International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

# DETECTION OF FAKE AND CLONE ACCOUNTS IN TWITTER USING CLASSIFICATION AND DISTANCE MEASURE ALGORITHMS

**[1]Mr. G Uday Kishore ,[2]V.Sanjana ,[3]B.Charitha, [4]Shaik Shoaib Ahmed**

[1]Assistant Professor, Department of Information Technology, CMR College of Engineering & Technology

[2, 3, 4] B-Tech, Department of Information Technology, CMR College of Engineering & Technology

**Abstract:**

Online Social Network (OSN) is a network hub where people with similar interests or real world relationships interact. As the popularity of OSN is increasing, the security and privacy issues related to it are also rising. Fake and Clone profiles are creating dangerous security problems to social network users. Cloning of user profiles is one serious threat, where already existing users details are stolen to create duplicate profiles and then it is misused for damagingthe identity of original profile owner. They can even launch threats like phishing, stalking, spamming etc. Fake profile is the creation of profile in the name of a person or a company which does not really exist in social media, to carry out malicious activities. In this paper, a detection method has been proposed which can detect Fake and Clone profiles in Twitter. Fake profiles are detected based on set of rules that can effectively classify fake and genuine profiles. For Profile Cloning detection two methods are used. One using Similarity Measures and the other using C4.5 decision tree algorithm. In Similarity Measures, two types of similarities are considered – Similarity of Attributes and Similarity of Network relationships. C4.5 detects clones by building decision tree by taking information gain into consideration. A comparison is made to check how well these two methods help in detecting clone profiles.

**INTRODUCTION :**

Fake and Clone profiles are creating dangerous security problems to social network users. Cloning of user profiles is one serious threat, where already existing user's details are stolen to create duplicate profiles and then it is misused for damaging the identity of original profile owner. They can even launch threats like phishing, stalking, spamming etc. Fake profile is the creation of profile in the name of a person or a company which does not really exist in social media, to carry out malicious activities. In this paper, a detection method has been proposed which can detect Fake and Clone profiles in Twitter. Fake profiles are detected based on set of rules that can effectively classify

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

fake and genuine profiles. For Profile Cloning detection two methods are used. 1.1 Problem Statement Social networks are an essential part of our life. Social networks are used for social, business purposes through Facebook, Twitter, or Instagram. Twitter is considered one of the largest social networks with a vast user base. A vast user base comes with certain drawbacks, Fake accounts. It becomes challenging for twitter to manage the users and detect fake profiles, andit still manages to do so. However, fake profiles are creating dangerous security problems for social network users. Fake profiles are created with credentials that do not exist in the world. making them a hoax. These fake accounts can also launch attacks on real users. Sharing personal data without reading privacy policies makes cyber-attacks easy. Cloning user profilesis another severe threat, where already existing users' detalls are stolen to create duplicate profiles. It is then misused for damaging the identity of the original profile owner. This paper proposes specific methods to detect fake accounts on a social media network like Twitter. Fake profiles are detected on a set of rules which classify Real and Fake accounts. This paper will not limit ourselves to specific algorithms, but we will use a couple of classification algorithms like Using distance measure and classification algorithm . Finally, a comparison is made to check how these algorithms detect fake accounts

**OBJECTIVE :**

Today, Fake and Clone profiles have become a very serious threat in social networks. So, a detection method is very much necessary to find these frauds who use people's faith to gather private information and create duplicate profiles. Many authors have worked in this area and have proposed methods to identify these types of profiles in social networks.

**IMPLEMENTATION OF FAKE PROFILE DETECTION IN SOCIAL MEDIA:**

Information is extracted from the user profile and a search is made in an online social network to find profiles which match to that of the user profile and a similarity score is calculated through the trained classifier. If the similarity score is greater then the profile is termed as fake. The feedback is used to train the classifier.

