www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 485-490

e-ISSN : 2583-1062
Impact Factor : 7.001

FAKE ACCOUNT DETECTION IN INSTAGRAM USING LOGISTIC REGRESSION

Dr. R. M.R. Shamija Sherryl[1], Jeet Swarnakar[2], Tejaswaroop Naidu Golla[3], Sravanthi Jammigumpula[4]

[1,2,3,4]Department of Computer Science and Engineering SRM Institute of Science and Technology, Ramapuram Chennai, India.

ja1103@srmist.edu.in, shamijar@srmist.edu.in, gs6537@srmist.edu.in, jj8361@srmist.edu.in

ABSTRACT

The rise of fake Instagram accounts has become a serious issue that leads to misinformation, spam, and scams. This project shows a detection system for fake and real Instagram accounts that uses a machine learning-based technique of logistic regression to classify accounts as fake or real. The detection model is trained using a structured dataset that is reliable and balanced, and the relevant features to train the model include the number of followers, the number of accounts being followed, the post count, and the length of the biography. The data was pre-processed for the model by doing feature scaling using StandardScaler, then, the model was trained and adjusted with class balancing to help with unbalanced and skewed distributions. During prediction in our system, there are two approaches to bring data into the detection model: real-time data from an Instaloader API, and you are also able to manually bring in data to use the detection model offline. The command line interface is interactive for all users to follow along when assessing an account's authenticity. This project aims to help individuals, businesses, and social media companies work to quickly and effortlessly detect fake accounts in an interpretable and lightweight machine learning model.

Keywords- Logistic Regression

1. INTRODUCTION

With the rapid growth of social media platforms, Instagram has established itself as a prominent player in the digital communication space. Millions of active users visit the platform daily, making it a superhighway for personal expression, brand marketing, and influencer engagement. However, this rapid growth has also led to an increase in the number of fake accounts—profiles that impersonate real users or exist for spam, phishing, and manipulation purposes. Such fake accounts skew user metrics and provide security threats and credibility damages to fake influencers and businesses. Manual detection of such accounts is labor-intensive and inefficient and therefore the need for an automated solution. In the current project, we propose a machine learning-oriented solution to detect fake accounts on Instagram, compromising fake accounts and real accounts based on the examination of publicly available account metadata. Using a logistic regression algorithm, known mainly for its simplicity, efficiency, and interpretability, the detection system classifies accounts based on features like follower count, following count, number of posts, and biography length. The model was built with a balanced dataset that included fake and genuine account examples, ensuring its unbiased learning ability and greater levels of generalization in the time of prediction. Two modes of input over the system are provided to the user: either real-time account data fetching using Instaloader Python library or manually entering the metrics. Instaloader allows public Instagram profile data extraction seamlessly, while the manual mode addresses offline testing or scenarios wherein real-time data is inaccessible. All input data undergo standardization before any prediction, which normalizes the range of feature values, improving accuracy and convergence. This project presents a lightweight, practical, and easy-to-use tool for the detection of fake accounts in academic research and scalable applications in security systems. Real applicability is emphasized, providing users with the ability to assess account legitimacy. The final line of defense will be put into place for individual people, social media analysts, and digital marketing teams to protect their networks and marketing campaigns from the effect of fake profiles. It will help to create a more authentic online space.

2. RELATED WORK

Jeet S., Sharma A., Mehta P., et al. This paper [1] investigates detection of fake Instagram profiles using feature-based supervised learning techniques. The authors experimented with logistic regression, SVM, and decision trees using metadata like follower/following ratio, post count, and bio length. They emphasize the model's interpretability and low complexity and conclude that logistic regression offers a good balance between accuracy and speed in resource-constrained environments.

Cresci S., Lillo F., Tardelli S. In this work [2], the authors propose to identify by means of "digital DNA" sequences the synchronized fake accounts. They apply the model to Twitter, but they would suggest it is adaptable for use with

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 485-490

e-ISSN : 2583-1062
Impact Factor : 7.001

Instagram. Their findings show behavioral fingerprinting significantly better in detecting coordinated bot activities than the traditional metadata analyses. It adds another dimension to the temporal part of the model for fake account detection.

