Fake Social Media Profile Detection Using Machine Learning

Dr. R. S. Khule[1], Pooja Gavande[2], Harshada Sonawane[3], Anushka Niphade[4], Pooja Phad[5]

[1] *Professor, Information Technology, Matoshri College of Engineering and Research Centre, Nashik*
[2] *Student, Information Technology, Matoshri College of Engineering and Research Centre, Nashik*
[3] *Student, Information Technology, Matoshri College of Engineering and Research Centre, Nashik*
[4] *Student, Information Technology, Matoshri College of Engineering and Research Centre, Nashik*
[5] *Student, Information Technology, Matoshri College of Engineering and Research Centre, Nashik*

ABSTRACT

Abstract— Online spoofing and fraudulent accounts are common on the social network, which is so crucial to our daily lives. Fake accounts that appear to have been created on behalf of businesses or individuals are the ones most likely to use issues related to social media, such as confidentiality, online abuse, misuse, bullying, etc., which can harm a person's reputation and decrease their number of likes and followers. On the other hand, the creation of false accounts is anticipated to harm more people than any other type of cybercrime. This problem inspires us to create a machine learning-based system for identifying fraudulent social media accounts. False profiles are frequently used by intruders in online social networks to engage in harmful activities like harassing individuals, identity theft, and privacy violations. As a result, one of the most challenging tasks on the online social network site is figuring out whether an account is legitimate or fraudulent. In this study, we introduced the Support Vector Machine algorithm and a deep neural network, among other classification methods. Additionally, it contrasts classification strategies using the Spam User dataset.

Keywords— social media, Fake accounts, Machine learning algorithms, Comprehensive Review, Support vector machine.

1. INTRODUCTION

Fake accounts that appear to have been created on behalf of businesses or individuals are the ones most likely to use issues related to social media, such as confidentiality, online abuse, misuse, bullying, etc., which can harm a person's reputation and decrease their number of likes and followers. On the other hand, the creation of false accounts is anticipated to harm more people than any other type of cybercrime. This problem inspires us to create a machine learning-based system for identifying fraudulent social media accounts.

Fake accounts that appear to have been created on behalf of businesses or individuals are the ones most likely to use issues related to social media, such as confidentiality, online abuse, misuse, bullying, etc., which can harm a person's reputation and decrease their number of likes and followers. On the other hand, the creation of false accounts is anticipated to harm more people than any other type of cybercrime. This problem inspires us to create a machine learning-based system for identifying fraudulent social media accounts.

2. LITERATURE SURVEY

The author[1] of this study put forth a fake profile detection model that uses sentiment-based attributes to distinguish between authentic and fraudulent OSN profiles. The study is based on the observation that real users' posts display a range of emotions based on their personal experiences, including joy, sadness, anger, fear, etc. Contrarily, fake users share posts to achieve a particular goal, so it is very likely that the content of their posts will include the same kinds

of emotions. To eliminate the outliers from the dataset, a noise removal technique is also presented. Finally, the detection model has been trained using a variety of machine learning techniques, including Support Vector Machine (SVM), Naive Bayes, and Random Forest.

To give an idea of profile cloning recognition in Online Social Networks (OSN) using Network Theory, this paper[2] makes an effort. Based on malicious users' most recent activities in the social network, this study investigates the Node Similarity Communication Matching algorithm using profile cloning recognition in online social networks. The various activities to be studied in this proposed method include things like Updates, Wall posts and comments, By recent activities, etc. Based on the comparison of threshold values of user-specific profile attribute values and network similarity analysis, malicious that steals users' identities are identified. Creating an account, user operation, monitoring, searching recent activity, detecting cloned profiles, choosing a profile to be examined, and determining whether a profile is real, or fake are some of the processes used in the research.

