International Journal of Research in Engineering, Science and Management
Volume 7, Issue 5, May 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Fake Account Detection on Instagram using Machine Learning

Tumuluru Shanmukha Nivas[1*], Pothireddy Sriramkrishna[2], Shavva Satya Keerthi Reddy[3], Pantham Veera Ram Gopal Rao[4], Rama Siva Prasad Komali[5]

[1,2,3,4,5] *B.Tech. Student, Department of Information Technology, Sagi Ramakrishnam Raju Engineering College, Bhimavaram, India*

***Abstract*: In the present generation, online social networks (OSNs) have become increasingly popular, people's social lives have become more associated with these sites. They use OSNs to keep in touch with each other's, share news, organize events, and even run their own e-business. The rapid growth of OSNs and the massive amount of personal data of its subscribers have attracted attackers, and imposters to steal personal data, share false news, and spread malicious activities. Recognizing the urgency of addressing these threats, researchers have embarked on a quest to develop effective strategies for detecting and thwarting abnormal activities and fake accounts within OSNs. This paper proposes a novel approach, ADB-CB, aimed at bolstering the detection of fake Instagram accounts. By integrating advanced feature selection and dimension reduction techniques, coupled with the utilization of three distinct machine learning algorithms - Adaboost, Catboost and Extra tree classifier. We aim to enhance the accuracy and reliability of identifying fraudulent accounts.**

***Keywords*: Adaboost, Catboost, Extra tree classifier, Fake account detection, Machine Learning.**

1. Introduction

In today's digital era, online social networks (OSNs) have emerged as integral components of modern social interaction, profoundly influencing individuals' daily lives. Platforms such as Instagram have become a hub for communication, information dissemination, and even commercial endeavors, facilitating connections, sharing news, organizing events, and fostering e-commerce ventures. However, the exponential expansion of OSNs has not only enhanced their societal impact, but has also attracted malicious individuals seeking to exploit these networks for nefarious purposes. The pervasive presence of personal data within OSNs renders them lucrative targets for attackers and imposters aiming to compromise user privacy, disseminate misinformation, and perpetrate various forms of cybercrime. Among these threats, the proliferation of fake profiles represents a significant challenge, as these accounts can be used to perpetuate scams, spread false information, manipulate public opinion, and engage in illicit activities. Recognizing the critical need to safeguard the integrity of online social networks, researchers have embarked on a quest to develop robust methods for detecting and mitigating the proliferation of inauthentic profiles. Traditional approaches to fake account detection have predominantly relied on the analysis of account features and the application of classification algorithms. However, the efficacy of these methods has been hindered by several challenges.

Firstly, certain account features may inadvertently contribute to misclassification or fail to adequately capture the nuanced characteristics of fake profiles. Additionally, employing standalone classification algorithms may not consistently yield satisfactory results, underscoring the necessity for innovative approaches to enhance detection accuracy. In response to these challenges, this paper proposes a novel algorithm, ADB-CB, tailored specifically for the detection of fake Instagram accounts. By integrating advanced feature selection and dimension reduction techniques, alongside the utilization of three distinct machine learning classifiers—AdaBoost, Cat Boost, and Extra Tree Classifier—we aim to bolster the accuracy and efficiency of fake account detection on Instagram. Through the meticulous examination and experimentation with various algorithms and feature engineering methodologies, this study seeks to provide insights into effective strategies for combating the proliferation of inauthentic profiles, thereby fostering a safer and more trustworthy online social environment.

In the contemporary digital era, online social networks (OSNs) have become ubiquitous, profoundly shaping the way individuals communicate, interact, and share information. These platforms, such as Instagram, Facebook, and Twitter, have transcended their initial purpose of facilitating social connections to become integral components of modern society, intertwining people's social lives with virtual communities.

However, the meteoric rise of OSNs has also ushered in a plethora of cybersecurity challenges, with attackers and imposters exploiting these platforms to perpetrate various malicious activities. From identity theft and data breaches to the proliferation of fake accounts and the dissemination of false information, the vulnerabilities inherent in OSNs have made them prime targets for nefarious actors seeking to exploit unsuspecting users.

