Fake User Profile Detection using Machine Learning- A Review

Dipanshu Sharma[1], Er. Ravinder Singh Madan[2]

1, M.Tech Scholar, IEC University Solan, Himachal Pradesh
2, HOD – CSE, IEC University Solan, Himachal Pradesh

Abstract— **There are several well-known websites, such as Facebook, Twitter, and Instagram, that fall under the category of social networking platforms that let users to engage with one another online. Users are consistently having conversations with people their own age via a number of platforms, including social networking websites. By using these social networks as a mode of communication, they exchange information that is both private and open to the general public. The appeal of social networking websites has led to a significant increase in the number of individuals making use of the services provided by these websites. This popularity makes things more difficult for the websites because it encourages users to establish phoney profiles on the platforms they use, which makes things more difficult for the websites. People who are in charge of phoney accounts steal the personal information of other users and then distribute it across other social media platforms. They do this by taking the information from the accounts that they control. In the following paragraphs, we will go through a few different methods that may be used to spot fake user accounts.**

Keywords—Fake user; profile; machine; Learning

I. Introduction

It is becoming more popular for people to utilise social networking websites on a regular basis in order to maintain contact with their personal and professional ties, a practise that is already fairly prevalent among individuals. In addition, this practise is growing more ubiquitous. The fact that users of these types of websites constantly share information about themselves and the activities that they get up to in the course of their everyday life is one of the most enticing characteristics of these kinds of websites. Between the years 2006 and 2022, both Facebook and Twitter saw a significant increase in the number of users actively participating in the platform's activities, in addition to a growth in the popularity of the platform [1]. During this same time period, the number of users who actively participated in the platform's activities also increased substantially. Users of any particular platform have the ability to communicate with one another, connect with their friends, and exchange a variety of information with one another. This information might be personal or it can be societal or economic or political or educational or business-related or anything else. In addition to this, they are able to deliver daily updates by means of the exchange of images, films, and a range of other kinds of communication through a wide variety of various types of media [2]. On the other hand, there are some individuals who do not consult these websites with a neutral and objective frame of mind. These people should be avoided at all costs. They do this by creating bogus accounts on a wide array of social networking platforms. Because fake accounts don't have a real identity of their own, we could refer to the people who use them as attackers in the future. This is because fake accounts don't have a real identity of their own. This is due to the fact that they provide a safety concern. Con artists utilise inaccurate information or statistics that belong to a real person in order to create a phoney account in the name of that person elsewhere on the globe. This allows the con artists to pose as the actual person. The attacker was able to earn the confidence of other users and exercise influence over them by utilising these bogus identities to propagate deceptive material that was then spread using these fictitious identities [3]. Protecting the privacy of highly sensitive user information is now one of the most serious difficulties that social media networks are working hard to find solutions for. In order to identify false accounts on social networking sites, a variety of different machine learning algorithms, such as Neural Network (NN), Naive Bayes, Markov Model, and Bayesian Network, have been developed [4]. The issue of cyberbullying prompted the development of these tactics as a means of addressing the problem. Recent studies have demonstrated that using these tactics results in improved results when seeking to detect bogus accounts. A neural network is a kind of computer system that processes information by using a group of distinct components that are all linked to one another. These components are all connected to one another. It approaches the task of making decisions in a manner that is analogous to the way in which a human brain processes information [5-7]. Classification is an example of an application that makes use of support vector machines, which are more often known to by their acronym, SVM. These are examples of approaches that may be grouped under the umbrella term of supervised machine learning. To get started with the categorizing process, the first thing that needs to be done is to locate the hyperplane. This will allow the procedure to begin. Because of their capacity to comprehend a significant quantity of random input, neural networks and support vector machines are especially well-suited for the task of recognizing false identities on social networking platforms. The answer that we provide to this enquiry will be dependent on a number of different elements that are related to the account. The Bayes theorem served as the basis for the building of a classifier that was later dubbed the Naive Bayes classifier after the theory that it was based on. It achieves this objective by making predictions on the probability that a specific variable belongs to a particular category [8-10]. This allows it to make more accurate forecasts.

