially held to detect faux accounts with up to 85% accuracy with the identification of 22 attributes. We have represented most of these attributes in Table 1. However, the high rate of accuracy has been achieved by smaller set of attributes which is discussed in the section of Data Methodology. Out of these 22 attributes 17 attributes were finalized, and out of 17 attributes 10 attributes were then proceeded for fake account detection. All these attributes are listed in Table 1; However, we have already mentioned that the results obtained from this research are not much optimistic and can be taken for granted for the identification of fake accounts on twitter and fake users’ profile, and it’s not sure that the fake users’ profile detection or fake account detection by the help of graph algorithms, is going to give results with higher rate of accuracy with absolute correctness and preciseness.

Although the research that has been conducted, have minimized the set of attributes from 10 to 6 in total. However, it has been already discussed that it can only detect the spammers that has some special qualities as we have mentioned before, those spammers are listed as baggers and posters. In our strategy we have decided to devise a minimum set in which the minimized attributes will be presented to detect all the types of spam bots. Moreover, one of these attributes will be requiring the strategy in which we will be analyzing the texts, so that we can detect the similarities among the texts will be measured, which does not actually play an important role in our methodology. In addition, the best results of detection were obtained from the Random Forest algorithm. Although most of the other research such as that claims the success of other algorithms such as support vector machine, and Bayesian network. However, our results matched the conclusions that has been deduced by them, so it made us capable of proving that by the practical presentation of all the sic attributes we can claim our successful results via the classification algorithm of Random Forest.

The authors have been contacted from different countries to ask for the set of data on which they have been working or worked on. And after this we then progressed and moved forward to work on this dataset to prove the authenticity of our work. The proposed methodology and details of this research paper are mentioned below in the data methodology.

# Data Methodology:

In this paper, we intend to successfully hit upon the faux accounts on the social community of Twitter with the viable minimal set of attributes. This proposed plan includes two fundamental steps, step one is to decide the primary elements that plays important role in an accurate detection of fake accounts and the second one step is to use a category in which there is defined a set of rules that makes use of the decided elements in the first step of an accurate detection of fake accounts. These main objectives of this research paper are to devise the least set of attributes that can detect the fake profiles with maximum accuracy. To find the best features that are less in number but can detect a greater number of fake Twitter accounts.

A lot of experiments have been conducted so that at least some of the ways could be discovered for the detection of fake accounts on twitter. The predominant intention of this task is to locate the excellent minimal characteristic set this is precisely primarily based at the considerable project of extracting, getting ready and studying those features. Therefore, locating the smallest ratio with the very best accuracy is taken into consideration as one of the powerful instructions for account identification.

Clone and fraudulent accounts have been a severe social threat. As information including mobile number, e-mail address, school and college identity, business name, address etc., are easily available in social networks, hackers can easily access this information and create new or clone accounts.

The operating plan that has been accomplished to come across the required features’ set are defined in steps, steps from 1 to 15 has proposed the operating plan in detail. A working plan to determine the required characteristics is gradually established various steps, a detailed plan is proposed below.

1. A survey has been carried out that described distinctive sets of features. wherein all of the features were accumulated that are proposed via researchers.
2. The experiment was carried out from different angles. In the first experimental series, the five most successful classification algorithms are applied to the data set.
3. The algorithms then passed to the series of validation that held upto 5-times, and the consequences of the validation series then accomplished.
4. The selected attributes then subjected to further testing and minimized. And then being applied a Gain Measure on, to find the weight of them.
5. In the third step of experimetation, the calculated weight of the attributes again used to validate further five classification of algorithms & then again applied to the data set.
6. Again the five cross validation of the algorithms conducted and then results have been compared.
7. The Gain Measure again applied on the selected and mininmized attributes.
8. By comparing the weight of all those attributes in the fourth step of experimentation, the five classification of algorithm again applied on the given dataset on the basis of the selected attributes
9. Again the five cross validation of the algorithms conducted and then results have been compared.
10. In the beginning we had 22 attributes in total, which then testified and validated and then minimized to 19, now in the fifth step of experimentations, attributes are being further validated and scrutinized to 7 attributes. This was according to the weight of those attributes that had an equivalent to or above to 50%. And after this process and the 5 classification algorithm again subjected to the dataset.
11. After this the previous step of five cross validation of the algorithms repeated and then consequences have been compared.
12. In the sixth step of expermentations, we tend to repeat the step of using weight which was calculated for 19 attributes, 5 classification algorithm again subjected to the dataset, according to the weight of those attributes that had an equivalent to or above to 50% and this minimized the attributes fro 7 to 6.
13. After this, again the previous step of five cross validation of the algorithms repeated and then consequences have been compared
14. The whole series of experiments lead us to collect the minimum and scrutinized set of attributes with high score of classification and the best algorithm were then selected, so that we could capture the unauthorized accounts of twitter.
15. The minimum set of attributes with maximum number of accuracy were reached and there is only 1% chance of failure, However the 99% chances shows that all the fakeness is going to be caught sooner or late.

Following is the table of research that has been conducted which lead to proposal of attributes that has been thenn subjected to experiments and minimized for the accurate detection of unauthorized twitter accounts.

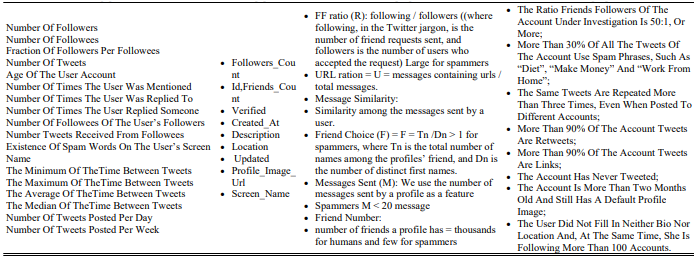


Table 1: Accurate Detection of Unauthorized Twitter Accounts

## Dataset :

Experiments have been applied to the dataset of Fake twitter accounts. There was a project named as “The Fake project”. This project mainly consisted of the data of those accounts which were verified and the data of those accounts which were fake. The data collected from the Fake project was collected from a source. And the collected data consisted of the following information:

Verified accounts:

* 1481 verified accounts that belongs to humans, collected from #elzioni2013.
* 469 verified human accounts have been collected by the “Fake Project” team.

Fake accounts:

* 1000 fakes accounts from <http://fastfollowerz.com>, at price $19.
* 1000 from http://intertwitter.com, at price $14.
* 1000 fake accounts from <http://twittertechnology.com>, at price $13.

Summarizing the above information, we get to know that from the project named as “The Fake Project” collected the information in total, was 3000 fake and unauthorized accounts and 1481 human accounts.which were authenticated and verified from twitter. This dataset will help us to compare the consequences of our work with the given methodology and will assist us in the correct validation of our proposed attributes. However we have proposed 22 attributes initially which then subjected to the series of experimentations, validated further and then minimized to the seven attributes. In this research paper, all the attributes are being collected so that we may apply a Gain Measure on them for calculating the weights of all those features. This will however, assist us in the 5 classfication of algorithm and will efficiently determine the height of its accuracy.

Table 1 - Data attributes and weights associated

|  |  |
| --- | --- |
| **Attributes/characteristics** | **Weights** |
| The account which has not more than 30 followers | 0 .52 |
| geo-localized accounts | 0 .86 |
| user’s favorites accounts | 0 .86 |
| Accounts having one hashtag used | 0 .95 |
| Uses iPhone to log in twitter | 0 .926 |
| using a mention, a mention in tweets | 1 |
| Account that tweeted 50 tweets | 0 .01 |
| Account that has been a part of another user’s list | 0 .43 |
| (2\*number followers) \_ (number of friends) | 0 .4 |
| Twitter Users’ having at least one Favorite list | 0 .18 |
| the users’ proﬁle containing a name | 0 .0 |
| the users’ proﬁle containing an image | 0 .0 |
| the users’ proﬁle containing a biography | 0 .0 |
| the users’ proﬁle containing a URL | 0 .0 |
| An account writing a tweet that uses punctuation | 0 .0 |
| Uses iPhone to log in twitter | 0 .0 |
| Uses Android device to log in twitter | 0 .0 |
| the users’ proﬁle containing a physical address | 0 .0 |
| Uses .com websites to log in twitter | N/A |
| An account that relates to Foursquare | N/A |
| Uses different clients to log in twitter | N/A |

However this table shows that the account that has followers less than 50, weighs upto 0.52, the geo-localized accounts and those accounts which are being added to favourites by other users weighs upto 0.86. Those accounts which uses hashtags in at least one tweet or more, and also the twitter account that has been logged in using an iphone weighs almost upto 0.9 and something. Mentions used by twitter user has the greates of all weights which is 1. The account that has tweeted atleast 50 times, weighs 0.01. The account that has been a part of following of any other user’s list is 0.43. Account having number of followers and friends\*2 weighs 0.4. Those users who have added atleast one account into their favorites does weigh about 0.18. However those profiles which contais a name, image, biography, URLs & physical address doesn’t count because they weigh 0.0. Also those accounts which uses puntuation in their tweets, are logged on twitter using iPhone and those whch are logged into twitter using an android device also weighs 0.0. Moreover the accounts that are logged in from twiter.com website, those which are connected to foursquare, those which are connected with instagram and those which are logged into twitter though different accounts are nil.

