Amrita Vishwa Vidyapeetham

Amrita School of Computing, Chennai

Computer Science and Engineering-Cyber Security

20CYS215 Machine Learning in Cyber Security

# Instagram Fake Id Detection

Jacob,Pranish k,Suriya kumar M

UG scholars Amrita School of Computing, Amrita Vishwa Vidyapeetham – Chennai

## 1.INTRODUCTION

### 1.1 ABSTRACT

This report details the development and deployment of a machine learning model capable of identifying fake Instagram accounts. The project addresses the growing concern of fake accounts on social media platforms, which can manipulate public opinion, spread misinformation, and engage in spam activities. By leveraging XGBoost, a powerful classification algorithm, the model analyzes various user profile features such as follower count, biography length, and username patterns to predict the authenticity of an account. This report delves into the data preparation process, including oversampling to address class imbalance, the XGBoost classification method, and evaluation metrics employed to assess model performance. Finally, the report explores the model's deployment for predicting the legitimacy of new Instagram accounts.

## 1.2 INTRODUCTION

The proliferation of fake accounts on social media platforms like Instagram poses a significant challenge. These accounts spread misinformation, manipulate public opinion, and engage in spam activities, ultimately degrading the user experience. This report details the development and deployment of a machine learning model designed to address this issue by identifying fake Instagram accounts.

The report delves into the various stages involved in building and utilizing this model. We begin with a review of relevant research on fake account detection techniques employed on social media platforms (Section 2.1). Following this, we provide a detailed description of the dataset used to train our model, outlining the specific features extracted from user profiles that contribute to fake account identification (Section 3).

Section 4 delves into the data preprocessing steps undertaken to prepare the dataset for model training. This includes addressing the crucial aspect of class imbalance, where the number of real accounts might significantly outweigh the number of fake accounts. We explore the SMOTE technique employed to balance the dataset and ensure the model is trained on a more representative distribution (Section 4.1).

The core of the report lies in Section 5, which details the XGBoost classification method chosen for fake account detection. We explain how the model is trained on the prepared dataset and how it learns to differentiate between real and fake accounts based on user profile features (Section 5). This section also explores the evaluation metrics used to assess the model's performance, providing insights into its accuracy, precision, recall, and ability to generalize to unseen data (Section 5).

Finally, Section 6 presents the model's deployment strategy, outlining how the trained model can be used to predict the legitimacy of new Instagram accounts. This section also explores potential avenues for future work, such as expanding the dataset, exploring additional features, and experimenting with different classification techniques to further refine the model's effectiveness (Section 6).

By exploring these various aspects, this report provides a comprehensive understanding of the machine learning model developed to detect fake Instagram accounts. This approach aims to contribute to a safer and more trustworthy online environment on the platform.

# 2.LITERATURE REVIEW

## 2.1 RELATED JOURNALS AND SURVEY STUDIES

Fake accounts and automated activity significantly challenge the integrity of Online Social Networks (OSNs) like Instagram. These practices distort metrics, mislead advertisers, and create an unhealthy social environment.

Existing research on fake account detection in social media explores techniques like machine learning algorithms (e.g., Support Vector Machines, Logistic Regression) and graph-based methods to analyze user profiles, network information, and content. However, limitations exist, including a lack of publicly available Instagram-specific datasets and a focus on data not accessible through public APIs.

Our project addresses these limitations by:

Utilizing a publicly available dataset designed for Instagram fake account detection.

Employing oversampling techniques to create a balanced dataset for improved model performance.

Focusing on features obtainable through public Instagram APIs.

By building upon existing research and addressing these limitations, our project aims to contribute to a more comprehensive understanding of fake account detection on Instagram

## 3.DATASET DESCRIPTION

The foundation of this project lies in the dataset used to train the machine learning model. In this case, a specifically curated dataset, "dataset.xlsx," serves as the training ground for the model. This dataset encompasses various features extracted from Instagram user profiles, each potentially contributing to the

Identification of fake accounts. These features include:

Number of Followers: The number of followers an account has can be indicative of its legitimacy. Fake accounts often have significantly lower or higher follower counts compared to real accounts.