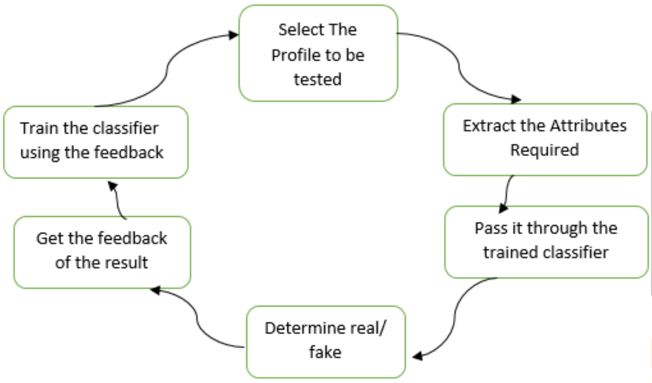
Select The Profile to be tested
Extract the Attributes Required
Pass it through the trained classifier
Determine real/ fake
Get the feedback of the result
Train the classifier using the feedback

Fig. 1 Process if Fake User Detection

Adopting a few strategies that have been shown to be successful in the past is one way in which the reputation of a social media account may be improved. Other techniques include: It is possible to achieve this objective through the use of a variety of different approaches, some of which include the application of bots, the acquisition of social metrics such as likes, comments, and followers through the purchase of said metrics, and the utilisation of particular platforms or networks that enable users to trade said metrics. All of these approaches are viable options. There is more than one way to achieve this objective, and all of these approaches are valid possibilities. Each of these strategies is an option that may be pursued if one is looking for a workable solution to the problem at hand. When speaking about a piece of software, the word "bot" is often reserved for referring to a programme that, when linked to the internet, would automatically carry out activities that are repetitious. This kind of application is known as a "task bot." About 95 million Instagram accounts are being handled by bots, according to the conclusions of research that was carried out by Ghost Data in the year 2018. The results of the investigation provided the foundation for these discoveries. The year 2016 was the first time when bots produced more internet traffic than genuine people did. This phenomenon occurred for the first time in 2016. The year 2016 saw the onset of this phenomenon for the very first time. This phenomenon occurred for the very first time in 2016, the year it was first observed. Additionally, in order to advertise their items, merchants may establish fictional accounts for themselves on social media platforms like Instagram and Facebook. This gives them the ability to sell likes and followers in a very short period of time. If, for example, you hired the services of a firm named IDigic, they would give you 50,000 followers on their website for a fee that was somewhere around $250, but you would have to pay them for the privilege. Using their firm would come with a number of benefits, including this one. Everything that has been discussed up to this point is an example of inorganic engagement, which is another phrase for "phoney participation." To put this another way, the term "fake engagement" refers to a broad array of behaviours that are shown by robots. These behaviours include things like smiling and talking. Following accounts, like and commenting on content, and submitting articles and stories are all examples of this kind of behaviour. The extra practise of gathering analytics on social media networks is a practise that runs the danger of being labelled as "fake engagement." [9] The detection of individuals who intentionally increase the size of their accounts is crucial for a variety of reasons, including the fact that such people should be avoided. It causes businesses to pay users more for advertising than the advertising is worth; it causes advertisers to reach the wrong audiences; it causes recommendation systems to function in an inefficient manner; and it makes it more difficult to acquire services and products of a high quality. Some of these reasons include the following: it causes it to be more difficult to acquire services and products of a high quality. The identification of fraudulent accounts and automated accounts, which are often referred to as bot accounts in certain areas, are two aspects of the problem of false interaction that are distinct from one another but are connected to one another. The term "bot account" may also be used to refer to accounts that are fully automated. Phony accounts, sometimes known as sock puppet accounts, may be found in this category. Sock puppet accounts are a term that is occasionally used to refer to fake accounts. As was just stated, bot accounts