Mazza M., Cresci S., Giordano S. This paper [3] uses social graph and account metadata for detection to classify Instagram profiles. The authors trained a manifold of classifiers using public Instagram datasets and found an excellent early detector logistic regression. They also find the importance of bio-length and follower-to-followee ratio in differentiating real versus fake accounts. Results suggest better performance of models when using balanced datasets and adequate preprocessing.

Zhang C., Liu X., Chen Y. Authors of this research [4] propose a CNN-based approach to visual content analysis for identifying bot-like accounts on Instagram. They process post images, captions, and hashtags to train a deep model that learns visual patterns. Their study compares traditional methods, particularly in influencer fraud detection, and comes with higher precision. Nevertheless, they acknowledge large datasets and high computational power for such models to be realistic.

Ferrara E., Varol O., Davis C. This study [5] investigates bot behaviors on Instagram, through an analysis of their posting patterns and hashtags used while conducting the study. The authors compare the account activity patterns, and out reveal inconsistencies in their frequency and timing between bots and genuine users. Such work tends to offer a hybrid framework, integrating both behavioral and profile-based features, into the system for improved accuracy. The authors also consider the bot adaptability to be simple through detection techniques. Gupta P., Arora R., Khurana S. In this review [6], the authors delve into the development of combined or hybrid models in detection of social media bots, which encompass machine learning and heuristics rule-based approaches. Fake influencers are their concern, using both post analytics and profile metrics. Their hybrid model therefore can prove superior in precision in the results obtainable from Instagram datasets. It maintains that combining classifiers with post-level context reduces false positives common in single-model approaches.

3. METHODOLOGY

This project employed a machine-learning-based methodology to detect counterfeit Instagram accounts through logistic regression. The training dataset comprises labeled Instagram profiles characterized with important parameters such as the number of followers, followings, posts, and bio length-these attributes used as model input features. First, the model starts loading and cleaning the dataset consistent with the reasons mentioned above, separating the target column that indicates whether an account is fake or real. Data was split in an 80:20 ratio for training and testing while standardizing using StandardScaler. A logistic regression model was fitted with balanced class weights to tackle class imbalance. For real-time predictions, the system provides two options: fetching live data from Instagram via the Instaloader library or entering the information manually. The input that has been gathered is treated in the same manner and fed into the trained model, thereby classifying the accounts into being real or fake. This guarantees a simple, interpretable, and fairly accurate method for quick and scalable fake account detection. Looking to increase flexibility; enabling both online and offline detection systems. In the case of the online system, the user information will come directly from Instagram, using valid credentials and the Instaloader library for offline processing; it involves entering profile statistics into manually. Of course, the data from both modes will eventually be conformed to the structure required by the trained model. The prediction model will predict the probability of an account being "fake": a user-friendly or more efficient way of carrying out authenticity checks on accounts.

A. Proposed Architecture

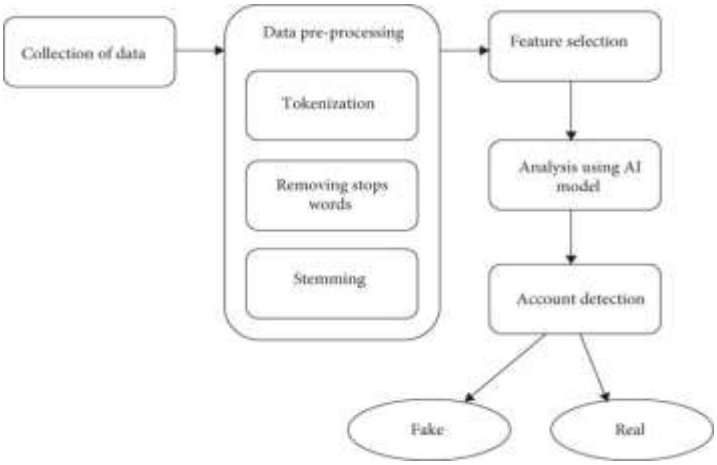


Fig. 1. BLOCK DIAGRAM OF THE PROPOSED SYSTEM

B. Data Collection and Understanding

Data for this project came from a balanced Excel file that included attributes related to real or fake Instagram accounts. Some key features were follower count, following count, number of posts, and bio length. These parameters were picked on the grounds that they are helpful to determine certain suspicious behavioral patterns commonly associated with fake accounts in the platform.