3. SYSTEM ARCHITECTURE AND METHODOLOGY

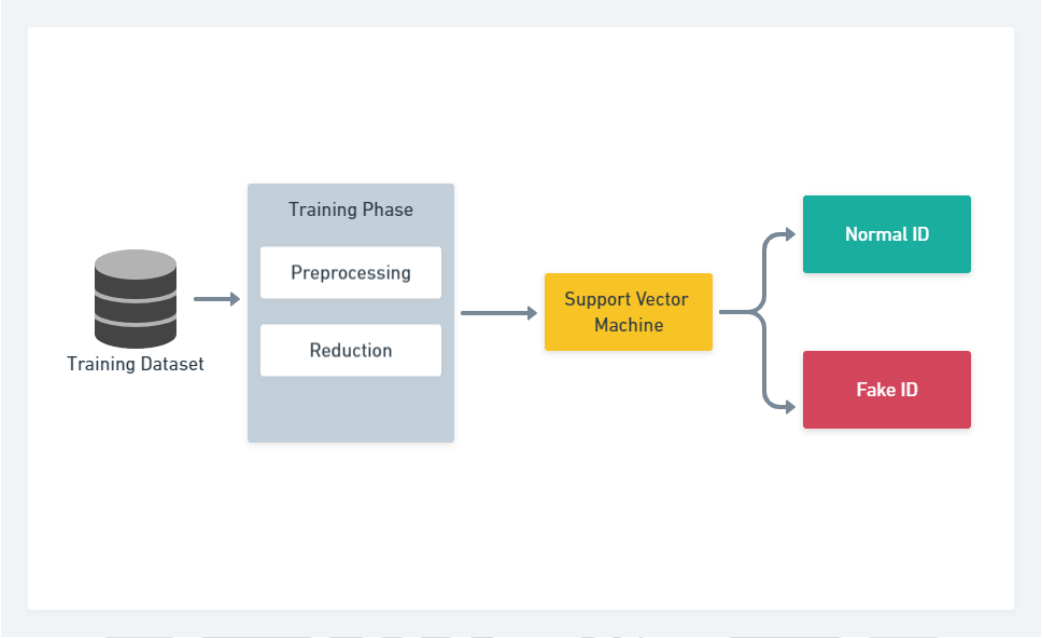

Fig -1: System Architecture

In our research work, a novel approach has been presented for the identification of fake profiles on social media using supervised machine learning algorithms. The proposed model has applied data pre-processing techniques to datasets before Analyzing them. A technique has been applied to identify the non-significant attributes in datasets and to do attribute reduction. For the dataset that includes both fake and real users, the proposed model is trained using supervised machine learning techniques on an individual basis. The prediction has been improved using an SVM classifier.

The model proposed in this proposed work demonstrates that Support Vector Machine (SVM) is an elegant and reliable method for binary classification in a large dataset. Despite the non-linearity of the decision boundary, SVM can accurately (>90%) distinguish between fake and real profiles. Any platform that needs binary classification on public profiles for various reasons can use this method. This project only uses publicly available information, which is convenient for organizations that want to avoid any breach of privacy, but organizations can also use private data to further extend the capabilities of the proposed model.

- In our research work, a novel approach has been presented for the identification of fake profiles on social media using supervised machine learning algorithms.

- The proposed model has applied data preprocessing techniques to datasets before analyzing them. A technique has been applied to identify the non-significant attributes in datasets and to do attribute reduction.

- The proposed model was trained using supervised machine algorithms individually for the dataset including fake and genuine users. An ensemble classifier has been used to make the prediction more accurate.

4. MATHEMATICAL MODEL

The problem of detecting fake social media accounts can be formulated as a binary classification task, where the goal is to classify each account as either fake or genuine. One approach to solving this problem is to use Support Vector Machines (SVMs), which are popular machine-learning algorithms for binary classification tasks.

Let's denote the set of training examples by X, where each example x_i is a feature vector that represents the characteristics of a social media account. The corresponding labels for the examples are denoted by y_i, where y_i = 1 indicates a genuine account and y_i = -1 indicates a fake account.

The SVM algorithm aims to find a hyperplane that separates the two classes in the feature space. The hyperplane is defined by the equation:

$$w^T x + b = 0$$

where w is the weight vector and b is the bias term. The decision boundary is given by the sign of the expression w^T x + b, which is positive for genuine accounts and negative for fake accounts.

The SVM algorithm seeks to find the weight vector w and the bias term b that maximize the margin between the decision boundary and the closest examples from each class. The margin is defined as the distance between the decision boundary and the closest examples.

