IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, August 2023

Evaluating and Detecting Fake Users in Social Media by Random Forest

Mahi Maanas Reddy[1], Shruti Sridhar[2], V. Maria Anu[3], Dr Punitha K[4]
UG Students, School of Computer Science and Engineering[1,2]
Associate Professor, School of Computer Science and Engineering[3,4]
Vellore Institute of Technology, Chennai, India

Abstract: *Currently, users have been engaging in conversations, sharing information and producing web content via social media platforms. But in recent times, many users have been using these platforms to conduct identity faults, payment frauds, and many more without the knowledge of the actual user. For example: - On Instagram, according to the latest analysis, there are around 95 million fake accounts compared to the total number of users, which amount to 1 billion. Therefore, there are nearly 10% of fake accounts active at present. The obtained dataset lies approximately in thousands. Hence, we used GANs and deep learning to broaden the data to around 1 lakh. The conventional methods used for distinguishing between real and fake accounts were ineffective. Adopting machine learning-based approaches allowed us to identify fake accounts that can mislead users. The dataset is pre-processed using several Python tools, and a comparison model is created to identify a practical solution appropriate for the dataset that has been provided.*

Keywords: Logistic Regression, Random Forest, XGBoost, Support Vector Machine, Generative adversarial networks.

REFERENCES

[1]. CagatayAkyon, Fatih, and EsatKalfaoglu. "Instagram fake and automated account detection." arXiv e-prints (2019): arXiv-1910.

[2]. Fong, Simon, Yan Zhuang, and Jiaying He. "Not every friend on a social network can be trusted: Classifying imposters using decision trees." In The First International Conference on Future Generation Communication Technologies, pp. 58-63. IEEE, 2012.

[3]. Conti, Mauro, Radha Poovendran, and Marco Secchiero. "Fakebook: Detecting fake profiles in on-line social networks." In 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1071-1078. IEEE, 2012.

[4]. Maniraj, S. P., G. Harie Krishnan, T. Surya, and R. Pranav. "Fake account detection using machine learning and data science." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 9, no. 1 (2019).

[5]. -Cao, Qiang, Michael Sirivianos, Xiaowei Yang, and Tiago Pregueiro. "Aiding the detection of fake accounts in large scale social online services." In Presented as part of the 9th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 12), pp. 197-210. 2012.

[6]. Yang, Chao, Robert Harkreader, and Guofei Gu. "Empirical evaluation and new design for fighting evolving twitter spammers." IEEE Transactions on Information Forensics and Security 8, no. 8 (2013): 1280-1293.

[7]. Shama, Sk, K. Siva Nandini, P. Bhavya Anjali, and K. Devi Manaswi. "Fake profile identification in online social networks." Int. J. Recent Technol. Eng 8, no. 4 (2019): 1119011194.

[8]. Fire, Michael, Dima Kagan, AviadElyashar, and Yuval Elovici. "Friend or foe? Fake profile identification in online social networks." Social Network Analysis and Mining 4 (2014): 1-23.

[9]. Jia, Jinyuan, Binghui Wang, and Neil Zhenqiang Gong. "Random walk based fake account detection in online social networks." In 2017 47th annual IEEE/IFIP international conference on dependable systems and networks (DSN), pp. 273-284. IEEE, 2017.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, August 2023

[10]. Tsikerdekis, Michail, and SheraliZeadally. "Multiple account identity deception detection in social media using nonverbal behavior." IEEE Transactions on Information Forensics and Security 9, no. 8 (2014): 1311-1321.

[11]. Lee, Kyumin, James Caverlee, and Steve Webb. "Uncovering social spammers: social honeypots+ machine learning." In Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval, pp. 435-442. 2010.

[12]. Wang, G. Alan, Hsinchun Chen, Jennifer J. Xu, and HomaAtabakhsh. "Automatically detecting criminal identity deception: an adaptive detection algorithm." IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 36, no. 5 (2006): 988-999.

[13]. Adikari, Shalinda, and Kaushik Dutta. "Identifying fake profiles in linkedin." arXiv preprint arXiv:2006.01381 (2020).

[14]. Wang, Binghui, Neil Zhenqiang Gong, and Hao Fu. "GANG: Detecting fraudulent users in online social networks via guilt-byassociation on directed graphs." In 2017 IEEE International Conference on Data Mining (ICDM), pp. 465-474. IEEE, 2017.

[15]. Tang, Rui, Luke Lu, Yan Zhuang, and Simon Fong. "Not every friend on a social network can be trusted: an online trust indexing algorithm." In 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, vol. 3, pp. 280-285. IEEE, 2012.

[16]. Aziira, A. H., N. A. Setiawan, and I. Soesanti. "Generation of synthetic continuous numerical data using generative adversarial networks." In Journal of Physics: Conference Series, vol. 1577, no. 1, p. 012027. IOP Publishing, 2020.

[17]. Zhao, Miaoyun, Yulai Cong, and Lawrence Carin. "On leveraging pretrained GANs for generation with limited data." In International Conference on Machine Learning, pp. 11340-11351. PMLR, 2020.

[18]. Bourou, Stavroula, Andreas El Saer, Terpsichori-Helen Velivassaki, Artemis Voulkidis, and Theodore Zahariadis. "A review of tabular data synthesis using GANs on an IDS dataset." Information 12, no. 09 (2021): 375.