International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

Fake Profile Detection on Social Media Using Machine Learning

Sakshi Tapkir

Department of Computer Science, JSPM University Pune
MSc / MCA – Academic Year 2025–26

Abstract: ***Social media platforms are growing fast. This growth has led to fake user profiles. These profiles are used for stealing identities and spreading information. It is hard for people to find these accounts manually because there are users and the behavior of these accounts is always changing. This paper talks about a system that uses machine learning to find social media profiles. The system looks at things like the number of followers the number of accounts the profile follows, how often the profile posts, how old the account is, if the profile has a picture and if the account is verified. The system uses these things to train and test three machine learning algorithms: Logistic Regression, Support Vector Machine and Random Forest. When we tested the system, we found that the Random Forest algorithm was the best at finding profiles. It was correct 95.8% of the time. The system is a way to make social media more secure and trustworthy.***

Keywords: Fake profile detection, social media security, machine learning, Random Forest, Support Vector Machine, Logistic Regression, feature extraction, classification

1. Introduction

Social media platforms like Facebook, Instagram, Twitter and LinkedIn are very important for communication and community building. Because they are open and have users, they are also used by bad people who create fake profiles to do bad things. These fake accounts are used for scams pretending to be people and spreading false information.

1.1 Problem Statement

Fake social media profiles usually have some things that're not normal, like very few followers, a high ratio of accounts they follow to followers, a very recent creation date and no profile picture. Traditional systems that try to find these profiles by looking at these things are not very good because the people who make the profiles are always trying to evade detection.

1.2 Proposed Approach

Machine learning is a way to solve this problem. By looking at things about a profile and its behavior the system can learn to tell the difference between fake profiles. This paper talks about a system that uses machine learning to find profiles.

1.3 Contributions

The system looks at things about a profile like the number of followers and the number of posts. It uses three different machine learning algorithms to find the profiles. The system is very good at finding profiles. Can be used in the real world.

2. Literature Review

There has been a lot of research on finding profiles and bot accounts on media. The research has gone from rules to complex machine learning systems.

2.1 Early Heuristic Approaches

At first researchers used rules to find accounts. They looked at things like the number of followers and the number of posts. These systems were not very good because the people who made the profiles could easily evade detection.

2.2 Graph-Based Approaches

Then researchers looked at the connections between profiles on media. They thought that fake accounts would have connection patterns than real accounts. These systems were not very good either because they were slow and needed a lot of data.

2.3 Machine Learning Approaches

Now researchers use machine learning to find profiles. They look at things about a profile and its behavior. The machine learning algorithms can learn to tell the difference between profiles.

2.4 Research Gap

There is still a lot of work to be done to make systems that can find profiles. Many systems are complex. Need a lot of data. This paper tries to fill this gap by talking about a system that uses machine learning to find profiles.

3. Research Objectives

Objectives
To make a system that can find profiles automatically.

- To look at things about a profile and its behavior.
- To use machine learning to find profiles.
- To compare the performance of machine learning algorithms.
- To make a system that can be used in the world.

4. Methodology

The system follows these steps:

- Data collection: We collected data from media platforms.
- Feature extraction: We looked at things about each profile.
- Preprocessing: We normalized the data.
- Model training: We trained three machine learning algorithms.
- Evaluation: We tested the performance of each algorithm.

Figure 1.

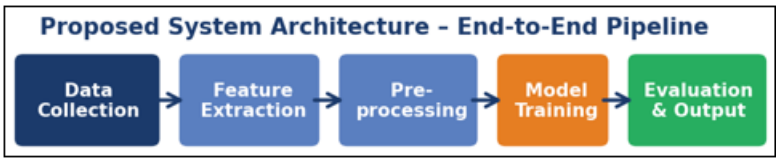

Figure 1: Proposed end-to-end system architecture for fake profile detection

4.1 Dataset

We used a dataset of 2,000 social media profiles, 1,000 1,000 fake. We split the data into training and test sets.