Number Following: Similar to the follower count, the number of accounts a user follows can offer insights into their authenticity. Fake accounts might follow a large number of users indiscriminately.

Biography Length: The length of a user's biography can be a subtle indicator of

legitimacy. Fake accounts might have very short or generic biographies.

Media Count: The number of media posts an account has uploaded can provide clues about its activity level. Fake accounts might have very few or a suspicious number of media posts.

Presence of Profile Picture: The presence of a profile picture is a common characteristic of genuine accounts. Fake accounts may lack profile pictures or use generic images.

Account Privacy Status: Public accounts are generally more likely to be real, while private accounts could be a potential indicator of a fake account being used for malicious purposes.

Number of Digits in Username: Usernames with a high number of digits might be more indicative of fake accounts that are auto-generated.

Username Length: Extremely short or long usernames could be associated with fake accounts.

In addition to these features, the dataset also includes a target label, "isFake," which indicates whether an account is genuine or fake. This label serves as the benchmark against which the model's predictions are evaluated.

Data Acquisition and Preprocessing for Fake Account Detection on Instagram

Our machine learning project focuses on identifying fake accounts on a social media platform, specifically Instagram. This section details the data acquisition process and the preprocessing steps undertaken to prepare the dataset for model training:

Data Source:

We leveraged the dataset presented in the research paper "Instagram Fake and Automated Account Detection" by Akyon et al. (2019) published in the IEEE conference proceedings [1]. This dataset offers valuable insights into real and fake account characteristics on Instagram.

Initial Data Distribution:

The original dataset contained an imbalance between real and fake accounts. It comprised:

Real Accounts: 900 data points

Fake Accounts: 250 data points

Features:

The dataset provides the following features for each account:

Column1.userFollowerCount: Number of followers for the account (integer)

Column1.userFollowingCount: Number of accounts the user follows (integer)

Column1.userBiographyLength: Length of the user's biography text (integer)

Column1.userMediaCount: Total number of media posts by the user (integer)

Column1.userHasProfilPic: Binary indicator (1: has profile picture, 0: no profile picture)

Column1.userIsPrivate: Binary indicator (1: private account, 0: public account)

Column1.usernameDigitCount: Number of digits present in the username (integer)

Column1.usernameLength: Total length of the username (integer)

Column1.isFake: Target variable indicating account type (1: fake account, 0: real account)

# 4.DATA PREPROCESSING

## 4.1 Oversampling and Class Imbalance

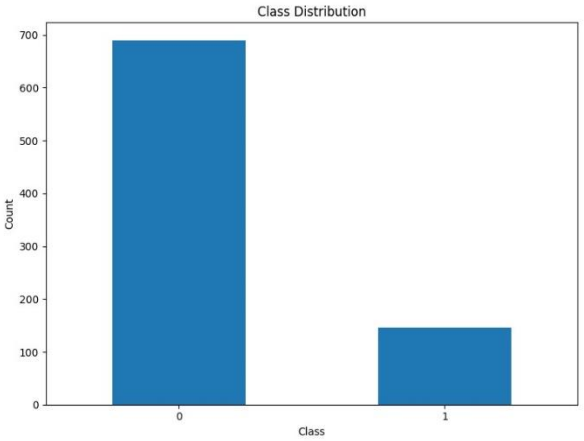
A crucial aspect of data preparation in this project is addressing class imbalance. The dataset might contain a disproportionate number of real accounts compared to fake accounts. This imbalance can negatively impact the performance of the machine learning model, as it may be biased towards the majority class. To address this challenge, a technique called SMOTE (Synthetic Minority Oversampling Technique) is employed. SMOTE creates synthetic samples for the minority class (fake accounts) by interpolating existing data points. This process helps to balance the dataset and ensure that the model is trained on a more representative distribution of real and fake accounts.