## Classification Algorithm selection step :

As we have already discussed in the methodology about the five best classification algorithms, Now in this step we have applied those, using the attributes that have being weighed

The algorithms includes:

1. Random Forest.
2. Support Vector Machine.
3. Naïve Bayes.
4. Neural Network.
5. Decision Tree.

For each algorithm,the standard indicators are devised. In this way outcomes have been summarized for each algorithm. However the four standard indicators that are introduced to the algorithms are as follows :

1. True Negative (TN).
2. True Positive (TP).
3. False Negative (FN).
4. False Positive (FP) .

After the step of applying standard indicators to the 5 best classification algorithm, we measure the three standard metrices which are :

1. Precision.
2. Recall.
3. F-Score.

# Experimental Results:

As I have already discussed in the dataset, about “The fake project”, and talked about the experiments which have been applied on the dataset. Now, in this section I will discuss about all those experiments which have been discussed earlier. Let me tell, that there are three steps in total in which we have conducted experiments and formed the subsections of each as Subsection A, Subsection B and Subsection C. Below is the briefing of each subsection:

## Subsection A:

This section is the one which would discuss about the 1st step, and in the 1st step we have applied the five-classification algorithm to all the attributes which were collected and then minimized, and I had presented the experimental results in this section too.

## Subsection B:

This section is the one which would discuss about the 2nd step, and in the 2nd step, what I have done is the application of five classification algorithms on the weighted attributes by the usage of a technique known as GAIN MEASURE and in this section, I had also presented the experimental results.

## Subsection C:

This section is the one which would discuss about the 3rd step, in which the classification of 5 algorithms were being applied successfully on the set of attributes which weighs more than 50% or were equal to 50% in total. Again, in this section, results are provided.

This discussion is finalized, when the complete set of attributes with high efficiency and accuracy but précised number of attributes will be provided. Presented experiments reached the standards of qualification as they are not being measured locally but by the standard metrics and all those were being presented to each set of experiment. The correctness of all the algorithms or you can say techniques and the set of characteristics or attributes, which was best of all were obtained by the 5 cross validation process, which was discussed earlier in the data methodology. The five cross validation was applied to all the experiments and the required metrics were calculated at each fold. After this process there was another step, in which the average of these results has been calculated and the results were being compared by the results of other ones.

### A: 1st Experiment:

Now it’s time to discuss the Subsection A, which is the one, which would discuss about the 1st step, and in the 1st step we have applied the five-classification algorithm to all the attributes which were collected and then minimized, and we had presented the experimental results in this section too.

Below is the table, which shows results for accuracy metrics, 5-fold cross validation for all the attributes

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Machine Learning  Algorithm | TN% | TP% | FN% | FP% | Precision% | Recall% | F-Measure% |
| Random  Forest | 84.67 | 94.21 | 17.43 | 3 .71 | 96 .18 | 71 .06 | 81 .76 |
| Decision Tree | 86.34 | 88.93 | 33.67 | 5 .75 | 93 .95 | 67 .05 | 78 .25 |
| Naïve Bayes | 78.23 | 82.06 | 58.93 | 8 .18 | 92 .78 | 62 .07 | 74 .48 |
| Neural Network | 77.17 | 87.31 | 34.58 | 8 .56 | 93 .31 | 66 .39 | 78 .86 |
| SVM | 56.63 | 99.52 | 11.47 | 15.97 | 87 .71 | 73 .89 | 78 .14 |

Table 3: 5-fold cross validation for attributes

This is the table which the results for accuracy metrics are clearly presented, and the results are obtained from the implementation of 5 classification algorithms on the dataset by the usage of all those attributes which were selected in the classification task.

Below is the table of attributes which are being classified with the attributes that are parallel to them.

|  |  |
| --- | --- |
| **Attributes/characteristics** | **Weights** |
| The account which has not more than 30 followers | 0 .52 |
| geo-localized accounts | 0 .86 |
| user’s favorites accounts | 0 .86 |
| Accounts having one hashtag used | 0 .95 |
| Uses iPhone to log in twitter | 0 .926 |
| using a mention, a mention in tweets | 1 |
| Account that tweeted 50 tweets | 0 .01 |
| Account that has been a part of another user’s list | 0 .43 |
| (2\*number followers) \_ (number of friends) | 0 .4 |
| Twitter Users’ having at least one Favorite list | 0 .18 |

Table 4: Classified parallel attributes

### B: 2nd Experiment:

In the 2nd step, what we have done is the application of five classification algorithms on the weighted attributes with the help of the Fake project & by the usage of a technique known as GAIN MEASURE. A weight that corresponds to each attribute is according to the importance of that attribute, it possesses in the dataset. GAIN MEASURES has discovered that except 10, all the rest of attributes have zero weight. Those 10 attributes possess positive weighting so experiments had to use the attributes of weight higher than 0%.

Below is the table which presents the results of. Five cross validation process for five classification algorithms, that was held with the help of weighted attributes.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Machine Learning  Algorithm | TN% | TP% | FN% | FP% | Precision% | Recall% | F-Measure% |
| Random  Forest | 97.46 | 92.21 | 17.43 | 0.71 | 98 .48 | 71 .06 | 83 .76 |
| Decision Tree | 97.43 | 97.93 | 5. 67 | 0.75 | 98 .49 | 73 .05 | 84 .25 |
| Naïve Bayes | 97.75 | 98.06 | 3. 93 | 0.81 | 98 .28 | 73 .07 | 84 .98 |
| Neural Network | 98.87 | 97.31 | 6. 58 | 0.56 | 98 .51 | 72 .39 | 83 .85 |
| SVM | 98.63 | 97.52 | 5. 47 | 0.47 | 98 .71 | 73 .89 | 84 .13 |

Table 5: Cross Validation for Five Classification algorithms

Table-5 clearly presents the results that are obtained with five cross validation process for five classification algorithms, that was held with respect to the weight of the attributes.

Below is the Table that shows those attributes that weighs equal to 50% or above than 50% parallel to their weights.

|  |  |
| --- | --- |
| Attributes/characteristics | Weights |
| The account which has not more than 30 followers | 0 .52 |
| geo-localized accounts | 0 .86 |
| user’s favorites accounts | 0 .86 |
| Accounts having one hashtag used | 0 .95 |
| Uses iPhone to log in twitter | 0 .926 |
| using a mention, a mention in tweets | 1 |
| Account that tweeted 50 tweets | 0 .01 |
| Account that has been a part of another user’s list | 0 .43 |
| (2\*number followers) \_ (number of friends) | 0 .4 |
| Twitter Users’ having at least one Favorite list | 0 .18 |

Table 6: Presents that these are some of the attributes

The above table presents that these are some of the attributes that weighs equal to or above then 50% and includes the account

* with at least 30 followers.
* that has been geo localized.
* that is included in another user’s favorite.
* that has used hashtag in atleast one tweet .
* that is mentioned by any twitter user .
* that is logged into twitter by using an iphone .

Here we have another table named as Table-7, that shows the validation results of 5 cross fold classification algorithms. Which obtained by using weighted attributes of 50% and more than 50%.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Machine Learning  Algorithm | TN% | TP% | FN% | FP% | Precision% | Recall% | F-Measure% |
| Random  Forest | 97.46 | 92.21 | 17.43 | 0.71 | 98 .48 | 71 .06 | 83 .76 |
| Decision Tree | 97.43 | 97.93 | 5. 67 | 0.75 | 98 .49 | 73 .05 | 84 .25 |
| Naïve Bayes | 97.75 | 98.06 | 3. 93 | 0.81 | 98 .28 | 73 .07 | 84 .98 |
| Neural Network | 98.87 | 97.31 | 6. 58 | 0.56 | 98 .51 | 72 .39 | 83 .85 |
| SVM | 98.63 | 97.52 | 5. 47 | 0.47 | 98 .71 | 73 .89 | 84 .13 |

**Table-7 .**

Table 7: Validation Results for Cross Classification Algorithm

### C: 3rd Experiment:

In the 3rd step of experiment , The C.5(five classification) algorithms were applied on the set of data using set of attributes which were filtered and whose weight were determined earlier. In this experiment the same techinique was used to measure weights as that of produced by application of GAIN MEASURES. GAIN measure was applied on the dataset;s each and every attribute, corresponding to their weights. And then the consequences were compared. In this experiment, we used some attributes in the classification task, however the attributes are specifically be those whose weight is more tha or equal to 50%. After this, it has been noticed that , the attributes that are presented in accordance with their correspondig attributes were only 7, which can also be seen in Table-6, whereas the table-6 shows those attributes that weighs equal to 50% or above than 50% corresponding with their weights.

However there are some other methods which we kept under consideration.

# An Overview of Other Methodologies.

## Detection of Fake Profile (Honeypot/Machine Learning):

Many other strategies have been developed to deal with the problem of spam in social networks. These techniques have used graphs theory methods on large scales to expose the assets of Sybil accounts (Danezis and Mittal, 2009). However, spammers proved to be very intellectual in this regard and as a response to this act, they used the perfect strategy of making authorized accounts, that were exactly similar and identical to the original ones, to avoid such detection and to be in their safe zone. They used the background information of the authentic users to complete their fake profiles instead of revealing their fakeness (Yang et al., 2011). Well, this somehow, further complicated the investigators to capture and restrict the fake account. The method they used for dodging proved to be a masterpiece and lead to the need of developing more strong detection technique in response.