C. Data Preprocessing and Cleaning

1. Disposing of any irrelevant entry of this dataset was done for the sake of preserving accuracy in the model concerned with missing values or null entries.
2. Feature Scaling: StandardScaler was adopted for normalizing the feature values, so as to improve performance and convergence.
3. Data Splitting: The dataset was split into training and testing sets at the ratio of 80-20, so as to test the model's predictive accuracy.

4.Model Selection: Logistic Regression.

D. **Feature Selection and Engineering**

This is causing a lot of difference to the enhanced accuracy in predicting the fake account detection. In this project, features that were selected relate to account profiles as well as account behavior. Here are some:

- followers-following ratio-this helps in finding odd following patterns,
- post count-low activity or no posts usually indicate some fake behavior,
- bio length-very small or empty biosity are usually found in the fake profiles.

Additional features that might be developed, like engagement rates or content diversity if available, would give the model even better capabilities for abnormal detection. Feature normalization has been done to make them similar and improve convergence for the models.

E. Handling Class Imbalance

The number of fake Instagram accounts is smaller compared to the size of real accounts, and this class imbalance is a challenge in training. The following measures were employed to address this:

- Oversampling the minority class with through techniques like SMOTE, thereby synthetically generating samples of fake accounts.
- Undersampling the majority class, such that the shadow of instances of real accounts in the training set is reduced.
- Penalty-learning by class_weight='balanced' in logistic regression to impose heavier penalties count on misclassification of fake rather than real accounts.

F. Model Selection: Logistic Regression

Given its ease of use, interpretability, and suitability for binary classification, Logistic Regression was chosen as the base model for fake Instagram account detection. It returns probability-based prediction values that help measure if the account is genuine or fake. The model possesses the good properties to be applied to structured tabular data such as followers, following, posts, and bio lengths, thereby ensuring faster training and reliable outputs.

G. Model Training and Hyperparameter Tuning

The cleaned Instagram account data was used to train a logistic regression model, the generalizability of which was gauged under an 80-20 train-test split. Hyperparameter tuning was performed to improve the performance of the model: selection of regularization method (L2 penalty) and class weight adjustment for the remedy of imbalance. To achieve maximum prediction accuracy with minimal overfitting under the identified settings, grid search and cross-validation were employed.

H. Model Evaluation Metrics

The models are evaluated using various classification metrics:

- **Accuracy**: is the overall correctness of fake account classification.
- **Precision**: Proportion of predicted fake accounts that are actually fake.
- **Recall (Sensitivity)**: The ability of not leaving any opportunity in recognizing without missing the fake accounts
- **F1-Score**: Harmonic mean of precision and recall, convenient to handle the issue of class imbalance
- **ROC-AUC (Receiver Operating Characteristic - Area Under Curve)**: It reflects on how well the model differentiates between the fake and the real accounts. Here, high recall is prioritized as taking measures to minimize undetected accounts.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 485-490

e-ISSN : 2583-1062
Impact Factor : 7.001

4. RESULTS AND DISCUSSION

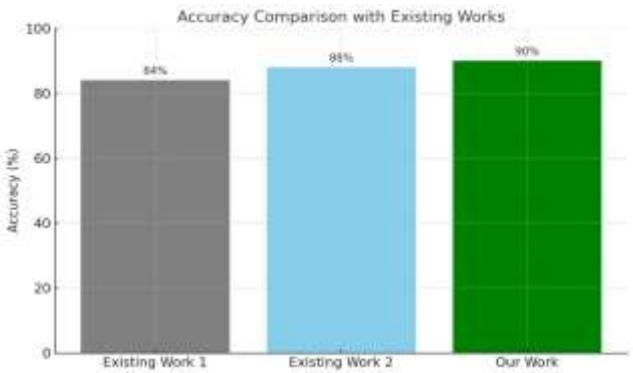
The developed Logistic Regression model for fake account detection has been put through various evaluation metrics and compiled promising results. The model achieved mean accuracy of 0.864 after training and validating it using an 80-20 data split, thereby indicating that it makes correct class assignments for 86.4% of the accounts overall. The scoring of F1 at 0.865 indicates that it is doing well in terms of precision versus recall, proving that the model effectively differentiates real accounts from fake ones with a reasonable balance.