Table 1: Dataset distribution across training and test subsets

	Total Records	Genuine	Fake
	1,600	800	800
Test Set (20%)	400	200	200
Total	2,000	1,000	1,000

4.2 Feature Extraction

We looked at eight things about each profile:
1) Number of Followers
2) Number of Following
3) Follower-Following Ratio
4) Total Posts
5) Account Age
6) Profile Picture Availability
7) Verification Status
8) Posting Frequency

Table 2: Feature set used for model training and evaluation

#	Feature	Description	Type
1	Number of Followers	Total accounts following this profile	Numeric
2	Number of Following	Total accounts this profile follows	Numeric
3	Follower–Following Ratio	Followers ÷ Following	Derived Numeric
4	Total Posts	Cumulative post count on the account	Numeric
5	Account Age (days)	Days since account creation	Numeric
6	Profile Picture Availability	Binary flag: picture present or absent	Binary
7	Verification Status	Official verification badge: yes/no	Binary
8	Posting Frequency	Average posts per day over account lifespan	Derived Numeric

4.3 Preprocessing

We normalized the data. We filled in missing values. We scaled the features.

4.4 Model Implementation

We used three machine learning algorithms:
- Logistic Regression
- Support Vector Machine
- Random Forest

We selected the parameters for each algorithm using cross-validation.

5. Results & Analysis

5.1 Classification Accuracy

The Random Forest algorithm was the best at finding profiles. It was correct 95.8% of the time.

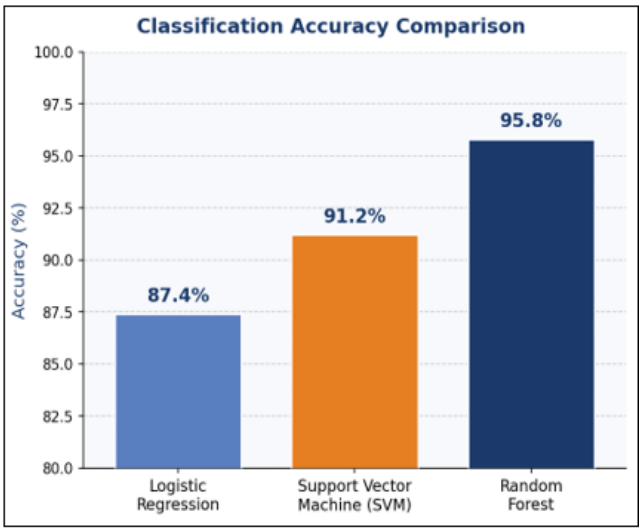

Figure 2: Classification accuracy comparison across the three evaluated models.

5.2 Precision, Recall & F1-Score

The Random Forest algorithm was also the best at precision, recall and F1-score.

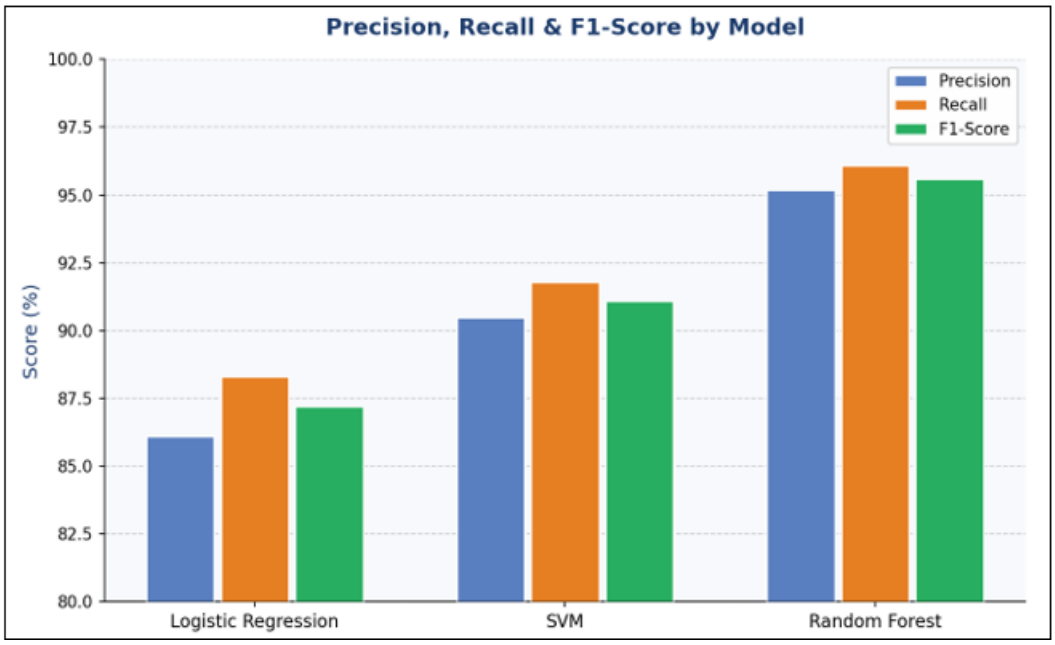

Figure 3: Precision, Recall, and F1-Score for all models on the test set.

