Evaluating and Detecting Fake Users in Social Media by Random Forest

Mahi Maanas Reddy1, Shruti Sridhar2, V. Maria Anu3, Dr Punitha K4

UG Students, School of Computer Science and Engineering1,2
Associate Professor, School of Computer Science and Engineering3,4
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