are user profiles that engage in automated behaviours such as following other users and enjoying content generated by audiences with similar interests in order to artificially inflate popularity metrics. These behaviours can include following other users and liking content generated by audiences with similar interests. Following other users and favoriting material created by audiences with similar interests are two examples of these behaviours. Two examples of these behaviours include the practise of following other users and the favoriting of content that was made by audiences who have similar interests. Following other users and watching material created by audiences with similar interests are two examples of behaviours that might fall under this category. Both of these instances can be found on social media. Twitter is a great place to participate in any of these two pastimes. When a person creates several social media profiles with the sole intention of boosting the metrics of one of their other accounts—particularly an account for which the user has paid for the service—the profiles in question are referred to as phoney. Ghost accounts are another name for fake or bogus accounts. There is a good chance that the number of likes, shares, and followers is included into all of these statistics. It is also conceivable to refer to these individuals as "phoney followers," which would attract even more attention to the issue that is now being addressed. This is one of the possibilities. There is also the option of referring to these individuals as "phoney followers." The most important distinction between automated accounts and fake accounts is that automated accounts improve the metrics of the account itself, whereas fake accounts improve the metrics of other users and contribute to an unhealthy environment on social media. In other words, automated accounts improve the metrics of the account itself, whereas fake accounts improve the metrics of other users. The metrics of the account itself may also be improved via the use of automated accounts. In addition, criminals are more likely to use automated accounts as opposed to regular ones. It is already fraught with danger to acquire personally identifiable information without proper authorization; however, the use of bots and phoney accounts raises the stakes of this risk to a whole new level. Bots are computer programmes that can gather information about a user even if the user is unaware that information is being collected about them. This is because bots can collect information even if the user is logged out of their account. This is due to the fact that bots are able to gather data even if the user is not currently signed into their account. This is because bots are able to acquire information even if the user is not actively logged into their account, which is the reason for this result. This behaviour occurs as a result of the fact that bots have the potential of collecting information even when the user is not signed into their account. This is the explanation for why this phenomenon occurs. If a person visits a website that maintains a record of the information that they enter, then there is a risk that they may experience something similar to what has been described in this article. Web scraping is the action that is being carried out in this scenario, which is a phrase that accurately reflects the technique that is being employed. The activity that is being carried out is known as web scraping. The current situation is far more precarious than it would have been under any other set of circumstances since the law not only allows for but also actively encourages the behaviour in question. The administrators of a social networking site have the option of either keeping bots hidden or enabling them to appear as false friend requests so that bots can access private information on the site. Both of these options allow bots to access information that is normally protected from public view. Both of these choices provide bots the ability to access information that is ordinarily shielded from the gaze of the general public.