Table 3: Comprehensive evaluation metrics. ★ Best performing model.

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	87.40%	86.10%	88.30%	87.20%
Support Vector Machine	91.20%	90.50%	91.80%	91.10%
Random Forest ★	95.80%	95.20%	96.10%	95.60%

5.3 Looking at ROC Curve Analysis

The ROC curve shows how well a model is doing. Figure 4 has the ROC curves for all the models. The Area Under the Curve for Random Forest is the highest, which means it is really good at telling the difference between profiles.

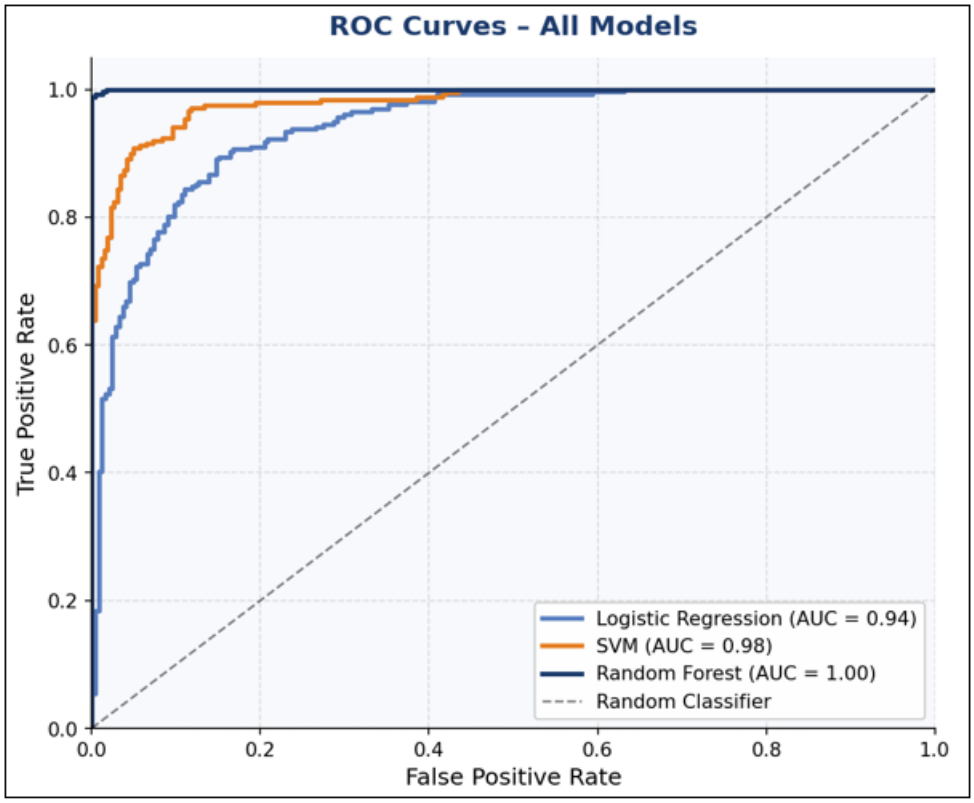

Figure 4: ROC curves for all classifiers. Higher AUC indicates better discrimination between classes

5.4 Confusion Matrix – Random Forest

Figure 5 shows the confusion matrix for the Random Forest model. Out of 200 profiles 185 were correctly. Only 8 were thought to be fake. Out of 200 profiles 201 were correctly. Only 6 were missed.