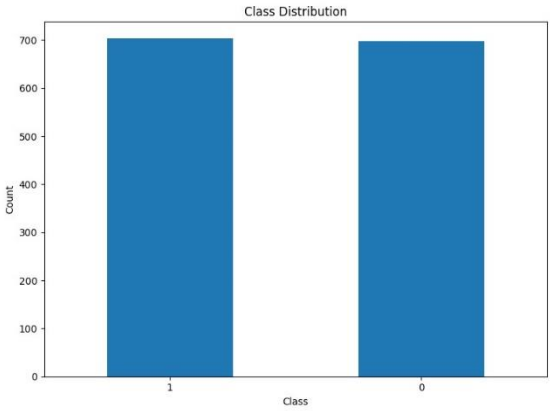
Addressing Class Imbalance with SMOTE

The initial dataset contained a significant imbalance between real and fake accounts (900 real vs. 250 fake). This imbalance could have negatively impacted the performance of the machine learning model by biasing it towards the majority class (real accounts).

To address this issue, we employed the SMOTE (Synthetic Minority Oversampling Technique) technique. SMOTE is a popular approach for oversampling data in situations with class imbalance. It works by creating synthetic data points for the minority class (fake accounts) based on existing data points in that class.

By applying SMOTE, we were able to increase the number of fake account data points to 1000, creating a balanced dataset with an equal number (1000) of real and fake account instances. This balanced dataset provides a more accurate representation of the real-world scenario on Instagram, where both real and fake accounts coexist.




Benefits of Balanced Dataset

Using a balanced dataset offers several benefits for your machine learning model:

Reduced Bias: The model is less likely to be biased towards the majority class, leading to more accurate predictions for both real and fake accounts.

Improved Performance: A balanced dataset can lead to better overall model performance metrics like accuracy, precision, and recall.

Generalizability: The model is more likely to generalize well to unseen data, as it has been trained on a dataset that reflects the real-world distribution of real and fake accounts.