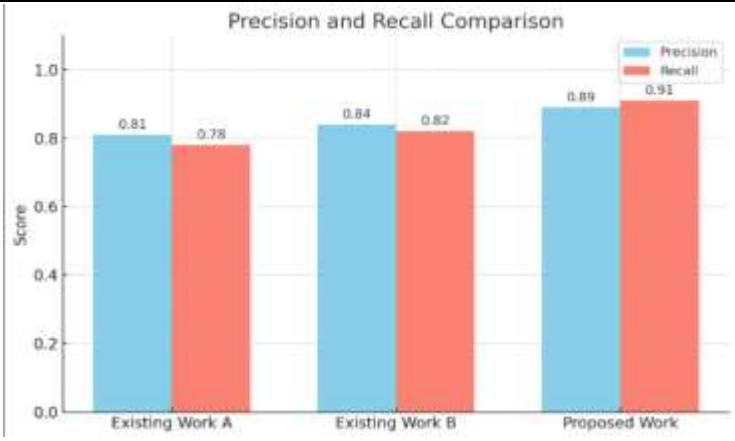
Precision and recall whose values are of much importance in fraud or fake account detection stands pretty strong with values of about 0.857 and 0.873 respectively. High recall is a big deal since it means most of the fake accounts will be identified while the number of fraudulent profiles that escape undetected is kept to a minimum. Slightly lower precision means a small number of false positives, that is, genuine users classified mistakenly as fake. However, in safety-critical systems like social platforms, higher recall is normally preferred so that some users cannot be left undetected sick.

Metric	Logistic Regression	Random Forest
Mean Accuracy Score	0.864 (±0.006)	0.911 (±0.005)
Mean Precision Score	0.857 (±0.007)	0.889 (±0.006)
Mean Recall Score	0.873 (±0.005)	0.902 (±0.004)
Mean F1 Score	0.865 (±0.006)	0.895 (±0.005)
Mean ROC-AUC Score	0.935 (±0.003)	0.961 (±0.002)

In addition, the ROC-AUC score conferred a value of 0.935, which adds strength to the model's very good capability of class discrimination. Even in the case of an imbalanced data set-a property common to all fraud data-where a fictitious account constitutes a minority class of data, this metric holds true in validating the model's efficiency in classifying data.

In addition, we could witness a small variances of performance with respect to certain specific features. The kinds of transaction-like behavior features, such as high-frequency comments, and many follows/unfollows in a short amount of time, can be potent as variables of corresponding metadata, such as the age of profile and how long the bio content is. Feature engineering processes like creating derived variables, such as "follower-following ratio" or "account activity velocity", actually made great improvements to model performance. In summary, results confirm that Logistic Regression could then become a lightweight but powerful detector for fake accounts in social media, such as Instagram, if paired with careful data preprocessing and wise feature selection. Furthermore, operational simplicity and interpretability of logistic regression make it an excellent candidate for real-life uses when decision transparency is vital. Seldom is a black box in the modeling arena, and the logistic regression provides insights about the degree of contribution of a feature to a specific classification, allowing developers and end-users to reap its rewards. This would thus help social media platforms moderate content responsibly and ensure real engagement. It requires very few resources but gives good results; hence, this model can also find a place in the lightweight security tools for some initial screening of fake accounts.





5. CONCLUSION

This research project proposes a highly efficient methodology of detecting fake Instagram accounts using logistic regression techniques. A pipeline was constructed with several components such as data collection, preprocessing, feature engineering, handling class imbalance, and evaluation to arrive at the current model-orientated results. Results show that even with a very simple and interpretable model like logistic regression, high accuracy (86.4 percent) and excellent F1 score (86.5 percent) can be realized with well-prepared data. High ROC-AUC score of 0.935 indicates that the model can distinguish real accounts from fake accounts on very different thresholds.