For detecting bogus Twitter accounts, this module, fake and unauthorized accounts are found here based on guidelines that essentially separate fake profiles from genuine profiles. Any of the rules used to spot fake portfolios of accounts usually do not have name or picture in fake profiles. No summary of the account is included. This area is geo-enabled, so it will not be seen in tweets. They usually make more tweets, or they often don't have tweets etc. them. The rules are added to the profile, and a counter is raised with each matching rule. If the counter value is higher than the pre-specified threshold, the profile is considered incorrect.

To verify the authenticity of any twitter account, techniques like machine learning and honeypots were devised, which very successfully allured the spammers to be trapped in their networking sites, aimly to get out the information from them. Utilizing machine learning techniques Profile information for spammers, obtained using approaches like a honeypot harvest, to educate oneself in understanding how spammers behave and developing sensing technologies, which help detecting their accounts. Other ID methods specific to Twitter spambots and fake profiles include. detection frame on the content of the tweets (e.g., “number of hashtags per hash word” of every tweet, and tweet usage/ the features of tweets as "reputational ratings"; “Double Tweet Count” and “No of URLs” and matching tweet links (URLs) to publicly listed URLs/fields.

Most up-to-date methods for detecting false accounts in Twitter and other OSNs have focused on screening for false consolidated accounts based on their business. For instance, Jiang et al. (2014, 2016) used a method, called Catch Sync for detecting suspect behaviors in Twitter based on user’s sync and anomaly activities & through this they were able to demonstrate that their parametric approach resulted in highly effective detection of fake accounts on twitter. Likewise, another man Clark proposed a method of classification scheme in which he used to detect messages which are auto generated, this helped him much in the detection of fake accounts, because in this method a random tweet history has been generated which leaves a mark of fakeness. The identification of false accounts is a blend of user’s profile information and the tweets he/she does. So, another technique using which, it has been shown that based on the fewest factors (including the number of followers), The availability of geographic information, the use of hashtags in tweets, etc. Fake Twitter accounts can be efficiently identified. Also, a controlled machine learning pipeline, which used to compare the frequency of text occurrences between functions such as names, emails, etc., and to identify twitter accounts as malicious or authorized.

## Similarity measures for clone profile detection:

Although there are many differences in social networking sites, but still there is a common concept between them, that is, each user presents a unique entity for viewing, and has a constantly changing profile that provides its image and attribute list to every user in the online world. This module recognizes clones based on the similarities of the attributes and the network. The technique of cloning a consumer profile in a social community includes replacing a targeted account with the aid of using developed first-class feasible copies of users' information in their social profile. In the simplest scenario, it consists of constructing relationships and sharing profile information within the identical way as that of the victim's profile, from every other social community. As a result, an attacker can extract the information and doing an identity theft of the targeted profile in every other social community. In simple words, the feedback is the user profile. The profile extracts user identification information. Check for accounts with characteristics that complement the profile of the user. The Similarity Index is determined, and the profile is called clone if the Similarity Index exceeds the threshold, otherwise normal (Chatterjee, 2019). However, if we put an eye on, why this is needed so we should know that identity clone is the attack that targets user’s identity in OSNs, the main thing it does is duplicates the user’s identity, misuses his information, and compromises his reputation. An automatic resistant model was created to detect clone profiles, this model used some techniques in OSNs and succeeded in reaping the information of user’s profile, for specific time. This model gave an almost appropriate an accurate information about users, in other words, exposed the user’s true credentials, and developed techniques of profile cloning attacks by giving and idea of activating a fake profile in real users’ communication. However, for this technique to be applied on, there’s a need of users’ trust mechanism along with the various activities of OSNs together.

Another technique was proposed by Carminati and Ferrari, in which they evaluated the connections and profile attributes of the users of social networks. Carminati and Ferrari however, succeeded in giving the measure of finding similarity of profiles, without observing the large fraction of users’ network. Also, the measure of detecting a similarity that is very common and known as “SEMENTIC SIMILARITY” among users, was devised and the technique to find the missing values of user’s profile was tied along with it.

A tool for the detection of clone profile was presented by Kontaxis, Polakis, Ioannidis and Markatos (2011). In this detection process, there three main steps which were taken into consideration by them. i.e.

* Information Distiller.
* Profile Hunter.
* Profile verifier.

The algorithm, that is used to compare attributes of the profiles was “string matching”. However, what challenged them was that they neglect the similarities of friend’s user’s networks and didn’t paid attention to the privacy and publicity of the attributes while comparing them. Another method that was explained by Takabi and Joshi, in which they inferred all the sections for detecting fake identities on OSNs. They only kept forced on the fact that extracting an identity from a profile is not just extracting the attributes of the profile, but also the friends they have in them can be fake.

For this purpose, they used “COSINE SIMILARITY MEASURE” which compared the similarity measure of the profiles successfully. The similarities that have been compared between the two types of accounts was based on the Basic profile similarity and the similarity of multiple types faked identities. Although their strategy helped them in detecting the clones, but they failed to achieve accuracy and the profiles’ strength of relationship was neglected by them and they did not use the right measure of comparing the similarities.

Nevertheless, research now-a-days is working hard on protecting the privacy of the users’ information and profile in the given operation networking system. Instead, some researchers also noted that there may be a probability of profile extinction in the last fancy social network. This is because the fake users are almost totally impersonating the original users in the operating system networks. Unauthorized Fake personal information can be used to establish online relationships with friends of victims of identity theft, with the purpose of stealing the victims' personal information through online interactions with the victims' friends.

### Similarity of attributes:

The similarity of attributes is determined based on the similarity between profile values. The similitude calculation attributes are name, screen name, language, position, and time-zone. Two tests of similarity calculate the similitude between the characteristics - Cosine similarity and Levenshtein distance. The resemblance of Cosine is used to find parallels between terms, and the distance between Levenshtein is used to find similarities between two sequences. The cosine similarity of two vectors is 1 when with the same emphasis, with a resemblance of 0 when with 90° and -1 when diametrically opposite (Chatterjee, 2019). The distance from Levenshtein is a similitude metric that measures the resemblance between two sequences. The minimum number of insert/delete/substituting operations appropriate for changing one sequence into another is the Levenshtein interval between two sequences. As we know that similarity of attributes measures is the calculation of similarity between two profiles, according to the they value they possess in their respective fields. Categorical data is the obligation of users’ profile while measuring the similarity of attributes. It returns 1 when both the attributes have same values (indicates that attributes are same or they are like some extent), however, when it returns 0, it means that they don’t have same values, means they have different values. (Indicates that similarity doesn’t exist). Organization of the item or we can say, the attributes of the profile in a respective network should be organized in an orthogonal dimension, Attributes can be of single values, and they can also be of multiple values.

### The similarity of the network:

The similarity of networks is determined based on relationships between networks (Chatterjee, 2019). In this case, the attribute Followers ids are used to find network resemblance between the profiles. Followers’ ids include a list of user-subsequent profiles. To prove that this profile is real, the clone profile still attempts to interact with the same set of users. But we can figure out whether they're identical in terms of network relation or not when we compare the followers' ids of two accounts. Similarity among the networks help in finding the users having similar attributes. Identifying the communities of the twitter users helps in gathering the data of them, by defining several similarity checks based on the content they shared, based on the people they follow and the interaction they have with them. Their interaction among the common follower or a friend, the hashtags they use, the calculation of the user mention similarity between two networks. By calculation the similarity of relationships they have established among two users. By computing the similarities of the hashtags, documentation of all the similarities and weighting some other attributes of networks. The calculation of the similarity of the way user’s respond to other networks. And number of calculated frequencies among them. Also, the similarity of mentioning account was calculated by the applying checklists on the number a user mentions other user Moreover, the sum-total of similarities was calculated as a Linear combination of all the individuals and a new methodology was proposed as model which was the basis of Dirichlet allocation of extracting the clusters of users discussing all the topics except trivial ones. For the identification of fake accounts, a network based on the on entities of user’s profile that matches or are exact copies of each other. Considering similarity of attributes and networks, there is a criteria of similarity check, which is evaluated via the help of network graph from the matrix known as adjacent matrix and used a method of PCA to extract new functionalities. Moreover, the data is balanced using the SMOTE technique and then being sent to classifier. After this the classification technique is applied and through the cross validation process the classifier is tested and validate further.

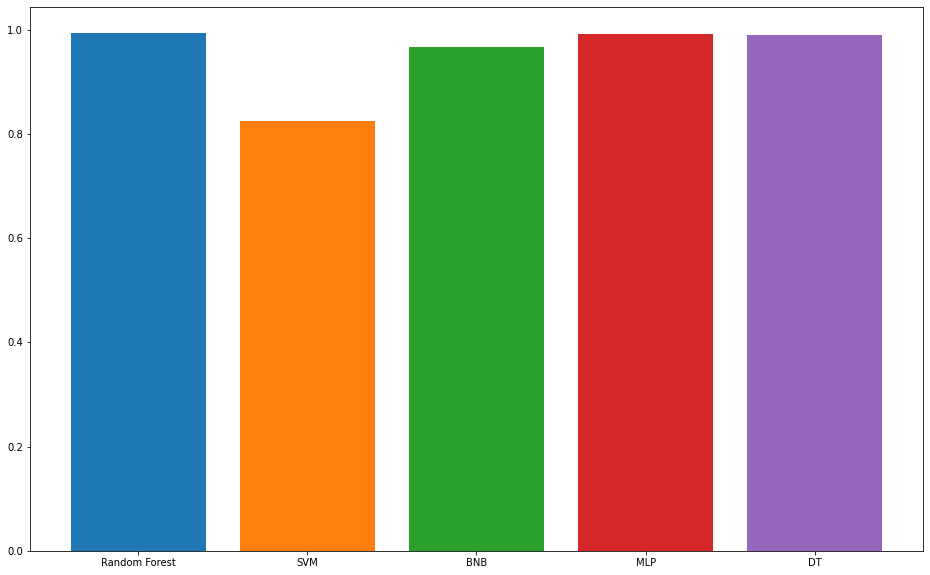
However, we are not going to investigate any of such approaches in such deep depth, as we have already discussed our methodology of fake profile detections in detail, in which we have used 5 classification algorithm that has been applied on the minimized weighted attributes successfully. What we have found during our research is discussed below in “Finding Analysis and Discussions”.