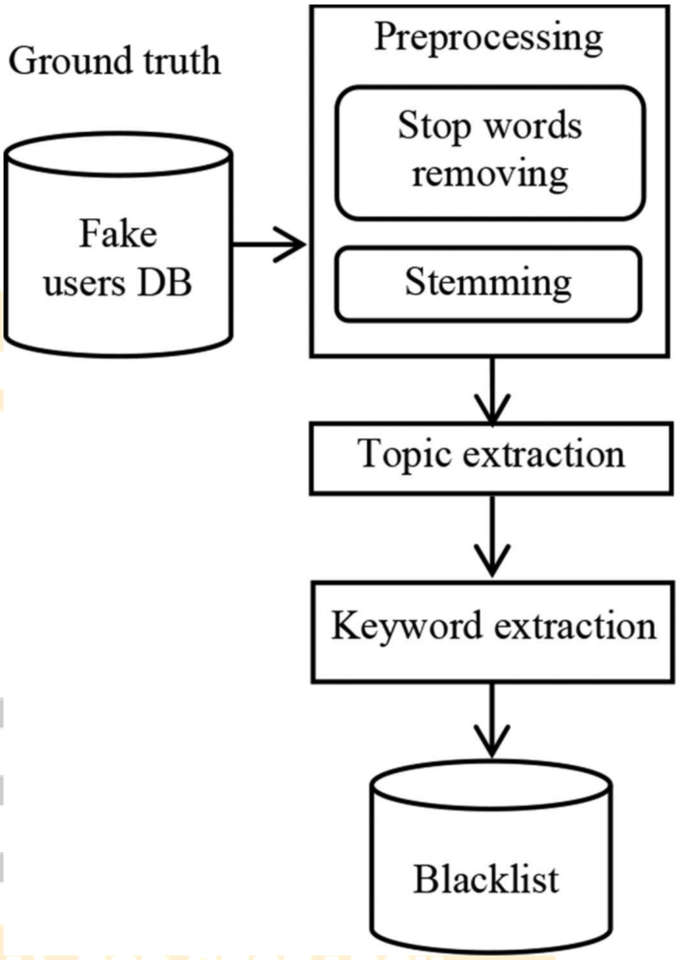

Fig. 2 Fake User Detection

The typical individual is now more vulnerable to crimes such as data breaches and possible identity theft as a result of the ever-increasing dependency on computer technology that exists in this day and age of the digital revolution. This is as a result of the fact that the digital revolution has reached its peak level of activity. It is conceivable that these attacks will take place with no previous notice whatsoever, and the persons whose private information has been stolen will often never realise what has taken place to them. However, it is possible that they will. At this moment in time, there are not a whole lot of reasons for social networks to raise the level of data protection that they provide for their members. This is because of the nature of the internet. In the not-too-distant future, this is an issue that has to be resolved as soon as possible. The vast majority of the time, social media networks such as Facebook and Twitter are the ones to be impacted by these kinds of security flaws and vulnerabilities. Additionally, they are able to aim their attacks towards monetary institutions such as banks and other types of financial enterprises. This article describes the different methods that have been used in the past to detect fake user profiles and the relevance of doing so. Additionally, the article explains why it is important to recognise false user profiles.

II. Review of PRevious Work

Hajdu et al. [1] The author takes use of machine learning, and more specifically an artificial neural network, in order to identify whether or not a friend request submitted on Facebook is genuine. In addition, the Author provides a synopsis of the schools and institutions that have relevant information about the topic under discussion. In addition to this, the author goes into detail about the sigmoid function, as well as the procedure of assigning weights and using them in the model. In conclusion, the Author takes into account the needs of the social network page, which are of the highest relevance in the solution that has been presented. [1]

Kulkarni et al. [2] Participation in online social networks is now intimately linked up with the day-to-day social life of everyone living in the modern world. Because of the impact that these websites have had on the writers' routines, the authors' social lives have been dramatically reshaped as a consequence of this influence. It is now more simpler to make new acquaintances and to keep in touch with those that one already knows than it was in times past. However, as a direct consequence of their meteoric rise in popularity, a plethora of brand-new challenges have emerged, such as dishonest persons, fake accounts, and online impersonation. These problems have arisen as a direct consequence of their skyrocketing popularity over the last several years. There is not, and never will be, a workable solution that can be implemented to address these problems, and there never will be. The author of this research study makes use of an artificial neural network to recognise fake profiles in a manner that is both efficient and automated. This approach is presented as part of the study. The author analyses the likelihood that a friend request on Facebook comes from a real person as opposed to a bot and provides his findings in the following paragraphs.

Khaled et al. [3] People of this age are shown a growing interest in engaging in activities that take place on online social networks (OSNs). As a consequence of this, the users' work and private lives are becoming more intertwined as a result of their usage of these services. They communicate with one another via the utilisation of online social networks (OSNs), plan events, exchange news, and even establish their very own online businesses with the assistance of OSNs. Cybercriminals and imposters whose primary goals are to steal personal data, spread false news, and engage in other types of disruptive behaviour have become increasingly interested in the rapid growth of online social networks (OSNs) and the vast amount of personal data that their users generate. This has attracted the attention of cybercriminals and imposters. On the other hand, researchers have begun looking at efficient methods to identify anomalous behaviours and phoney accounts based on the characteristics of accounts and classification algorithms. These approaches are intended to be used in conjunction with the characteristics of accounts. This is in contrast to the scenario that existed before, in which no study of this kind had been carried out. On the other hand, some aspects of the account that have been explored either make a contribution that is negative to the overall results or have no influence at all. In addition, the use of categorization algorithms that function in isolation from one another does not always provide results that are satisfactory.

Patel et al. [4] to protect oneself against unwanted messages known as spam, as well as unlawful acts and cyberbullying, both of which are often conducted out through fake accounts. These actions represent a threat to the privacy limits that are maintained by the communities that are a part of social networking sites and pose a threat to such communities. These bogus accounts are the ones responsible for spreading incorrect information across a variety of social networks. In this particular field of research, a significant lot of work is being done in order to recognise fake profiles, duplicate accounts, spam accounts, and bot accounts. Finding the vast majority of the false accounts required the employment of an algorithm that made use of machine learning, which was successful in doing so.