In response to these threats, researchers have increasingly focused on developing efficient techniques to detect and combat abnormal activities and fake accounts within OSNs. Leveraging the rich trove of data available on these platforms,

*Corresponding author: tshanmukhanivas@gmail.com researchers explore innovative approaches that integrate account features and machine learning classification algorithms to distinguish between genuine and fraudulent accounts.

Nevertheless, the effectiveness of existing methodologies is hindered by several challenges. Certain exploited features may inadvertently contribute to misclassification, while standalone classification algorithms may not consistently yield satisfactory results. In light of these limitations, this paper proposes a novel algorithm, ADB-CB, designed to enhance the detection of fake Instagram accounts. By incorporating advanced feature selection and dimension reduction techniques, coupled with the utilization of three distinct machine learning classification algorithms – AdaBoost, Cat Boost, and Extra Tree Classifier – our approach aims to improve accuracy and reliability in identifying fraudulent accounts. Through empirical evaluation, we aim to contribute to the ongoing efforts to fortify the security and integrity of online social networks.

A. Objective of the Study

The primary objective of this study is to enhance the detection accuracy of fake Instagram accounts by leveraging advanced machine learning techniques, including AdaBoost, Cat Boost, and Extra Tree Classifier algorithms. Through the integration of feature selection and dimension reduction methods, our aim is to develop a robust algorithm, ADB-CB, capable of efficiently distinguishing between genuine and inauthentic profiles, thereby improving the overall integrity of the platform.

B. Scope of the Study

The scope of this study encompasses the development and evaluation of the ADB-CB algorithm for detecting fake Instagram accounts. It includes the exploration of various feature selection and dimension reduction techniques to optimize classification accuracy. The study focuses specifically on the application of machine learning algorithms, namely AdaBoost, Cat Boost, and Extra Tree Classifier, within the context of Instagram account authentication, aiming to provide insights into effective detection strategies.

C. Adaboost

AdaBoost, short for Adaptive Boosting, is a popular ensemble learning algorithm used for classification tasks. It works by combining multiple weak learners (classifiers that perform slightly better than random guessing) to create a strong learner with improved predictive accuracy. AdaBoost operates iteratively, adjusting the weights of misclassified data points at each iteration to focus on those instances that are harder to classify correctly. During training, AdaBoost assigns equal weights to all data points initially.

It then trains a base classifier on the data and evaluates its performance. Data points that are misclassified by the base classifier are given higher weights, while correctly classified points receive lower weights. In subsequent iterations, the algorithm places more emphasis on the misclassified points, effectively forcing the classifier to focus on the most challenging instances. Each new base classifier is trained on a modified dataset where the weights of the data points are adjusted accordingly.

The final classifier is a weighted combination of all the base classifiers, with each classifier's weight determined by its performance in classifying the data. AdaBoost gives more weight to classifiers with higher accuracy, effectively allowing them to have a greater influence on the final decision. One of the key advantages of AdaBoost is its ability to handle complex datasets and improve classification accuracy without overfitting. By focusing on difficult instances and continuously refining the model, AdaBoost achieves high generalization performance.

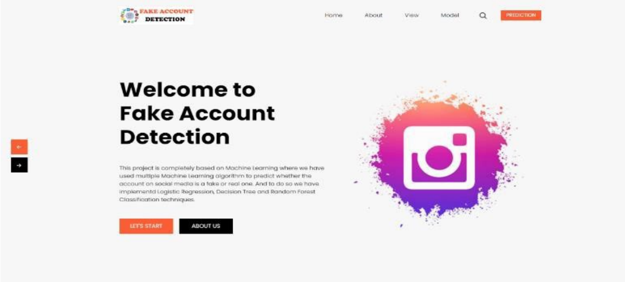
D. Catboost

The Cat Boost is a state-of-the-art gradient boosting algorithm specifically designed for handling categorical features in machine learning tasks. Developed by Yandex researchers, Cat Boost stands out for its ability to automatically handle categorical variables without the need for extensive pre-processing or one-hot encoding, making it particularly useful for real-world datasets where categorical features are prevalent.