# Finding Analysis and Discussion:

In this section, we will discuss about our findings about this research, as we have discussed earlier in experimental results in subsection A, subsection B & subsection C, that the set of attributes have been minimized, on which the classification task was applied. The focus of this research was minimizing the set of attributes, by rigorous validation and testing so that we can collect the best ones and apply the C.5(five classification) algorithm on them. However, these attributes are not usually available for experiments. The classification task of detecting fake twitter accounts has revealed that the minimized set of attributes were more towards the accuracy and preciseness.

However, by applying different classification algorithms, we have achieved the high score of efficiency in detecting the fake profiles and obtained satisfactory results. The main assistance was gotten by the minimized attributes which weighs equal to or above then the 50% corresponding to their weights. And these attributes get themselves to be ready for the application of classification algorithm, to get the results with maximum accuracy.

Moreover, to clarify our research we have devised a graph whose snapshot is pasted below, in this graph we have made a comparison of the result.

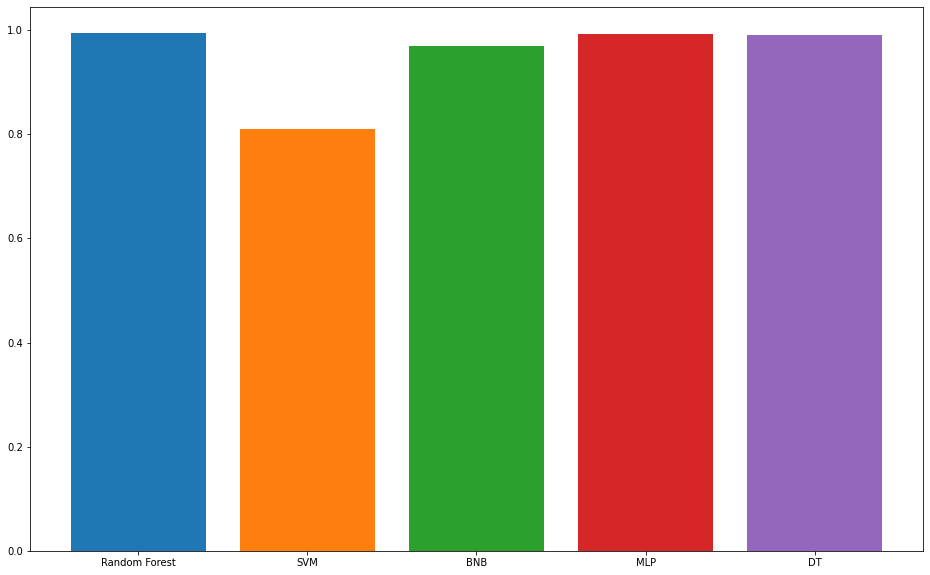


Graph I: Accuracy score for five Classification Algorithms

The accuracy scores for all the algorithms are being represented graphically in Graph 1. The values for these accuracies are:

* Random Forest Classifier Accuracy: 0.9940898345153665
* Support Vector Classifier Accuracy: 0.8262411347517731
* Bernoulli Naive Bayes Classifier Accuracy: 0.966903073286052
* Multi-layer Perceptron Accuracy: 0.9929078014184397
* Decision Tree Classifier Accuracy: 0.9905437352245863

By the formation of this graph, we have found that, the methodology we have proposed, measures the accuracy of the project at its maximum rate with the minimum effort of analysis. The second performance metric is the f1-score represented by Graph 2.



Graph 2: F1-Score for five classification algorithms

The values for the f1-scores for these algorithms are as follows:

* Random Forest Classifier F1-score: 0.994535519125683
* Support Vector Classifier F1-score: 0.8093385214007782
* Bernoulli Naive Bayes Classifier F1-score: 0.9689578713968958
* Multi-layer Perceptron F1-score: 0.9934354485776805
* Decision Tree Classifier F1-score: 0.9912280701754386

# Conclusion:

In this research paper, we have proposed a methodology for the detection of fake profile that are being penetrating in twitter’s network, which is a site of social networking. The focus of the proposed approach was to identify the effective functionalities, that could help us in detecting the fake and unauthorized profiles. During this research, we collected 22 attributes which were then subject to the series of experimentations, testified and validated various time to collect the best ones and then minimized attributes were weighted. The weighted attributes were then sent forward to be applied 5 classification algorithms. The way to the application of C.5 (five classifications.) algorithm was started when the attributes were filtered to seven attributes with high results of accuracy. And then used for the detection of fake profiles.

Although, we can declare that these attributes can effectively be used in the detection of faux accounts of the sites such as, Twitter. Also, we know that these attributes or characteristics can also detect the fake and unauthorized accounts of other social sites, such Facebook & Instagram, with minute changes, yet we need to stance to stand this claim, and must devise a proper dataset to prove that these attributes are helpful in detecting fake Facebook, Instagram and any social media accounts too. Also, we had studied the other method in our research, such the honey pot and machine learning algorithms, and analyzed the similarity measure for clone profile detection, yet we are satisfied with the application C.5 (five classification) algorithms on minimized weighted attributes. Moreover, we can see, that results we have obtained are completely accurate and can provide more accurate consequence while proceeding toward the detection process.

# References

1. BRÓDKA, P., SOBAS, M. & JOHNSON, H. (2014) Profile cloning detection in social networks. European Network Intelligence Conference, 2014. IEEE, 63-68.
2. CHATTERJEE, M (2019) Detection of Fake and Cloned Profiles in Online Social Networks.
3. Dr. Ponnurangam Kumaraguru (Chair), Dr Arun Balaji Buduru, Dr. Siddhartha Asthana (2019) Detecting Fake Profiles on Online Matrimony.
4. Nitika Kadam, Harish Patidar (2020) Social Media Fake Profile Detection Technique Based on Attribute Estimation and Content Analysis Method.
5. CONTI, M., POOVENDRAN, R. & SECCHIERO, M (2012) Fakebook: Detecting fake profiles in online social networks. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2012. IEEE, 1071-1078. International Journal of Control and Automation. 13, No. 1s, (2020), pp. 95-10095ISSN: 2005-4297IJCACopyright ⓒ2020SERSCDETECTING FAKE ACCOUNT ON SOCIAL MEDIA USING MACHINE LEARNING ALGORITHMS1, Jan,2020
6. CRESCI, S., DI PIETRO, R., PETROCCHI, M., SPOGNARDI, A. & TESCONI, M. 2014. A Fake Follower Story: improving fake accounts detection on Twitter. *IIT-CNR, Tech. Rep. TR-03*.
7. CRESCI, S., DI PIETRO, R., PETROCCHI, M., SPOGNARDI, A. & TESCONI, M. 2015. Fame for sale: Efficient detection of fake Twitter followers. *Decision Support Systems,* 80**,** 56-71.
8. DEVMANE, M. & RANA, N (2014) Detection, and prevention of profile cloning in online social networks. International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014) IEEE, 1-5.
9. ELAZAB, A. 2016. Fake accounts detection in twitter based on minimum weighted feature. *World*.
10. ELAZAB, A. & MAHMOUD, M. 2018. Fraud news detection for online social networks. *Artif Intell Syst Mach Learn,* 10**,** 177-182.
11. ERŞAHIN, B., AKTAŞ, Ö., KıLıNÇ, D. & AKYOL, C (2017) Twitter fake account detection. 2017 International Conference on Computer Science and Engineering (UBMK). IEEE, 388-392.
12. FIRE, M., GOLDSCHMIDT, R. & ELOVICI, Y. 2014. Online social networks: threats and solutions. *IEEE Communications Surveys & Tutorials,* 16**,** 2019-2036.
13. GRIER, C., THOMAS, K., PAXSON, V. & ZHANG, M (2010). @ spam: the underground on 140 characters or less. Proceedings of the 17th ACM conference on Computer and communications security. 27-37.
14. GUPTA, A. & KAUSHAL (2017), R. Towards detecting fake user accounts in Facebook. 2017 ISEA Asia Security and Privacy (ISEASP). IEEE, 1-6.
15. KHALIL, A., HAJJDIAB, H. & AL-QIRIM, N. 2017. Detecting fake followers in twitter: A machine learning approach. *International Journal of Machine Learning and Computing,* 7**,** 198-202.
16. KHAYYAMBASHI, M. R. & RIZI, F. S (2013) An approach for detecting profile cloning in online social networks. 7th International Conference on e-Commerce in Developing Countries: with focus on e-Security. IEEE, 1-12.
17. KIRUTHIGA, S. & KANNAN, A (2014) Detecting cloning attack in Social Networks using classification and clustering techniques. International Conference on Recent Trends in Information Technology, 2014. IEEE, 1-6.
18. KOLIAS, C., KAMBOURAKIS, G. & GRITZALIS, S. 2012. Attacks and countermeasures on 802.16: Analysis and assessment. *IEEE Communications Surveys & Tutorials,* 15**,** 487-514.
19. KONTAXIS, G., POLAKIS, I., IOANNIDIS, S. & MARKATOS, E. P (2011) Detecting social network profile cloning. IEEE international conference on pervasive computing and communications workshops (PERCOM Workshops), 2011. IEEE, 295-300.
20. PELED, O., FIRE, M., ROKACH, L. & ELOVICI, Y (2013) Entity matching in online social networks. International Conference on Social Computing, 2013. IEEE, 339-344.
21. WANG, A. H (2010). Don't follow me: Spam detection in twitter. 2010 international conference on security and cryptography (SECRYPT). IEEE, 1-10.
22. Asha Gowda Karegowda, A. S. Manjunath, and M.A. Jayaram (2010), "Comparative Study of Attribute Selection Using Gain Ratio," International Journal of Information Technology and Knowledge Management, vol. 2, no. 2, pp. 271-277.
23. Tatsunori Mori, Miwa Kikuchi, and Kazufumi Yoshida (2002), "reweighting Method based on Information Gain Ratio for Summarizing Documents retrieved by IR systems," Journal of Natural Language Processing, vol. 9, no. 4, pp. 3--32.
24. S. Cresci, M. Petrocchi, and R. Di Pietro (2014), "A criticism to Society (As seen by Twitter analystics)," in IEEE 34th international conference on distributes computing systems workshopS.
25. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi (2014), "A Fake Follower Story: improving fake accounts detection on Twitter".
26. Bas Van Den Beld. (2012, September) Stateofsearch.com. (Online). <http://goo.gl/YZbVf>
27. Manuel Fern\_andez Delgado, Eva Cernadas, Sen\_en Barro, and Dinani Amorim (2014), "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? “Journal of Machine Learning Research, vol.15, pp. 3133-3181.