Singh et al. [5] Because of the broad acceptance of various social media platforms in the contemporary world and the extensive use of these platforms on a daily basis, social media has become a vital aspect of the life of writers. The number of individuals who use social media platforms for harmful purposes is growing at a rate that is far faster than it was in the past. This is because more people are using these platforms. On more than one occasion, successful differentiation of fake accounts has been accomplished by the use of machine adaption procedures. On the other hand, the quantity of research labour required to identify fake characters created by people is not very extensive. This is due to the fact that counterfeit characters often utilise pre-existing characters as inspiration. In the case of bots, machine learning models take into consideration a variety of different aspects in order to compute the proportion of an account's number of followers to the number of friends that the account has on various social networking sites (SOCIAL MEDIA PLATFORMSs). No account's rights were infringed upon in any way during the process of compiling this information, and the ratio of an account's friends to the number of its followers is readily apparent on the profile of the account that is in issue.

Bhattacharya and others, [6] [7] These days, there are around one billion individuals who use social networking sites in order to contact with other people from all over the globe and to trade images, opinions, and other information through the internet. These people may do this by using the internet. During this trying time brought on by pandemics, authors are depending on social media to a higher degree than they ever have before, which has resulted in a rise in the amount of engagement authors have with social media. It has provided us with information regarding the global spread of the coronavirus, provided us with amusement, assisted us in maintaining relationships with friends and acquaintances who reside in different locations, and even assisted many individuals in establishing and developing their own online small businesses. The fact that social media platforms are available online brings with it not only a great number of advantages for users, but also a great number of drawbacks that users must overcome in order to fully take advantage of the many benefits that these platforms provide. Impersonation and the creation of fraudulent accounts are two problems that have become more widespread on social media platforms over the last several years. These problems might have been caused by cyborgs, bots made by humans, or bots generated by computers, all of which are possibilities. These accounts are created with the sole purpose of wreaking havoc and raising a commotion in the community. In addition, there is no workable remedy that can be found that can be implemented for such a challenging circumstance that can be considered.

Aydin et al. [7] It is impossible to quantify the extent to which the development of social networks has altered the lives of a great number of people living in the contemporary world. As a consequence of people using social networks, many new activities have begun to be carried out. These activities include communication, promotion, advertising, news gathering, and the establishment of agendas. There are a lot of malevolent accounts that are used for a variety of nefarious objectives, including the spread of false information and the creation of agendas. These accounts may be found on Twitter. When it comes to social networks, this is one of the most fundamental issues that might arise at any point in time. The identification of bogus accounts is of utmost relevance since this is a direct result of its importance. In this research, machine learning-based technologies were used to search for false accounts that had the potential to trick consumers. These accounts had the capacity to mislead people. These accounts had the possibility of causing damage. In order to accomplish this goal, the dataset that had just been produced was first pre-processed, and then methods from the field of machine learning were used in order to identify bogus accounts. When detecting whether or not accounts are fraudulent, the process makes use of a broad variety of computer techniques. Some of these techniques include decision trees, logistic regression, and support vector machines, amongst others. The data reveal that the logistic regression approach is superior to the other methods that are reviewed and compared when it comes to classification, and there are many distinct procedures that are assessed and contrasted.

Harris et al. [8] The ever-increasing number of individuals who use the internet has resulted in the platform known as Instagram being seen as a highly significant one for the purposes of advertising, marketing, and social engagement. This is due to the ever-increasing number of individuals who have access to and make use of the internet. Even if it is utilised by millions of people, there is a problematic subset of those users who engage in the questionable practise of misusing the network by fabricating identities for themselves in order to hide their true intentions. Even if the Internet has shown to be helpful in recent years, online social networks are still prone to the hazards that are posed by spammers and other sorts of cybercriminals. Users resort to a wide variety of inappropriate strategies in order to promote bigger profile followers as a direct result of the fact that the popularity of users on social media is based on the number of followers they have. This is because the popularity of users is determined by the number of followers they have. For the many varied uses of social media, researchers have developed a broad array of tactics that are not only successful but also easily implementable. This article's goal is to offer a technique for the automatic detection of fake accounts for the purpose of discovering phoney Instagram profiles in order to safeguard the users of Instagram with respect to their social life. The approach will be shown in this post. This article also includes a technique for the automatic detection of bogus accounts and contains related information. Using supervised machine learning algorithms, it is possible to create solid predictions regarding whether or not Instagram accounts are false. This is a currently active research area. After being placed into the appropriate categories, phoney profile IDs are then added to a data dictionary in order to provide further assistance to the relevant authorities in taking the appropriate actions against bogus social media identities. This is done in order to be of more support to those in need.