The SVM problem can be formulated as an optimization problem:

$$\text{minimize } (1/2) \|w\|^2 + C \sum_{i=1}^{n} \max(0, 1 - y_i(w^T x_i + b))$$

$$\text{subject to } y_i(w^T x_i + b) \geq 1 \text{ for all } i = 1, ..., n$$

where C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error. The term $C \sum_{i=1}^{n} \max(0, 1 - y_i(w^T x_i + b))$ is the hinge loss function, which penalizes the classifier for misclassifying examples.

The SVM algorithm can be solved using various optimization techniques, such as gradient descent, quadratic programming, or interior point methods. Once the weight vector w and the bias term b are learned, the SVM classifier can be used to predict the label of new social media accounts by evaluating the sign of the expression w^T x + b.

5. CONCLUSIONS

In this proposed work, we propose a machine-learning pipeline for detecting fraudulent accounts in online social networks. The use of datasets containing fake profiles effectively eliminates the difficulty of detecting fake profiles. This work presented a technique for detecting fake accounts created with NLP and machine learning. A classification algorithm for detecting fake profiles on social networks is presented. For classifying the fake and genuine profiles, we use SVM / Naive Bays.

6. REFERENCES

[1] S. Revathi and D. M. Suriakala, "Profile Similarity Communication Matching Approaches for Detection of Duplicate Profiles in Online Social Network," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2018, pp. 174-182

[2] Romanov, A., Semenov, A., Veijalainen, J.: Revealing fake profiles in social networks by longitudinal data analysis. In: 13th International Conference on Web Information Systems and Technologies, January 2017

[3] Song, J., Lee, S., Kim, J.: CrowdTarget: target-based detection of crowdturfing in online social networks. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, CCS 2015, pp. 793–804. ACM, New York (2015)

[4] Nazir, A., Raza, S., Chuah, C.-N., Schipper, B.: Ghostbusting Facebook: detecting and characterizing phantom profiles in online social gaming applications. In: Proceedings of the 3rd Conference on Online Social Networks, WOSN 2010. USENIX Association, Berkeley, CA, USA, p. 1 (2010)

[5] Adikari, S., Dutta, K.: Identifying fake profiles in Linkedin. Presented at the Pacific Asia Conference on Information Systems PACIS 2014 Proceedings (2014)

[6] Stringhini, G., Kruegel, C., Vigna, G.: Detecting spammers on social networks. In: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC 2010, pp. 1–9 (2010)

[7] Yang, C., Harkreader, R.C., Gu, G.: Die free or live hard? Empirical evaluation and new design for fighting evolving Twitter spammers. In: Proceedings of the 14th International Conference on Recent Advances in Intrusion Detection, RAID 2011, pp. 318–337. Springer, Heidelberg (2011)

[8] Elyusufi, Y., Seghiouer, H., Alimam, M.A.: Building profiles based on ontology for recommendation custom interfaces. In: International Conference on Multimedia Computing and Systems (ICMCS) Anonymous IEEE, pp. 558–562 (2014)

[9] Elyusufi, Y., Alimam, M.A, Seghiouer, H.: Recommendation of personalized RSS feeds based on ontology approach and multi-agent system in web 2.0. J. Theor. Appl. Inf. Technol. 70(2), 324–332 (2014) Social Networks Fake Profiles Detection 39

[10] Elyusufi, Z., Elyusufi, Y., Ait Kbir, M.: Customer profiling using CEP architecture in a Big Data context. In: SCA 2018 Proceedings of the 3rd International Conference on Smart City Applications Article No. 64, Tetouan, Morocco, 10–11 October 2018. ISBN: 978-1-4503- 6562-8

[11] Granik, M., Mesyura, V.: Fake news detection using naive Bayes classifier. In: Conference: IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), May 2017

[12] Ameena, A., Reeba, R.: Survey on different classification techniques for detection of fake profiles in social networks. Int. J. Sci. Technol. Manage. 04(01), (2015)

[13] Beatriche, G.: Detection of fake profiles in Online Social Networks (OSNs), Master's degree in Applied Telecommunications and Engineering Management (MASTEAM), (2018)