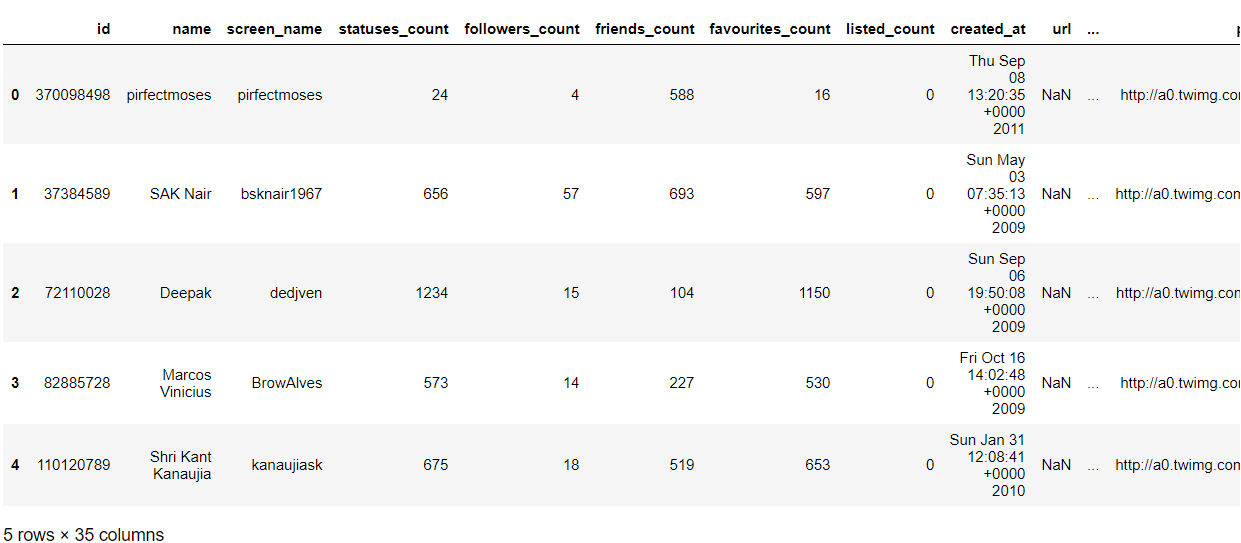
# Appendix

import pandas as pd  
import numpy as np  
from matplotlib import pyplot as plt  
import seaborn as sns  
  
*# Scikit-learn*  
from sklearn.model\_selection import train\_test\_split  
from sklearn.preprocessing import LabelEncoder  
from sklearn.ensemble import RandomForestClassifier  
from sklearn.svm import SVC  
from sklearn.naive\_bayes import BernoulliNB  
from sklearn.neural\_network import MLPClassifier  
from sklearn.tree import DecisionTreeClassifier  
from sklearn.metrics import confusion\_matrix, accuracy\_score, f1\_score  
  
plt.rcParams["figure.figsize"] = [16, 10]

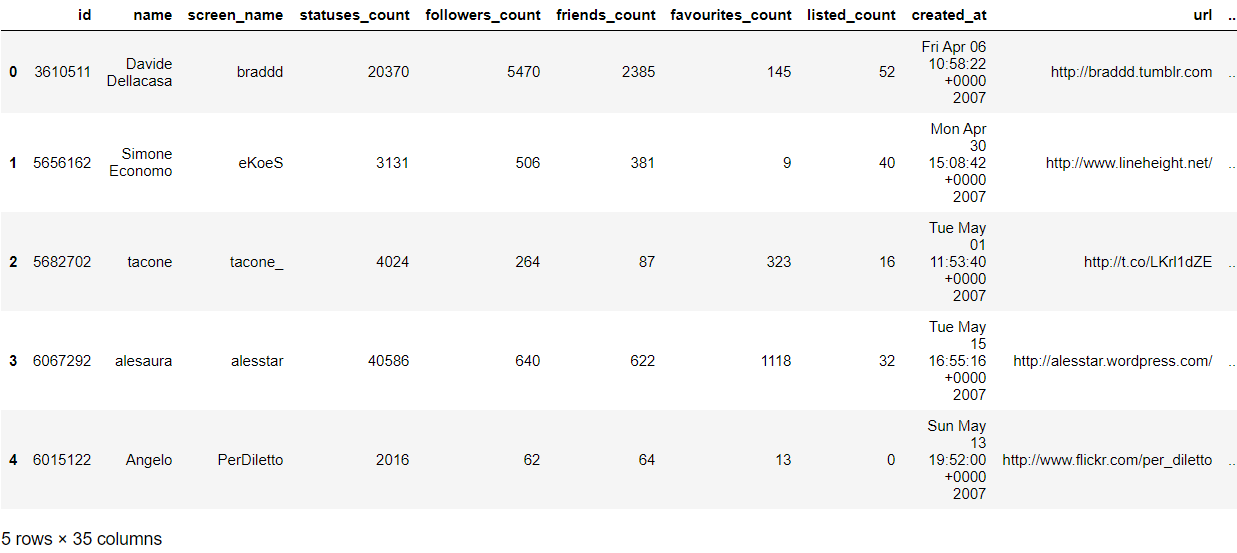
**Loading Data**

fake = pd.read\_csv('DATASET/fusers.csv')  
legit = pd.read\_csv('DATASET/users.csv')

fake.head()



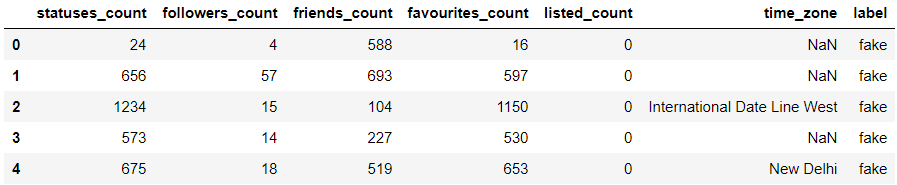
legit.head()



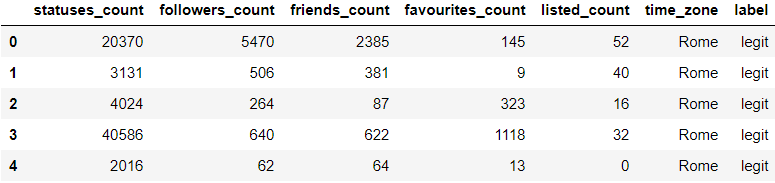
**Removing unnecessary columns**

fake = fake.drop(columns=["id", "name", "screen\_name", "created\_at", "lang", "location", "default\_profile", "default\_profile\_image", "geo\_enabled", "profile\_image\_url", "profile\_banner\_url", "profile\_use\_background\_image", "profile\_background\_image\_url\_https", "profile\_text\_color", "profile\_image\_url\_https", "profile\_sidebar\_border\_color", "profile\_background\_tile", "profile\_sidebar\_fill\_color", "profile\_background\_image\_url", "profile\_background\_color", "profile\_link\_color", "utc\_offset", "protected", "verified", "dataset", "updated", "description", "url"], axis=1)  
legit = legit.drop(columns=["id", "name", "screen\_name", "created\_at", "lang", "location", "default\_profile", "default\_profile\_image", "geo\_enabled", "profile\_image\_url", "profile\_banner\_url", "profile\_use\_background\_image", "profile\_background\_image\_url\_https", "profile\_text\_color", "profile\_image\_url\_https", "profile\_sidebar\_border\_color", "profile\_background\_tile", "profile\_sidebar\_fill\_color", "profile\_background\_image\_url", "profile\_background\_color", "profile\_link\_color", "utc\_offset", "protected", "verified", "dataset", "updated", "description", "url"], axis=1)

fake.head()



legit.head()