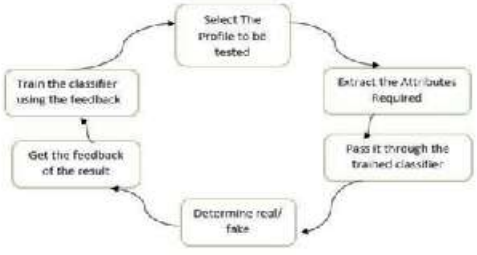


Fig - 1: Fake Profile detection in Social Media

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

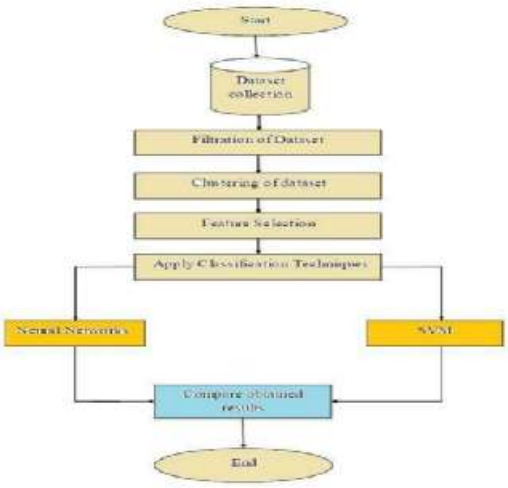


Fig - 2: Detection of Fake Accounts on Social Media using Neural Network

**PROPOSED SYSTEM:**

Objective of proposed model Fake and clone profiles have become a major social issue. Because information such as phone numbers, email addresses, school or college names, corporate names, and locations are publicly available on social media sites, hackers can easily use this data to construct bogus or clone identities. They then attempt to perpetrate various attacks such as phishing, spamming, cyber bullying, and so on. They even go so far as to try to discredit the legitimate owner or organization. So, in order to make users social lives more secure, a detection approach has been presented that can detect both fake and clone profiles. 3.2 Algorithms Used for Proposed Model 3.2.1 Support Vector Machine Algorithm SVM or Support Vector Machine is a linear model for classification and regression problems. It can solve linear and non-linear problems and work well for many practical problems. The idea of SVM is simple: The algorithm creates a line or a hyper plane which separates the data into classes

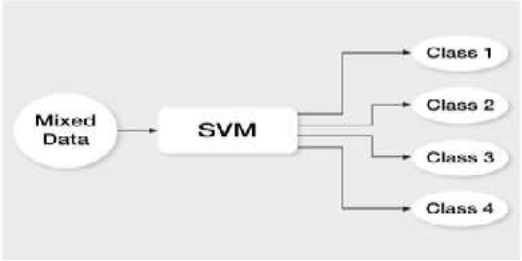


Fig - 3: Support Vector Machine Algorithm

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyper plane.SVM chooses the extreme points/vectors that help in creating the hyper plane.

Designing System Architectural Design:

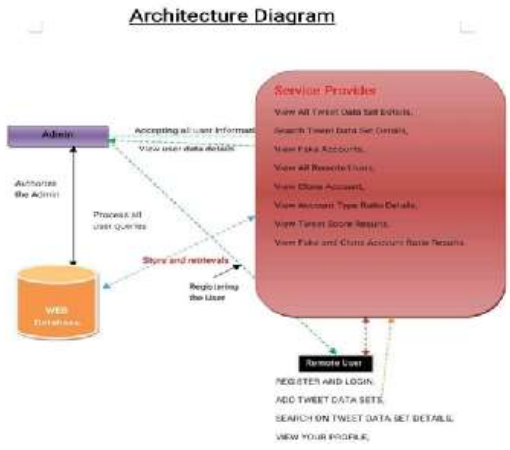


Fig:-4

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

**RESULTS:**

Performance Metrics Precision, is a measure of fraction of positive predictions, that are actually positive. Higher precision implies that this model is better to identify the needy people for help. The evaluation metrics considered are

1. Accuracy which gives the ratio of number of correct results to the total number of inputs

2. Precision which gives the proportion of positive detection that was actually correct

3. Recall which gives the proportion of actual positives that was detected correctly

4. F1 Score which takes into account both precision and recall to compute the score. F1- score is given by harmonic mean of precision and recall. If F1-score is 1, then it is best value and worst is 0.