Cat Boost employs a variant of gradient boosting called gradient boosting with categorical features support. During training, Cat Boost builds a series of decision trees sequentially, with each tree aiming to correct the errors made by the previous ones. Unlike traditional gradient boosting algorithms, Cat Boost uses a novel technique called ordered boosting, which reduces the overfitting risk by choosing the best split points for categorical features.

One of the key features of Cat Boost is its efficient handling of categorical variables. It employs an innovative algorithm for encoding categorical features, known as the target encoding or probability encoding, which takes into account the target variable to encode categorical values. This approach helps prevent overfitting and improves the model's predictive accuracy.

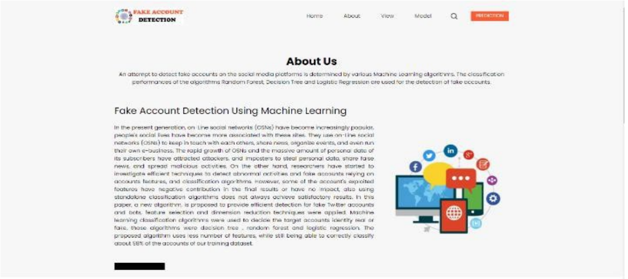
2. Results



Welcome to Fake Account Detection

This project is completely based on Machine Learning where we have used multiple Machine Learning algorithm to predict whether the account on social media is a fake or real one. And to do so we have implementd Logistic Regression, Decision Tree and Random Forest Classification techniques.

Fig. 1. Home Page



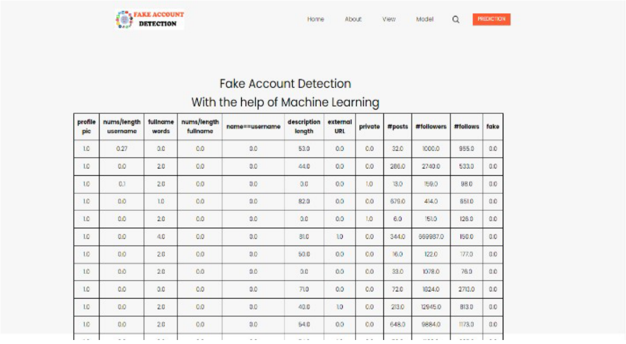
About Us

An attempt to detect fake accounts on the social media platforms is determined by various Machine Learning algorithms. The classification performances of the algorithms Random Forest, Decision Tree and Logistic Regression are used for the detection of fake accounts.

Fake Account Detection Using Machine Learning

In the present generation, on- Line social networks (OSNs) have become increasingly popular, people's social lives have become more associated with these sites. They use on-Line social networks (OSNs) to keep in touch with each others, share news, organize events, and even run their own e-business. The rapid growth of OSNs and the massive amount of personal data of its subscribers have attracted attackers, and imposters to steal personal data, share false news, and spread malicious activities. On the other hand, researchers have started to investigate efficient techniques to detect abnormal activities and fake accounts relying on accounts features, and classification algorithms. However, some of the account's exploited features have negative contribution in the final results or have no impact, also using standalone classification algorithms does not always achieve satisfactory results. In this paper, a new algorithm, is proposed to provide efficient detection for fake Twitter accounts and bots, feature selection and dimension reduction techniques were applied. Machine learning classification algorithms were used to decide the target accounts identity real or fake, those algorithms were decision tree , random forest and logistic regression. The proposed algorithm uses less number of features, while still being able to correctly classify about 98% of the accounts of our training dataset.