**Changing all NaN to 0**

fake = fake.fillna(0)  
legit = legit.fillna(0)

fake.info()

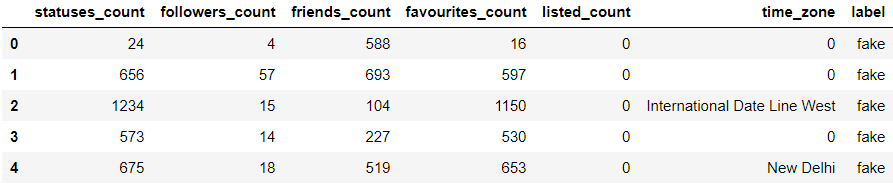
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1337 entries, 0 to 1336  
Data columns (total 7 columns):  
 # Column Non-Null Count Dtype   
--- ------ -------------- -----   
 0 statuses\_count 1337 non-null int64   
 1 followers\_count 1337 non-null int64   
 2 friends\_count 1337 non-null int64   
 3 favourites\_count 1337 non-null int64   
 4 listed\_count 1337 non-null int64   
 5 time\_zone 1337 non-null object  
 6 label 1337 non-null object  
dtypes: int64(5), object(2)  
memory usage: 73.2+ KB

legit.info()

<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1481 entries, 0 to 1480  
Data columns (total 7 columns):  
 # Column Non-Null Count Dtype   
--- ------ -------------- -----   
 0 statuses\_count 1481 non-null int64   
 1 followers\_count 1481 non-null int64   
 2 friends\_count 1481 non-null int64   
 3 favourites\_count 1481 non-null int64   
 4 listed\_count 1481 non-null int64   
 5 time\_zone 1481 non-null object  
 6 label 1481 non-null object  
dtypes: int64(5), object(2)  
memory usage: 81.1+ KB

**Creating merged dataset from legit and fake profile dataset.**

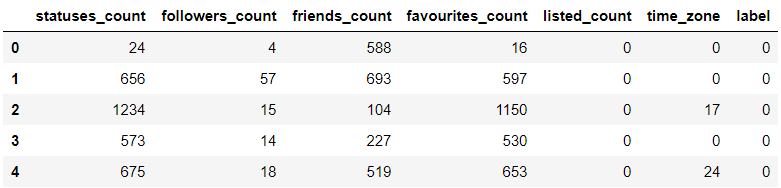
data = [fake, legit]  
dataset = pd.concat(data)  
dataset.head()



**Encoding Features with "Object" Datatype**

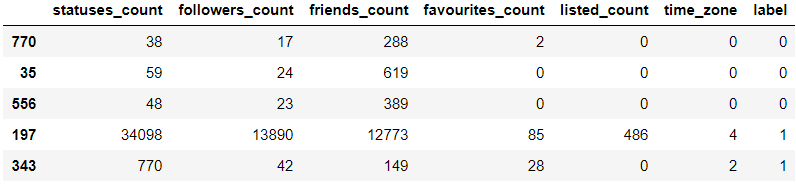
le = LabelEncoder()

dataset['time\_zone'] = le.fit\_transform(dataset['time\_zone'].astype(str))  
dataset['label'] = le.fit\_transform(dataset['label'].astype(str))  
dataset.head()



**Shuffling Data Instances to Mix Up the Fake and Legit Instances**

dataset = dataset.sample(frac=1)  
dataset.head()



**Splitting Label and Features**

y = dataset['label']  
X = dataset.drop(columns='label')

**Splitting Dataset into Train and Test**

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.3, random\_state=42)

**Initializing Models**

rfc = RandomForestClassifier()  
svm = SVC()  
bnb = BernoulliNB()  
mlp = MLPClassifier(max\_iter=250)  
dt = DecisionTreeClassifier()

**Training the Models**

rfc.fit(X\_train, y\_train)  
svm.fit(X\_train, y\_train)  
bnb.fit(X\_train, y\_train)  
mlp.fit(X\_train, y\_train)  
dt.fit(X\_train, y\_train)

DecisionTreeClassifier()

**Predicting with the Models**

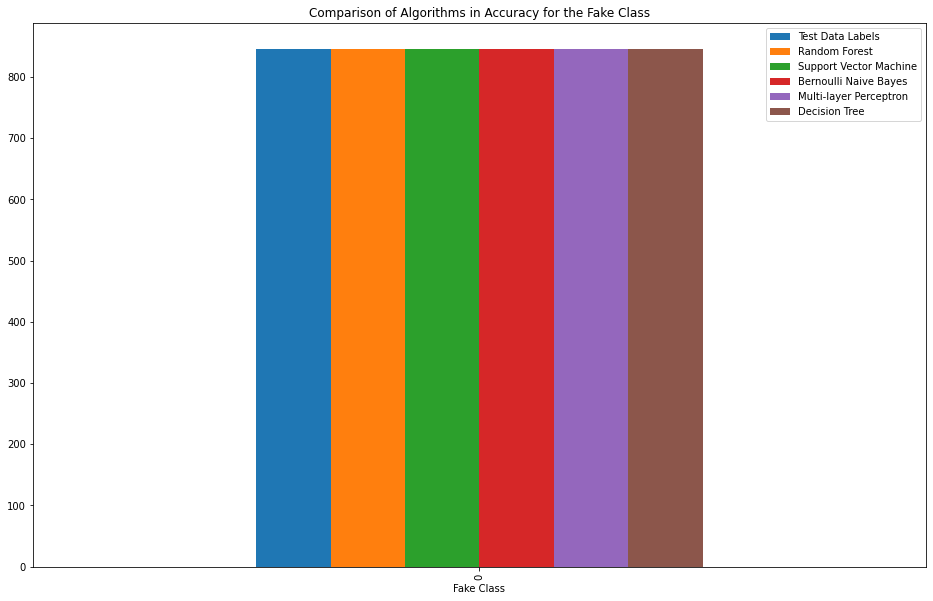
rfc\_pred = rfc.predict(X\_test)  
svm\_pred = svm.predict(X\_test)  
bnb\_pred = bnb.predict(X\_test)  
mlp\_pred = mlp.predict(X\_test)  
dt\_pred = dt.predict(X\_test)

**Performance Metrics**

fgraph = pd.DataFrame([[0, len(y\_test == 0), len(rfc\_pred == 0), len(svm\_pred == 0), len(bnb\_pred == 0), len(mlp\_pred == 0), len(dt\_pred == 0)]],  
 columns=['Fake Class', 'Test Data Labels', 'Random Forest', 'Support Vector Machine', 'Bernoulli Naive Bayes', 'Multi-layer Perceptron', 'Decision Tree'])

fgraph.plot(x='Fake Class', kind='bar', stacked=False, title='Comparison of Algorithms in Accuracy for the Fake Class')

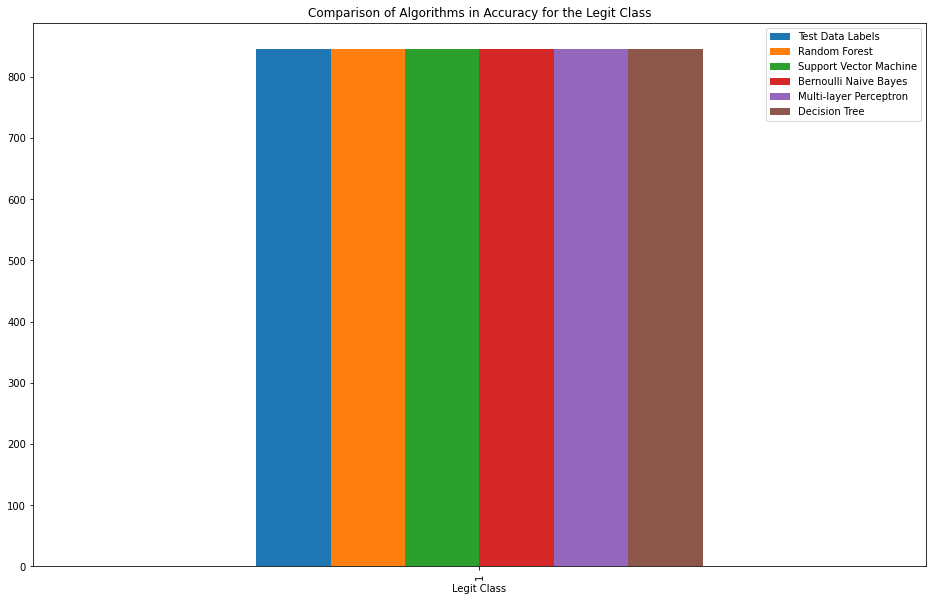
<AxesSubplot:title={'center':'Comparison of Algorithms in Accuracy for the Fake Class'}, xlabel='Fake Class'>



lgraph = pd.DataFrame([[1, len(y\_test == 1), len(rfc\_pred == 1), len(svm\_pred == 1), len(bnb\_pred == 1), len(mlp\_pred == 1), len(dt\_pred == 1)]],  
 columns=['Legit Class', 'Test Data Labels', 'Random Forest', 'Support Vector Machine', 'Bernoulli Naive Bayes', 'Multi-layer Perceptron', 'Decision Tree'])

lgraph.plot(x='Legit Class', kind='bar', stacked=False, title='Comparison of Algorithms in Accuracy for the Legit Class')