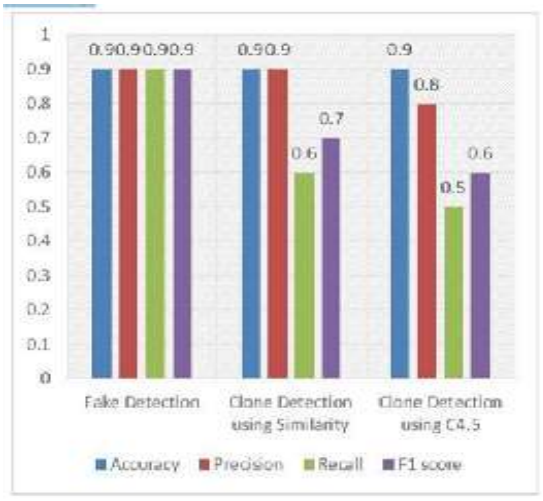


Fig:-5

**CONCLUSION :**

Fake and clone profiles have become a very serious problem in online social networks. We hear some or the other threats caused by these profiles in everyday life. So a detection method has been proposed which can find both fake and clone Twitter profiles. For fake detection, a set of rules were used which when applied can classify fake and genuine profiles. In future, weare intrigued to expand the work with some profound learning models.

**FUTURE ENHANCEMENT:**

Clone detection was carried out using Similarity Measures and C4.5 algorithm and a comparison was made to check the performance. Clone detection using Similarity Measures worked better than C4.5 and was able to detect most of the clones which were fed into the system. In this work we have considered only the profile attributes for fake and clone detection. In future this work can be extended by taking tweets also into consideration by applying some NLP techniques

**REFERENCES :**

[1] Sowmya P and Madhumita Chatterjee ," Detection of Fake and Cloned Profiles in Online Social Networks", Proceedings 2019: Conference on Technologies for Future Cities (CTFC) [2] Georgios Kontaxis, Iasonas Polakis, Sotiris

Ioannidis and Evangelos P. Markatos, "Detecting Social Network Profile Cloning", 2013

[3] Piotr Bródka, Mateusz Sobas and Henric Johnson, "Profile Cloning Detection in Social Networks", 2014 European Network Intelligence Conference

[4] Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angello Spognardi, Maurizio Tesconi, "Fame for sale: Efficient detection of fake Twitter followers", 2015 Elsevier's journal Decision Support Systems, Volume 80

[5] Ahmed El Azab, Amira M Idrees, Mahmoud A Mahmoud, Hesham Hefny, "Fake Account Detection in Twitter Based on Minimum Weighted Feature set", World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering Vol:10, 2016

[6] M.A.Devmane and N.K.Rana, "Detection and Prevention of Profile Cloning in Online Social Networks", 2014 IEEE International Conference on Recent Advances and Innovations in Engineering

[7] Kiruthiga. S, Kola Sujatha. P and Kannan. A, "Detecting Cloning Attack in Social Networks Using Classification and Clustering Techniques" 2014 International Conference on Recent Trends in Information Technology

[8] Reddy, b. Venkata ramana, nageshbabu dasari, and k. Venkateswararao. "A steganography system with gausian markov random fields and error detection codes." (2021).

[9] Dr. S.Balamurugan, & Aurchana, Aurchana & Gurumoorthi Elangovan, Dr & Govindharaj, I. (2022). Augmentation of Decision Tree Characteristics for Agri-Food Supply Chain using Internet of Things.

[10] Vinay, R., Soujanya, K. L. S., & Singh, P. (2019). Disease prediction by using deep learning based on patient treatment history. Int. J. Recent Technol. Eng, 7(6), 745-754.

[11] Latha, C. M., & Soujanya, K. L. S. (2018). Enhancing end-to-end device security of internet of things using dynamic cryptographic algorithm. Int. J. Civil Eng. Technol, 9(9), 408-415.

[12] Debnath, S., Islam, M., 2022, Disinfection chain: A novel method for cheap reusable and chemical free disinfection of public places from SARS-CoV-2, ISA Transactions, 10.1016/j.isatra.2021.03.040

[13] Kavati, V., Kavati, V., Varma, K.P.V.K., 2022, The Effect of Pine Oil Emulsifier on Gasoline-Alcohol Blends in a Spark Ignition Engine with Multi-Point Fuel Injection, Trends in Sciences, 10.48048/tis.2022.3427

[14] Narasimha, V., Dhanalakshmi, M., 2022, Detection and Severity Identification of Covid-19 in Chest X-ray Images Using Deep Learning, International Journal of Electrical and Electronics Research, 10.37391/IJEER.100250

[15] Madhavi, K.R., Kora, P., Reddy, L.V., Avanija, J., Soujanya, K.L.S., Telagarapu, P., 2022, Cardiac arrhythmia detection using dual-tree wavelet transform and convolutional neural network, Soft Computing, 10.1007/s00500-021-06653-w

.