L. P et al. [9] It is possible for the function that social media platforms play in today's society to have a significant impact on every facet of an individual's life. The vast majority of individuals spend the bulk of their time, 24 hours a day, seven days a week, communicating on various social media platforms. This includes both their waking hours and their sleeping hours. The number of people who have accounts on these social networking sites is quickly growing on a daily basis, and a substantial proportion of those persons are talking with other users regardless of the time or location of the other users. This interaction may take place anywhere in the world. This phenomenon is referred regarded as "globalisation," and its description uses that word. These social networking sites each come with their own individual set of perks and cons, and in addition, they pose possible dangers to the privacy of not just our information but also the information that our writers have shared with us. Author needs to organise these social networking accounts in a way that enables Author to differentiate between genuine accounts and fake ones before he can examine who is making threats on these networking sites. This is necessary so that Author can examine who is making threats on these networking sites. The author has, in the past, made use of a broad variety of classification algorithms in order to identify bogus accounts on a number of different social networking platforms. On the other hand, Author has to speed up the pace at which they can recognise fake accounts on these many sites.

F. C. Akyon et al. [10] One of the most significant issues that can arise in online social networks is the practise of creating fake interactions with other users in online social networks (OSNs), particularly those that are used to artificially enhance the level of popularity of an account. This is one of the most problematic things that can take place in online social networks. The identification of fraudulent interactions is very important due to the fact that these interactions may result in monetary losses for organisations, poor audience targeting in advertising, incorrect product prediction algorithms, and an unpleasant environment inside social networks. The purpose of this research is to discover false and automated profiles on Instagram, the existence of which may create the impression that there is a greater level of involvement than really exists. To the best of the Author's knowledge, there is no publicly accessible dataset that comprises false accounts or accounts that were produced by a computer programme. This is the case despite the fact that such datasets are widely available. In order to accomplish this goal, two datasets that were developed specifically for the purpose of identifying automated accounts and fake accounts respectively have been established. These datasets are designed to be used in combination with one another, as described in the previous sentence. In order to conduct an analysis of these reports, a number of different machine learning methods, including Naive Bayes, logistic regression, support vector machines, and neural networks, to name a few, have been used.

Ekosputra et al. [11] Because of the vast number of celebrities and fan pages that use the platform as the location for them to engage with one another, Instagram has acquired a tremendous degree of popularity as a consequence of this fact. Instagram is by far the most popular social media platform that is used for marketing a broad range of companies. This is due to the fact that it allows users to post a number of different sorts of content in a variety of different ways and that it is extremely easy to use. Even though it is the most visited website, Instagram may sometimes include photographs of "fake" people in its feeds. Regrettably, there are some individuals who utilise false accounts in order to participate in negative behaviours, such as mimicking artists or influencers, spreading rumours, and making nasty comments. These activities may be found on several social media platforms. Mimicking prominent personalities is one of the other types of behaviours. These acts are being carried out with the goal of spreading virally over the internet. As a direct result of this, the purpose of this study is to identify phoney Instagram users by conducting an analysis of the profiles of persons that fulfil the aforementioned criteria. Before a legitimate account can be accurately detected, several steps must first be completed. These processes begin with the pre-processing of data, continue with the selection of an appropriate classification model, and then go on to the execution of classification assessment. It is essential that each of these stages be completed effectively. When constructing a supervised machine learning model, some of the machine learning algorithms that are used include Bernoulli-naive bayes, random forest, support vector machine, and artificial neural network. Logical regression is another one of the machine learning methods that is utilised. Support vector machine and logistic regression are two other examples of machine learning techniques that are put to use (ANN).

Tiwari et al. [12] Recent developments in technology could be directly responsible for a meteoric rise in the number of people using social networking sites. This phenomenon has been seen recently. On Facebook, there are around 1.5 billion active users at any one time. More than ten million new likes and shares are generated every single day on Facebook and other social media platforms. The number of people using several extra social media sites like LinkedIn, Instagram, Pinterest, and Twitter, amongst others, is increasing at a rapid pace. The proliferation of social networking sites has resulted in the creation of a large number of fictitious user identities that are then put to use for deceptive or harmful purposes. These accounts are accessible on several websites, including Facebook, Twitter, and Instagram, among others. There are a variety of names for fake accounts, including Sybils and social Bots, to name just a few of the more common ones. There are a lot of these personalities out there, and a lot of them try to befriend people who aren't paying attention with the intention of eventually gaining access to their private information by hacking their accounts. There are a lot of these personas floating about. In practically every online social network, the most common kind of security risk is known as social engineering, which refers to the process of manipulating others (OSN).