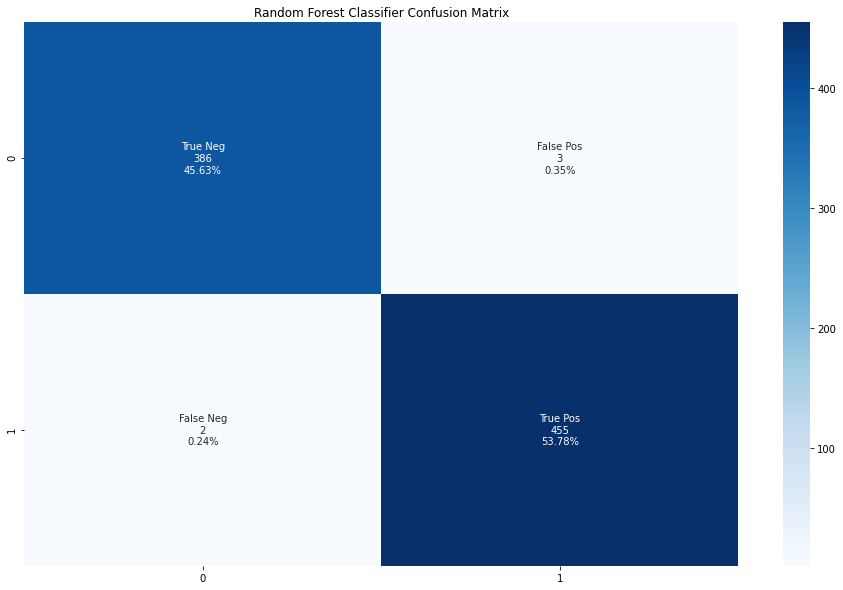
<AxesSubplot:title={'center':'Comparison of Algorithms in Accuracy for the Legit Class'}, xlabel='Legit Class'>



rfc\_cm = confusion\_matrix(y\_test, rfc\_pred)  
svm\_cm = confusion\_matrix(y\_test, svm\_pred)  
bnb\_cm = confusion\_matrix(y\_test, bnb\_pred)  
mlp\_cm = confusion\_matrix(y\_test, mlp\_pred)  
dt\_cm = confusion\_matrix(y\_test, dt\_pred)

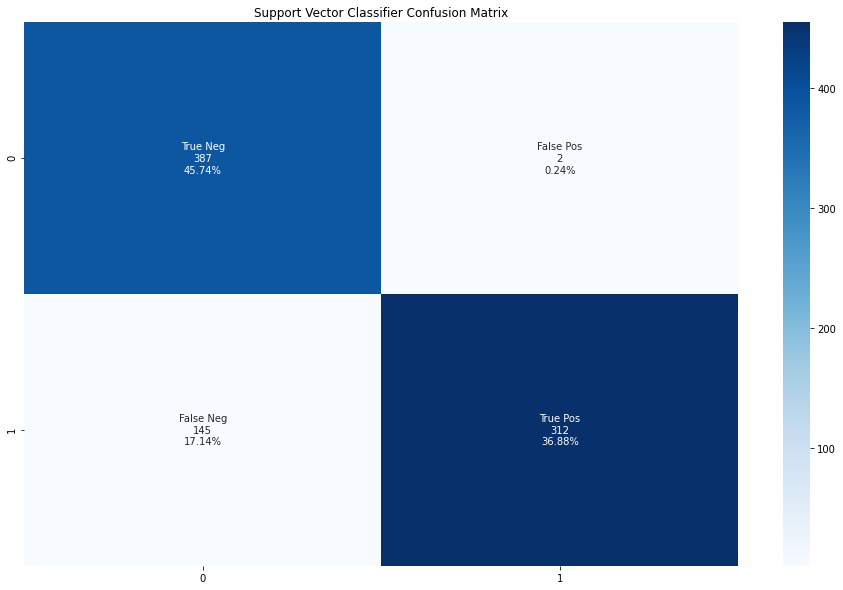
group\_names = ['True Neg', 'False Pos', 'False Neg', 'True Pos']  
group\_counts = ["{0:0.0f}".format(value) **for** value in rfc\_cm.flatten()]  
group\_percentages = ["{0:.2%}".format(value) **for** value in rfc\_cm.flatten()/np.sum(rfc\_cm)]  
labels = [f"{v1}\n{v2}\n{v3}" **for** v1, v2, v3 in zip(group\_names, group\_counts, group\_percentages)]  
labels = np.asarray(labels).reshape(2, 2)  
plt.title('Random Forest Classifier Confusion Matrix')  
sns.heatmap(rfc\_cm, annot=labels, fmt='', cmap='Blues')

<AxesSubplot:title={'center':'Random Forest Classifier Confusion Matrix'}>



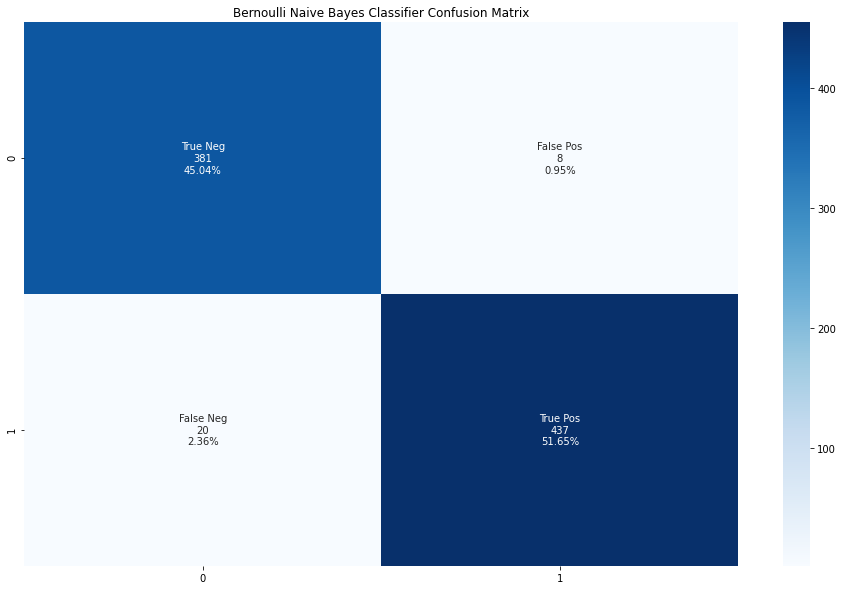
group\_names = ['True Neg', 'False Pos', 'False Neg', 'True Pos']  
group\_counts = ["{0:0.0f}".format(value) **for** value in svm\_cm.flatten()]  
group\_percentages = ["{0:.2%}".format(value) **for** value in svm\_cm.flatten()/np.sum(svm\_cm)]  
labels = [f"{v1}\n{v2}\n{v3}" **for** v1, v2, v3 in zip(group\_names, group\_counts, group\_percentages)]  
labels = np.asarray(labels).reshape(2, 2)  
plt.title('Support Vector Classifier Confusion Matrix')  
sns.heatmap(rfc\_cm, annot=labels, fmt='', cmap='Blues')

<AxesSubplot:title={'center':'Support Vector Classifier Confusion Matrix'}>



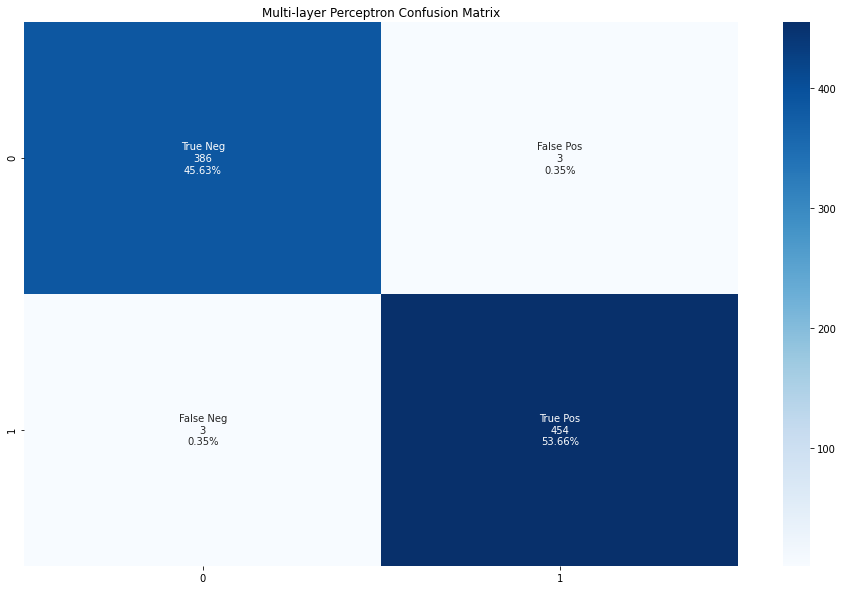
group\_names = ['True Neg', 'False Pos', 'False Neg', 'True Pos']  
group\_counts = ["{0:0.0f}".format(value) **for** value in bnb\_cm.flatten()]  
group\_percentages = ["{0:.2%}".format(value) **for** value in bnb\_cm.flatten()/np.sum(bnb\_cm)]  
labels = [f"{v1}\n{v2}\n{v3}" **for** v1, v2, v3 in zip(group\_names, group\_counts, group\_percentages)]  
labels = np.asarray(labels).reshape(2, 2)  
plt.title('Bernoulli Naive Bayes Classifier Confusion Matrix')  
sns.heatmap(rfc\_cm, annot=labels, fmt='', cmap='Blues')