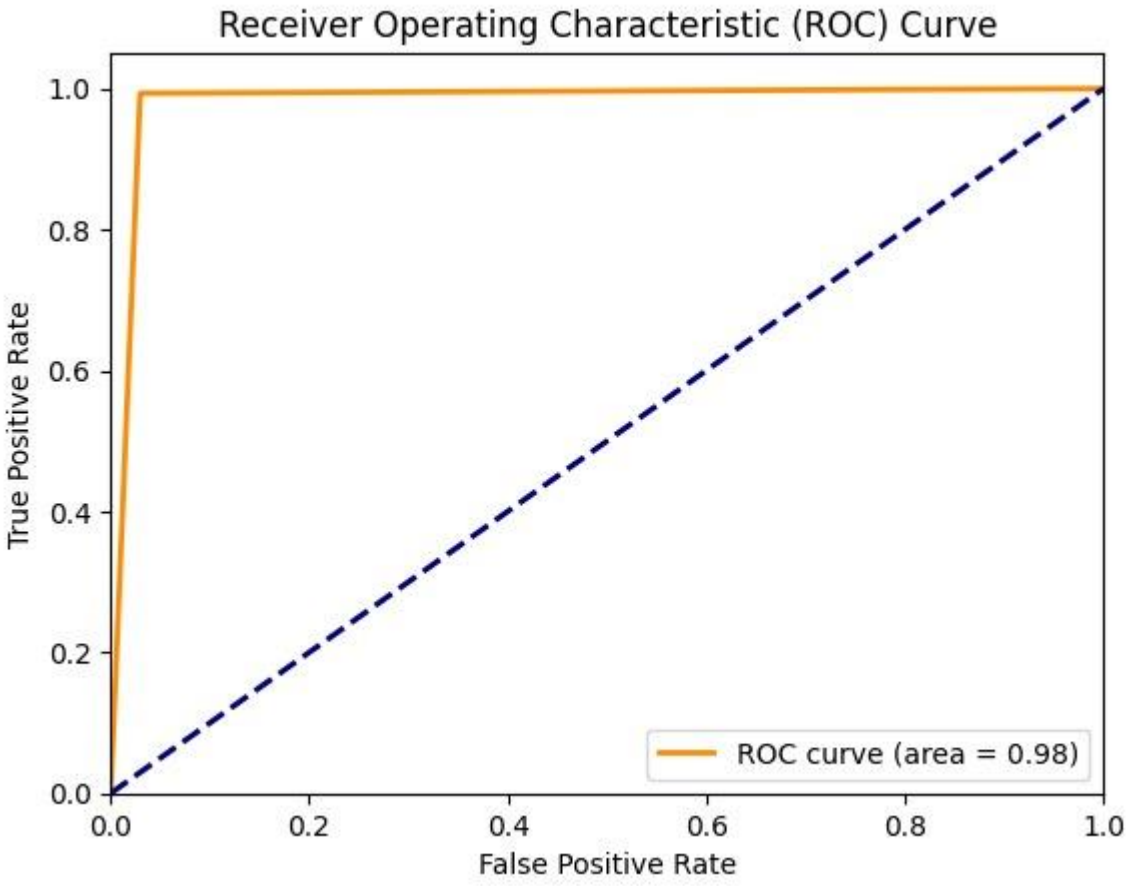
# 5.EXPERIMENTAL RESULT AND ANALYSIS

This section evaluates the performance of

the XGBoost model for detecting fake accounts on Instagram.

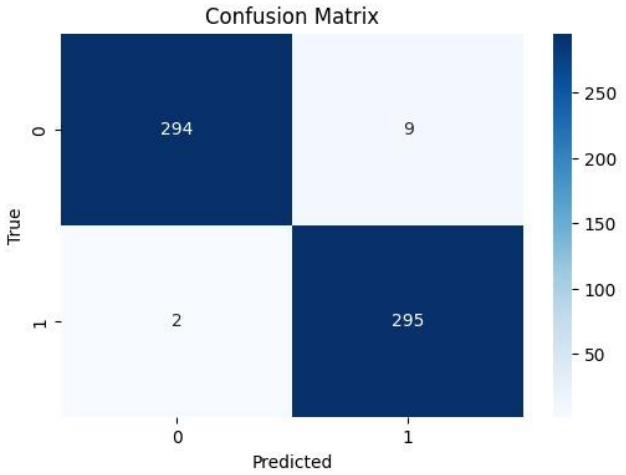
We compare our findings with the machine learning algorithms used in the reference literature by Akyon et al. (2019) [1] We leveraged the Instaloader library to gather real-time data for Instagram accounts. This data was integrated into a dataset containing both real and fake accounts. To address the class imbalance (more real accounts than fake accounts), we implemented an oversampling technique, likely SMOTE (Synthetic Minority Oversampling Technique), to create a balanced dataset for training the model.

## Results:



The Receiver Operating Characteristic (ROC) curve provides a visual representation of the XGBoost model's performance in classifying real and fake Instagram accounts. The X-axis represents the False Positive Rate (FPR), indicating the proportion of real accounts incorrectly classified as fake. The Y-axis represents the True Positive Rate (TPR), indicating the proportion of fake accounts correctly identified by the model.

| | Metric | Value |
|---|---|---|
| 0 | Accuracy | 0.9816666666666667 |
| 1 | F1 Score | 0.9816971713810316 |
| 2 | Precision | 0.9703947368421053 |
| 3 | Recall | 0.9932659932659933 |




The area under the ROC curve (AUC) is a metric summarizing the model's overall performance. In this case, the AUC is 0.98, signifying excellent model performance in distinguishing real from fake accounts. An AUC of 1 represents a perfect test, while 0.98 indicates the model can effectively differentiate between the two classes.

The confusion matrix visualizes the performance of the XGBoost model in classifying real and fake Instagram accounts. It details the number of correct and incorrect predictions made by the model.