Anklesaria et al. [13] Since the advent of the internet, there has been a significant rise in the amount of dishonest and unlawful behaviour that takes place online. This trend is expected to continue. This encompasses both criminal activity committed online as well as identity theft. The overall level of criminal behaviour on the internet has grown as a direct result of this. The prevalence of social media, which acts as one of the most common venues for fraudsters to target their victims, is only helping to contribute to the rise in the number of online scams that have been reported in recent times. This development came about as a direct result of the rise in the use of mobile devices to access the internet. [Example in point:] the alarming increase in the number of con artists operating online that has been documented in recent years. [Case in point:] Fraudsters have made social media platforms like Instagram, Facebook, and Whatsapp their principal operations centres, and these websites make it easy for scammers to engage in risky behaviour. Instagram and Facebook are both owned by Facebook. Con artists are the ones who are in charge of carrying out these ruses. They use fictitious internet personas, which they construct themselves, in order to conceal their real identities from prying eyes. Author need a piece of software that is able to differentiate between phoney accounts and real accounts in order to protect oneself from falling subject to scams of this sort and prevent themselves from being taken advantage of.

Gupta et al. [14] People have formed a substantial dependency on online social networks (OSNs), which has drawn the attention of cybercriminals who are interested in participating in a variety of illegal activities. Cybercriminals are interested in exploiting the dependence that people have built on OSNs. As a direct consequence of the fact that it is now practicable to obtain services that are predicated on the use of phoney accounts, a whole sector of the shadow economy has surfaced as a direct result of this development. As a direct result of this, the primary focus of Author's research and development efforts is now centred on the identification of phoney accounts on Facebook, a very well-liked (but challenging for data collecting) online social network. The following is a synopsis of the author's work's most significant contributions, in the order of their importance. In order to get the necessary information, the first thing that we have done is do research on both actual and fraudulent accounts that are currently active on Facebook. Collecting data from user accounts has become very difficult as a result of Facebook's stringent privacy settings and its constantly improving application programming interface (API), both of which introduce additional limitations with each new iteration. Nevertheless, this data can still be collected. This is due to the fact that each of these aspects make it more difficult with each subsequent edition. The development of a thorough set of 17 characteristics that play a vital role in differentiating fake users from genuine users on Facebook is the author's second major contribution. This contribution involves the utilisation of information from user feeds on Facebook to analyse user profile activity. The fact that the Author was able to put both of these areas of knowledge to work made it possible for both of these areas of contribution to be made. The author's last contribution consisted of using these variables to determine which machine learning-based classifiers did particularly well in the detection task. This was the author's final contribution. There were a total of twelve distinct classifiers that were used in this study. The discovery of certain kinds of behaviours (such as like, commenting, tagging, sharing, and so on) that contribute the most to the detection of fake users is the fourth contribution.

Van Der Walt et al. [15] A rising number of individuals are utilising social media platforms (SMPs) to participate in activities online while concealing their identities with the objective of engaging in behaviour that is either illegal or immoral. The purpose of these actions may be found in the following sentence: To the author's dismay, very little study has been conducted to determine how to identify fraudulent identities created by people, particularly in relation to SMPs. This is a particular area of interest for the author. On the other hand, there have been a number of instances in which fraudulent accounts that were established by bots or computers have been effectively uncovered by utilising machine learning models. In these instances, the bots or computers used to construct the accounts were successfully identified. These incidences have taken place a great deal of times in the past. These occurrences have been captured on record. When it came to bots, the success of these machine learning models was determined in large part by the use of a set of predetermined characteristics. The "friend-to-followers ratio" is a good illustration of this concept. These capabilities were developed by using features such as "friend-count" and "follower-count," both of which are conveniently located inside the user profiles on SMPs. Both of these counts are readily accessible to users. These figures represent the total number of friends and followers for a particular individual.