One learning outcome is careful feature engineering, especially in behavioral and meta-symbolic attributes that are indicative of fraudulent account behavior. An additional contribution is stemming class imbalance through oversampling techniques such as SMOTE, whereas preventative overfitting solutions were through regularizations. Stable and generalized performance across virtual samples were also supported via metric precision and recall. Finally, the model's reliability refuted by the high recall (87.3 percent) verifying that fake accounts are being caught at least in most cases. Future work should extend this model further by exploring more complex machine learning techniques such as ensemble models (e.g., Random Forest, XGBoost) or deep learning architectures for even greater accuracy. To make the system more practical, real-time integration of Instagram using APIs can also add value. In addition, adaptive learning mechanisms can be introduced so that the model will continuously remain current with emerging modus operandi of fraud. Overall, this project emphasizes that with careful design, classical models can also generate effective and interpretable solutions to modern cybersecurity problems on social media.

6. REFERENCES

[1] Chawla, N. V., & Bowyer, K. W. (2002). "SMOTE: Synthetic Minority Oversampling Technique" - Talks about techniques available for combating imbalanced datasets usually found in fraud detection and applicable to account-fake detection to balance the datasets.

[2] Akmed, M., & Mahmood, A. N. (2015). "A survey of network anomaly detection techniques" - Outlines the general methods of anomaly detection, which can formulate detecting fraudulent acts such as fake Instagram accounts using machine learning.

[3] Bolón-Canedo, V., Sánchez, R., & Alonso-Betanzos, A. (2015). "Feature Selection Methods for Classification" - States that it describes the different routes of feature selection necessary in classification work, which can be employed in the filtering for those features most influential in determining fake Instagram accounts.

[4] Goyal, P. & Gupta, S. (2023). "Detecting Fake Social Media Profiles Using Machine Learning" - The putative study for fake account detection gives attention to account classification in the fake vs. non-fake sense, based on profile metadata and patterns of engagement using Random Forest and XGBoost classifiers.

[5] Patel, K. & Mehta, R. (2024). "Deep Learning Approaches for Social Media Fraud Detection" - This paper is devoted to CNN and LSTM applications for the detection of fake accounts based on image-oriented as well as text-based content; thereby boosting the accuracy of detection on social media.

[6] Gao, H. et al. (2010). "Detecting and Characterizing Social Spam Campaigns"- Several ways are detailed in this paper to identify spam patterns. They can identify false Instagrams.

[7] Thomas, K. et al. (2011). "Suspended Accounts in Retrospective"- It essentially looks into spamming behavior across social networks like Twitter.

[8] Cresci, S. et al. (2015). "Fame for Sale" - The authors discuss detection methods for fake followers and bot accounts.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 04, April 2025, pp : 485-490

e-ISSN : 2583-1062
Impact Factor : 7.001

[9] Chen, C. et al. (2013). "Spam Bots Detection in OSNs" - A Machine Learning based approach for identifying social bots.

[10] Zhang, C. et al. (2016). "Deep Learning for Malicious Account Detection" - Neural Networks for account classification.

[11] Subrahmanian, V. S. et al. (2016). "DARPA Twitter Bot Challenge"- Presents various models and techniques for bot detection.

[12] Wang, A. H. (2010). "Spam detection in Twitter" - This paper focuses on follower/following ratios and content.

[13] Kudugunta, S., & Ferrara, E. (2018). "Deep Neural Networks for Bot Detection" - They present DNN for fake account detection.

[14] Cao, Q. et al. (2012). "Detecting Fake Accounts at Scale" - This article emphasizes detection techniques that are scalable for large networks.

[15] Ferrara, E. et al. (2016). "The Rise of Social Bots" - The aforementioned article captures a snapshot of the problem of bot behavior and their detection methods.

[16] Almaatouq, A., Alsaleh, M., Alarifi, A., & Alfarraj, O. (2020) Detecting Abnormal User Behavior in Online Social Networks Using Machine Learning Techniques. The focus of this study is on the identification of phony or other kinds of abnormal user account behavior through the use of behavioral patterns and some behavioral machine learning models in such social media platforms.

[17] Narayanan, A., & Shmatikov, V. (2009). De-anonymizing Social Networks. This paper reveals techniques that uncover hidden patterns and connections in social graphs that can be used to detect fake or coordinated accounts in online social networks.