<AxesSubplot:title={'center':'Bernoulli Naive Bayes Classifier Confusion Matrix'}>



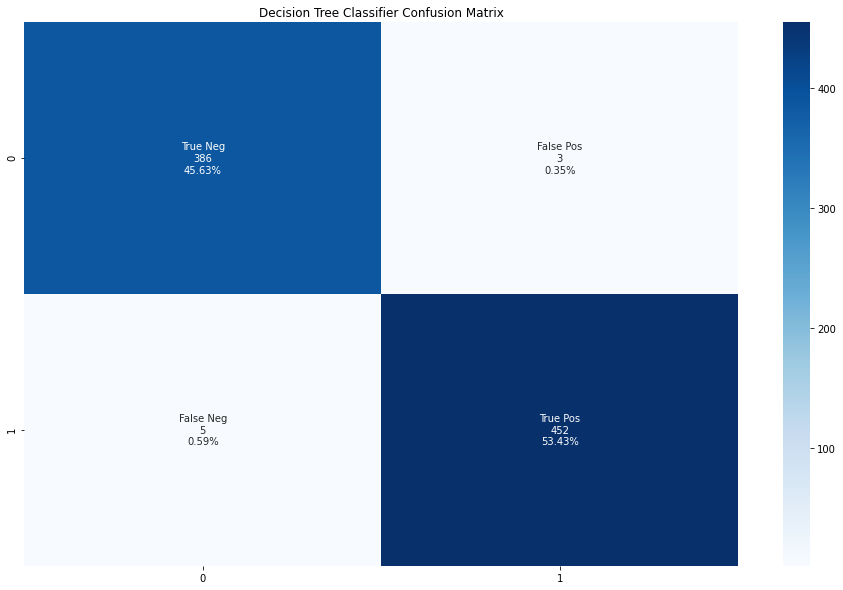
group\_names = ['True Neg', 'False Pos', 'False Neg', 'True Pos']  
group\_counts = ["{0:0.0f}".format(value) **for** value in mlp\_cm.flatten()]  
group\_percentages = ["{0:.2%}".format(value) **for** value in mlp\_cm.flatten()/np.sum(mlp\_cm)]  
labels = [f"{v1}\n{v2}\n{v3}" **for** v1, v2, v3 in zip(group\_names, group\_counts, group\_percentages)]  
labels = np.asarray(labels).reshape(2, 2)  
plt.title('Multi-layer Perceptron Confusion Matrix')  
sns.heatmap(rfc\_cm, annot=labels, fmt='', cmap='Blues')

<AxesSubplot:title={'center':'Multi-layer Perceptron Confusion Matrix'}>



group\_names = ['True Neg', 'False Pos', 'False Neg', 'True Pos']  
group\_counts = ["{0:0.0f}".format(value) **for** value in dt\_cm.flatten()]  
group\_percentages = ["{0:.2%}".format(value) **for** value in dt\_cm.flatten()/np.sum(dt\_cm)]  
labels = [f"{v1}\n{v2}\n{v3}" **for** v1, v2, v3 in zip(group\_names, group\_counts, group\_percentages)]  
labels = np.asarray(labels).reshape(2, 2)  
plt.title('Decision Tree Classifier Confusion Matrix')  
sns.heatmap(rfc\_cm, annot=labels, fmt='', cmap='Blues')

<AxesSubplot:title={'center':'Decision Tree Classifier Confusion Matrix'}>

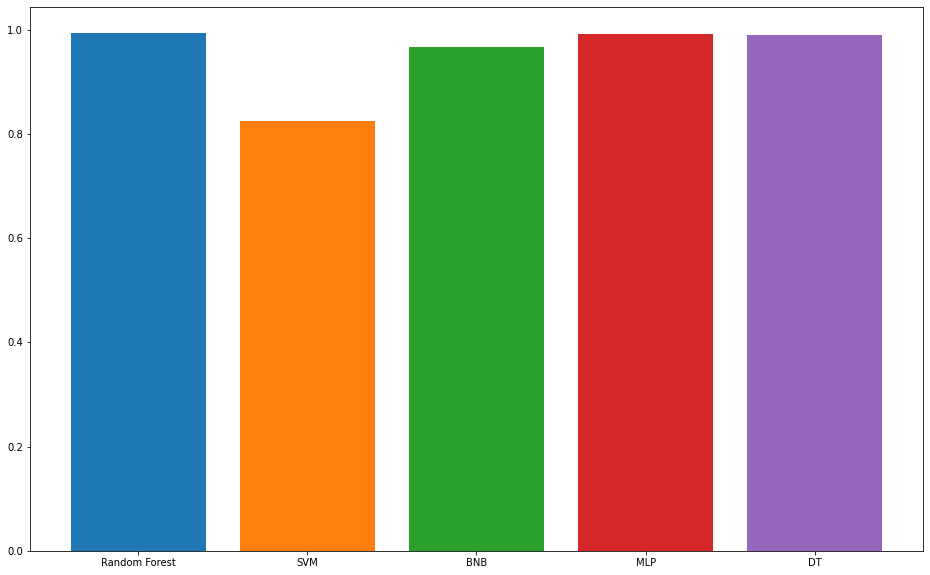


print("Random Forest Classifier Accuracy: ", accuracy\_score(y\_test, rfc\_pred))  
print("Support Vector Classifier Accuracy: ", accuracy\_score(y\_test, svm\_pred))  
print("Bernoulli Naive Bayes Classifier Accuracy: ", accuracy\_score(y\_test, bnb\_pred))  
print("Multi-layer Perceptron Accuracy: ", accuracy\_score(y\_test, mlp\_pred))  
print("Decision Tree Classifier Accuracy: ", accuracy\_score(y\_test, dt\_pred))

Random Forest Classifier Accuracy: 0.9940898345153665  
Support Vector Classifier Accuracy: 0.8262411347517731  
Bernoulli Naive Bayes Classifier Accuracy: 0.966903073286052  
Multi-layer Perceptron Accuracy: 0.9929078014184397  
Decision Tree Classifier Accuracy: 0.9905437352245863

plt.bar('Random Forest', accuracy\_score(y\_test, rfc\_pred))  
plt.bar('SVM', accuracy\_score(y\_test, svm\_pred))  
plt.bar('BNB', accuracy\_score(y\_test, bnb\_pred))  
plt.bar('MLP', accuracy\_score(y\_test, mlp\_pred))  
plt.bar('DT', accuracy\_score(y\_test, dt\_pred))

<BarContainer object of 1 artists>

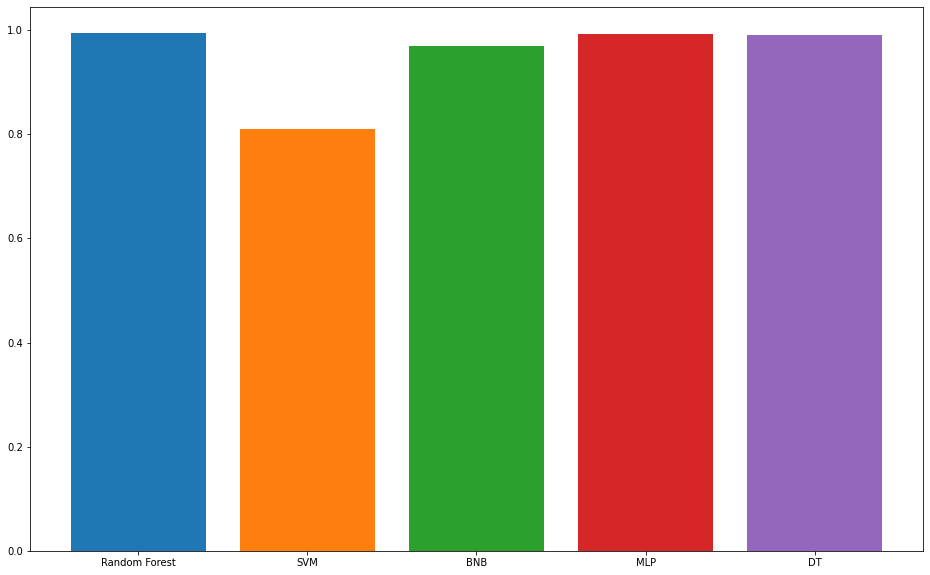


print("Random Forest Classifier F1-score: ", f1\_score(y\_test, rfc\_pred))  
print("Support Vector Classifier F1-score: ", f1\_score(y\_test, svm\_pred))  
print("Bernoulli Naive Bayes Classifier F1-score: ", f1\_score(y\_test, bnb\_pred))  
print("Multi-layer Perceptron F1-score: ", f1\_score(y\_test, mlp\_pred))  
print("Decision Tree Classifier F1-score: ", f1\_score(y\_test, dt\_pred))

Random Forest Classifier F1-score: 0.994535519125683  
Support Vector Classifier F1-score: 0.8093385214007782  
Bernoulli Naive Bayes Classifier F1-score: 0.9689578713968958  
Multi-layer Perceptron F1-score: 0.9934354485776805  
Decision Tree Classifier F1-score: 0.9912280701754386

plt.bar('Random Forest', f1\_score(y\_test, rfc\_pred))  
plt.bar('SVM', f1\_score(y\_test, svm\_pred))  
plt.bar('BNB', f1\_score(y\_test, bnb\_pred))  
plt.bar('MLP', f1\_score(y\_test, mlp\_pred))  
plt.bar('DT', f1\_score(y\_test, dt\_pred))

<BarContainer object of 1 artists>

**