III. CONCLUSION

Facebook, Twitter, and Instagram are just some of the social networking websites that have received a considerable amount of prominence in recent years. Users are continually engaged in dialogue with their contemporaries through a range of media, including social networking websites. They communicate information that is private as well as information that is available to the public by utilising these social networks. Many individuals log on to social networking websites on a daily basis as a consequence of the pervasive attractiveness of these online communities. Since of this reputation, the websites encounter issues since it encourages the installation of fraudulent accounts. The operators of fraudulent accounts steal the personal information of other users and spread the stolen data throughout the different social media networks. This study discussed a few strategies and reviews of prior work in fake user profile identification and its relevance.

References

[1] G. Hajdu, Y. Minoso, R. Lopez, M. Acosta and A. Elleithy, "Use of Artificial Neural Networks to Identify Fake Profiles," 2019 IEEE Long Island Systems, Applications and Technology Conference (LISAT), 2019, pp. 1-4, doi: 10.1109/LISAT.2019.8817330.

[2] V. Kulkarni, D. Aashritha Reddy, P. Sreevani and R. N. Teja, "Fake profile identification using ANN," 4th Smart Cities Symposium (SCS 2021), 2021, pp. 375-380, doi: 10.1049/icp.2022.0372.

[3] S. Khaled, N. El-Tazi and H. M. O. Mokhtar, "Detecting Fake Accounts on Social Media," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 3672-3681, doi: 10.1109/BigData.2018.8621913.

[4] K. Patel, S. Agrahari and S. Srivastava, "Survey on Fake Profile Detection on Social Sites by Using Machine Learning Algorithm," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 1236-1240, doi: 10.1109/ICRITO48877.2020.9197935.

[5] N. Singh, T. Sharma, A. Thakral and T. Choudhury, "Detection of Fake Profile in Online Social Networks Using Machine Learning," 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), 2018, pp. 231-234, doi: 10.1109/ICACCE.2018.8441713.

[6] A. Bhattacharya, R. Bathla, A. Rana and G. Arora, "Application of Machine Learning Techniques in Detecting Fake Profiles on Social Media," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-8, doi: 10.1109/ICRITO51393.2021.9596373.

[7] İ. AYDIN, M. SEVİ and M. U. SALUR, "Detection of Fake Twitter Accounts with Machine Learning Algorithms," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1-4, doi: 10.1109/IDAP.2018.8620830.

[8] P. Harris, J. Gojal, R. Chitra and S. Anithra, "Fake Instagram Profile Identification and Classification using Machine Learning," 2021 2nd Global Conference for Advancement in Technology (GCAT), 2021, pp. 1-5, doi: 10.1109/GCAT52182.2021.9587858.

[9] L. P, S. V, V. Sasikala, J. Arunarasi, A. R. Rajini and N. Nithiya, "Fake Profile Identification in Social Network using Machine Learning and NLP," 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT), 2022, pp. 1-4, doi: 10.1109/IC3IOT53935.2022.9767958.

[10] F. C. Akyon and M. Esat Kalfaoglu, "Instagram Fake and Automated Account Detection," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), 2019, pp. 1-7, doi: 10.1109/ASYU48272.2019.8946437.

[11] M. J. Ekosputra, A. Susanto, F. Haryanto and D. Suhartono, "Supervised Machine Learning Algorithms to Detect Instagram Fake Accounts," 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2021, pp. 396-400, doi: 10.1109/ISRITI54043.2021.9702833.

[12] V. Tiwari, "Analysis and detection of fake profile over social network," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 175-179, doi: 10.1109/CCAA.2017.8229795.

[13] K. Anklesaria, Z. Desai, V. Kulkarni and H. Balasubramaniam, "A Survey on Machine Learning Algorithms for Detecting Fake Instagram Accounts," 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021, pp. 141-144, doi: 10.1109/ICAC3N53548.2021.9725724.

[14] A. Gupta and R. Kaushal, "Towards detecting fake user accounts in facebook," 2017 ISEA Asia Security and Privacy (ISEASP), 2017, pp. 1-6, doi: 10.1109/ISEASP.2017.7976996.

[15] E. Van Der Walt and J. Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," in IEEE Access, vol. 6, pp. 6540-6549, 2018, doi: 10.1109/ACCESS.2018.2796018.