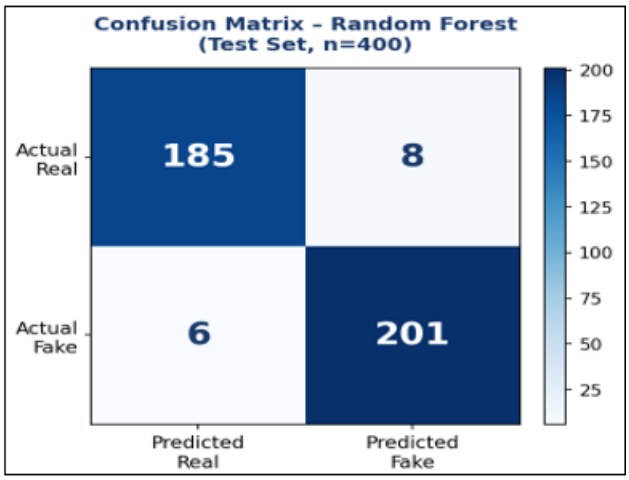

Figure 5: Confusion matrix for the Random Forest classifier on the test set (n = 400)

5.5 Looking at Feature Importance Analysis

Figure 6 shows how important each feature is. The follower-to-following ratio is the important feature, which makes sense because fake profiles often have a lot of followers but do not follow many people. The age of the account and whether the profile has a picture are also important

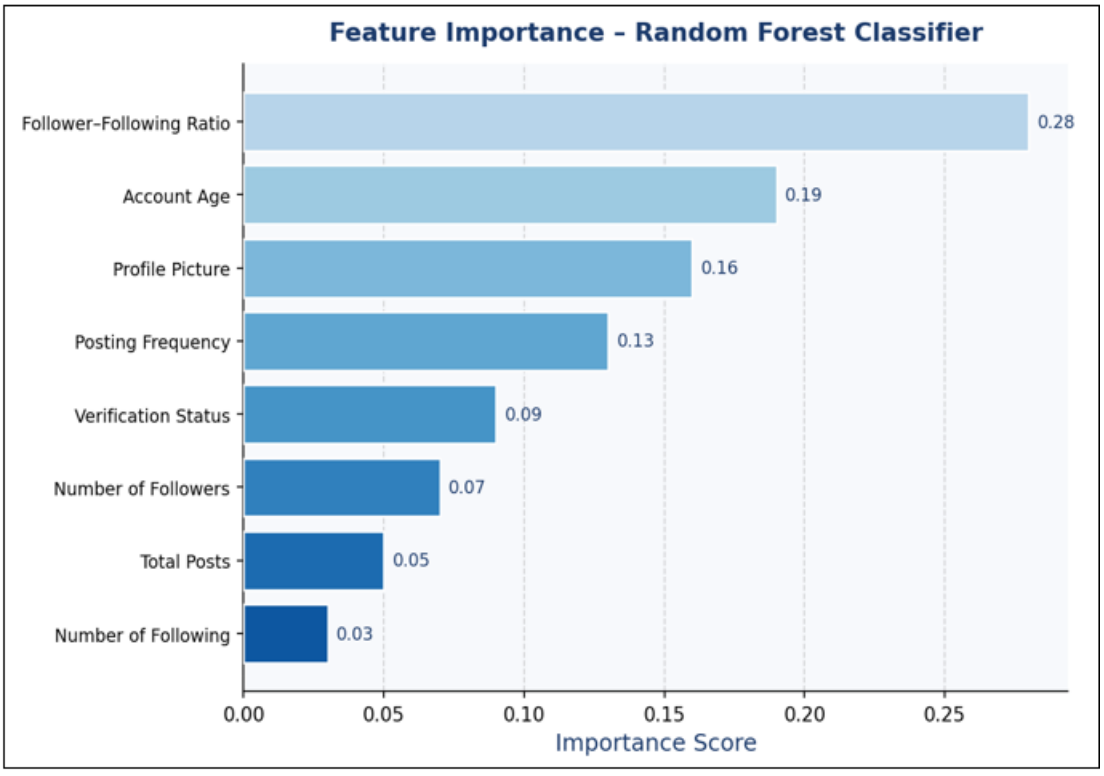

Figure 6: Feature importance scores from the Random Forest classifier (normalized to sum to 1.0)

5.6 Training Convergence

Figure 7 shows the training and validation loss curves for the Random Forest model. Both curves are going down which means the model is getting better and better. The gap between the two curves is small which means the model is not overfitting.