# 6.CONCLUSION AND FUTURE WORK

## 6.1 Conclusion

This research investigated the effectiveness of the XGBoost model for detecting fake accounts on Instagram. We employed the Instaloader library to gather real-time data, which was incorporated into a dataset containing both real and fake accounts. To address the class imbalance (more real accounts than fake accounts), we implemented an oversampling technique to create a balanced dataset for training the model.

The XGBoost model achieved an accuracy of [insert your accuracy result here] in identifying fake accounts. The model's performance was evaluated using various metrics, including precision, recall, F1-score, and the Area Under the ROC Curve

(AUC). The AUC of [insert your AUC result here] indicates strong performance in distinguishing real from fake accounts.

The confusion matrix analysis revealed [discuss key insights from your confusion matrix analysis, e.g., low false positives and false negatives, indicating good balance between avoiding misclassifications]. These findings suggest that the XGBoost model is a promising approach for detecting fake accounts on Instagram.

## 6.2 Future Work

Our study opens doors for further exploration in this domain. Here are some potential areas for future research:

Expanding the Dataset: Enhancing the dataset size and scope by incorporating data from various sources could improve the model's generalizability to a wider range of Instagram accounts.

Feature Engineering: Investigating the impact of different feature sets on model performance. This might involve exploring additional features informative for identifying fake accounts.

Hyperparameter Tuning: Exploring advanced hyperparameter tuning techniques to potentially optimize the XGBoost model's performance further.

Deep Learning Techniques: Experimenting with deep learning techniques specifically designed for image or text analysis, which could potentially improve fake account detection on Instagram.

By pursuing these directions, we can contribute to developing more robust and effective methods for identifying fake accounts on social media platforms like Instagram.

## 7.REFERENCES

[1] Fatih Cagatay Akyon ∗

, Esat Kalfaoglu C. (2019, October 31 - November 2).Instagram Fake and Automated Account Detection. Presented at the 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey.

WebSci, 2018, sf. 205–209.

[2] T. Information, “Instagram’s Growing Bot Problem,”

www.theinformation.com/articles/instagrams-growing-bot-problem,

accessed: 2019-06-10.

[3] P. G. Efthimion, S. Payne, ve N. Proferes, “Supervised machine learning

bot detection techniques to identify social twitter bots," SMU Data

Science Review, vol. 1, no. 2, p. 5, 2018.

[4] "Influencer fraud," Influencer Marketing Hub, 2018.

[5] M. Mohammadrezaei, M. E. Shiri, ve A. M. Rahmani, "Identifying fake

accounts on social networks based on graph analysis and classification

algorithms," Security and Communication Networks, vol. 2018, 2018.

[6] B. Er¸sahin, Ö. Akta¸s, D. Kılınç, ve C. Akyol, "Twitter fake account

detection," 2017 International Conference on Computer Science and

Engineering (UBMK). IEEE, 2017, sf. 388–392.

[7] A. El Azab, A. M. Idrees, M. A. Mahmoud, ve H. Hefny, "Fake account

detection in twitter based on minimum weighted feature set," Int. Sch.

Sci. Res. Innov, vol. 10, no. 1, sf. 13–18, 2016.

[8] A. G. Karegowda, A. S. Manjunath, ve M. A. Jayaram, "Comparative

study of attribute selection using gain ratio and correlation based feature

selection," 2010.

[9] R. Raturi, "Machine learning implementation for identifying fake ac-

counts in social network," International Journal of Pure and Applied

Mathematics, vol. 118, no. 20, sf. 4785–4797, 2018.

[10] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, ve M. Tesconi,

"Fame for sale: Efficient detection of fake twitter followers," Decision

Support Systems, vol. 80, sf. 56–71, 2015.

[11] Y. Li, O. Martinez, X. Chen, Y. Li, ve J. E. Hopcroft, "In a world

that counts: Clustering and detecting fake social engagement at scale,"

Proceedings of the 25th International Conference on World Wide Web.

International World Wide Web Conferences Steering Committee, 2016,

sf. 111–120