Fig. 2. Data



Fake Account Detection
With the help of Machine Learning

profile pic	nums/length username	fullname words	nums/length fullname	name==username	description length	external URL	private	#posts	#followers	#follows	fake
1.0	0.27	0.0	0.0	0.0	53.0	0.0	0.0	32.0	1000.0	955.0	0.0
1.0	0.0	2.0	0.0	0.0	44.0	0.0	0.0	286.0	2740.0	533.0	0.0
1.0	0.1	2.0	0.0	0.0	0.0	0.0	1.0	13.0	159.0	98.0	0.0
1.0	0.0	1.0	0.0	0.0	82.0	0.0	0.0	679.0	414.0	651.0	0.0
1.0	0.0	2.0	0.0	0.0	0.0	0.0	1.0	6.0	151.0	126.0	0.0
1.0	0.0	4.0	0.0	0.0	81.0	1.0	0.0	344.0	569987.0	150.0	0.0
1.0	0.0	2.0	0.0	0.0	50.0	0.0	0.0	16.0	122.0	177.0	0.0
1.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	33.0	1078.0	76.0	0.0
1.0	0.0	0.0	0.0	0.0	71.0	0.0	0.0	72.0	1824.0	2713.0	0.0
1.0	0.0	2.0	0.0	0.0	40.0	1.0	0.0	213.0	12945.0	813.0	0.0
1.0	0.0	2.0	0.0	0.0	54.0	0.0	0.0	648.0	9884.0	1173.0	0.0

Fig. 3. Model selection

3. Conclusion

This paper presented fake account detection on Instagram using machine learning.

References

[1] Political advertising spending on Facebook between 2014 and 2018. Internet draft. [Online]. Available: https://www.statista.com/statistics/891327/political-advertising spending-face book-by-sponsor-category/

[2] J. R. Douceur, "The Sybil attack," in International workshop on peer-to-peer systems. Springer, 2002, pp. 251-260.

[3] Facebook shares drop on news of fake accounts. Internet draft. [Online]. Available: http://www.cbc.ca/news/technology/facebook-shares-drop- on news-of-fake-accounts-1.1177067

[4] R. Kaur and S. Singh, "A survey of data mining and social network analysis-based anomaly detection techniques," Egyptian informatics journal, vol. 17, no. 2, pp. 199–216, 2016.

[5] L. M. Potgieter and R. Naidoo, "Factors explaining user loyalty in a social media-based brand community," South African Journal of Information Management, vol. 19, no. 1, pp. 1–9, 2017.

[6] Quarterly earning reports. Internet draft. [Online]. Available: https://investor.fb.com/home/default.aspx

[7] Twitter: number of monthly active users 2010-2018. Internet draft. [Online]. Available: https://www.statista.com/statistics/282087/number-of- monthly active-twitter-users/

[8] Y. Boshmaf, M. Ripeanu, K. Beznosov, and E. Santos- Neto, "Thwarting fake accounts by predicting their victims," in Proceedings of the 8th ACM Workshop on Artificial Intelligence and Security. ACM, 2015, pp. 81–89.

[9] Facebook publishes enforcement numbers for the first time. Internet draft. [Online]. Available: https://newsroom.fb.com/news/2018/05/enforcement- numbers/

[10] Banque populaire dis-moi combien damis tu as sur facebook, je te dirai si ta banque va taccorder un prłt. Internet draft.[Online]. Available: http://bigbrowser.blog.lemonde.fr/2013/09/19/popularitedis-moi-combien-damis-tu-as-sur-facebook-je-te-dirai-si-ta-banqueva-taccorder-un-pret/

[11] S. T. Sun, Y. Boshmaf, K. Hawkey, and K. Beznosov, "A billion keys, but few locks: the crisis of web single sign-on," in Proceedings of the 2010 New Security Paradigms Workshop. ACM, 2010, pp. 61–72.

[12] S. Fong, Y. Zhuang, and J. He, "Not every friend on a social network can be trusted: Classifying imposters using decision trees," in Future Generation Communication Technology (FGCT), 2012 International Conference on. IEEE, 2012, pp. 58–63.

[13] K. Thomas, C. Grier, D. Song, and V. Paxson, "Suspended accounts in retrospect: an analysis of twitter spam," in Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference. ACM, 2011, pp. 243–258.

[14] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "The socialbot network: when bots socialize for fame and money," in Proceedings of the 27th annual computer security applications conference. ACM, 2011, pp. 93–102.

[15] J. Ratkiewicz, M. Conover, M. Meiss, B. Goncalves, S. Patil, A. Flammini, and F. Menczer, "Truthy: mapping the spread of astroturf in microblog streams," in Proceedings of the 20th international conference companiossn on World wide web. ACM, 2011, pp. 249–252.