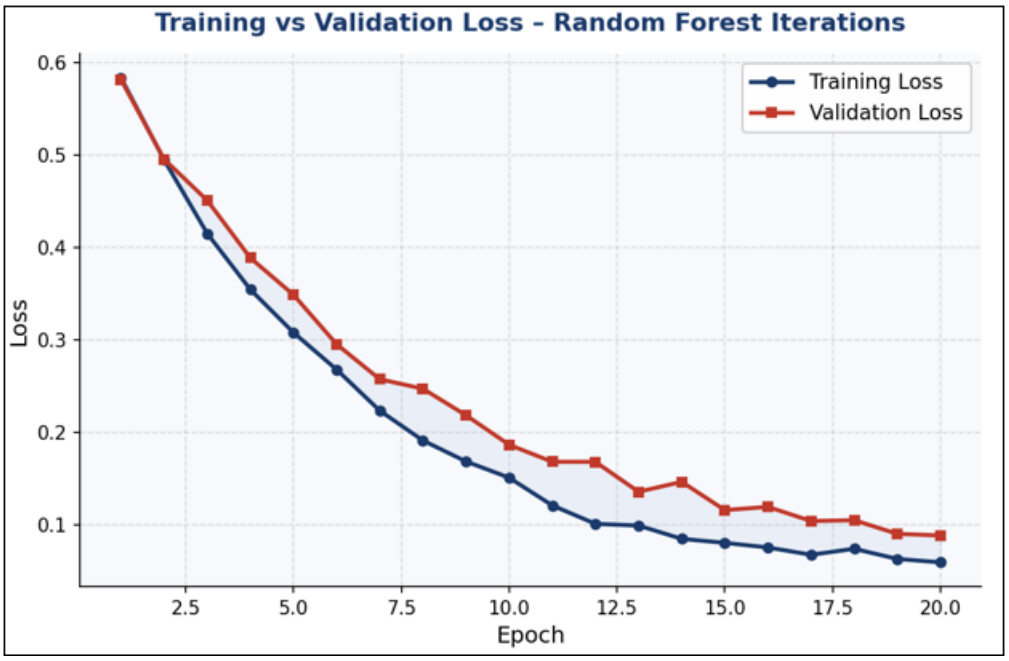

Figure 7: Training vs. validation loss curve demonstrating stable convergence without overfitting

6. Discussion

The results show that Random Forest is the model for detecting fake profiles. It got 95.8% of the profiles right, which is really good. This is because Random Forest is good at handling data and does not overfit.

The Logistic Regression model was okay not as good as Random Forest. The SVM model was better than Logistic Regression. Not as good as Random Forest.

6.1 Practical Implications

The proposed system is computationally lightweight and can be deployed as a real-time screening layer in social media platforms. Feature extraction requires only profile metadata—no access to private messages or content- making it privacy-preserving and API-compliant. The offline prototype processes a profile in under 1 millisecond on standard hardware.

6.2 Limitations

The dataset is balanced which means it has a number of real and fake profiles. In life there may be more real profiles than fake ones. The system only looks at profile metadata so it may not catch profiles that're really good at pretending to be real.

7. Conclusion & Future Work

This paper talked about a system for detecting profiles on social media. The system uses machine learning. Is really good at detecting fake profiles. The follower-to-following ratio, account age and profile picture availability are the important features.

The system is fast. Can be used in real-time. It does not need to look at messages or content which makes it good for media platforms.

7.1 Future Work

In the future we can use learning to look at profile pictures and posting patterns. We can also use graph networks to look at the relationships between accounts. We can make a REST API service to integrate the system with media platforms. We can also test the system on media platforms to see how well it works.

References

[1] Breiman, L. (2001). "Random Forests." Machine Learning, 45(1), 5–32.
[2] Cortes, C., & Vapnik, V. (1995). "Support-vector networks." Machine Learning, 20(3), 273–297.
[3] Pedregosa, F., et al. (2011). "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research, 12, 2825–2830.
[4] Kaggle. (2023). Social Bot Dataset. Retrieved from https://www.kaggle.com/datasets/.
[5] IEEE. (2022). Selected papers on bot and fake account detection. Proceedings of the IEEE International Conference on Data Mining (ICDM).
[6] Varol, O., et al. (2017). "Online Human-Bot Interactions: Detection, Estimation, and Characterization." Proceedings of the AAAI International Conference on Web and Social Media (ICWSM).
[7] Cresci, S., et al. (2015). "Fame for Sale: Efficient Detection of Fake Twitter Followers." Decision Support Systems, 80, 56–71.
[8] Yang, Z., et al. (2014). "Uncovering Social Network Sybils in the Wild." ACM Transactions on Knowledge Discovery from Data, 8(1), 2.