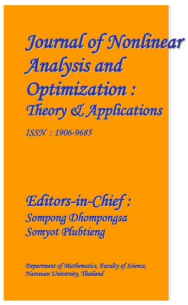
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1, No.1 : 2025
ISSN : **1906-9685**



IDENTIFYING FAKE DEALERS ON INSTAGRAM

Dr. Santhosh J, Assistant Professor, Department Of Computer Applications, Sri Krishna Adithya College Of Arts And Science, Coimbatore-641042
M. Kaviya, B. Santhosh, III BCA, Sri Krishna Adithya College Of Arts And Science, Coimbatore-641042

Abstract
The proliferation of e-commerce has led to a surge in online shopping activities, particularly on social media platforms like Instagram. However, this growth has also resulted in an increase in fraudulent dealers exploiting unsuspecting customers. This study proposes a systematic approach to identifying fake dealers on Instagram by analyzing key factors such as account activity, customer reviews, product authenticity, and interaction patterns. Using machine learning techniques, sentiment analysis, and metadata examination, the framework classifies accounts into genuine and suspicious categories. This research aims to contribute to building trust in social media marketplaces by providing practical tools and methodologies to distinguish genuine dealers from fraudulent ones. The study utilizes machine learning algorithms and natural language processing to analyze dealer profiles and identify suspicious behaviors. A framework is proposed to enhance user awareness and facilitate safe online transactions. The study utilizes machine learning algorithms and natural language processing to analyze dealer profiles and identify suspicious behaviors. A framework is proposed to enhance user awareness and facilitate safe online transactions. The study identifies key indicators of fraudulent accounts, such as unusual follower-to-engagement ratios, inconsistent or plagiarized content, excessive promotional posts, and the absence of verifiable contact details.
Keywords: Instagram, Fake dealers, social media fraud, Fraud detection, Account activity analysis, Customer reviews.

INTRODUCTION:

Instagram, as a leading social media platform, has evolved beyond its original purpose of photo sharing to become a dynamic hub for e-commerce and business activities. Its visually engaging format and vast global user base have made it a preferred platform for brands, influencers, and small businesses to market and sell products. the platform's rapid commercialization has also attracted fraudulent actors seeking to exploit its features and users for financial gain. Online fraud on Instagram typically manifests in various forms, including counterfeit product sales, fake giveaways, phishing scams, and the creation of fraudulent business profiles. These scams often target users through deceptive practices such as using fake followers and engagement to appear legitimate, plagiarizing content from authentic businesses, and offering deals that seem too good to be true.

PROBLEM STATEMENT:

The rapid growth of e-commerce on social media platforms like Instagram has provided immense opportunities for businesses and consumers. However, this growth has also given rise to a significant challenge: the proliferation of fake dealers. These fraudulent accounts deceive users by posing as

legitimate sellers, often exploiting Instagram's visual and informal nature to appear credible.

Lack of Verification Mechanisms:
Instagram lacks robust verification systems for small-scale sellers, making it difficult for users to distinguish between legitimate and fake accounts.

Manipulated Social Proof:
Fake dealers often inflate their credibility by purchasing fake followers, likes, and comments, creating an illusion of trustworthiness.

Limited Consumer Awareness:
Many users are unaware of red flags such as inconsistent pricing, unverified reviews, or lack of contact information, leaving them vulnerable to scams.
Absence of Transaction Security

LITERATURE REVIEW:

The problem of fake dealers on Instagram is not unique to this platform; social media has seen an increase in fraudulent activity, including scams, counterfeit goods, and false advertisements. Various strategies and technologies have been proposed to detect and mitigate such activities:

FAKE ACCOUNT DETECTION:

A considerable amount of research has focused on detecting fake accounts on social media platforms. Techniques like image recognition (Choi et al., 2019) and behavioral analysis (Zhang et al., 2020) have been used to detect suspicious activities. Fake accounts often show distinct patterns such as sudden spikes in followers, generic profile pictures, and minimal engagement with posts. Machine learning classifiers (e.g., SVM, Decision Trees) are commonly used to flag these accounts.

FRAUDULENT CONTENT DETECTION:

Content analysis plays a key role in identifying fake dealers, as fraudulent sellers tend to post misleading or deceptive images of products. AI techniques such as image classification (Bansal et al., 2021) and text analysis (Peng et al., 2022) have been used to detect counterfeit products and misleading advertising. Deep learning models, particularly Convolutional Neural Networks (CNNs), can be trained to distinguish between real and fake images, based on features such as image quality and common patterns in counterfeit items.

SOCIAL NETWORK ANALYSIS (SNA):

Social Network Analysis (SNA) is widely used to identify relationships between users and detect suspicious patterns of interaction. Fake dealers often operate in clusters, relying on multiple accounts to amplify their reach. By analyzing the graph structure of Instagram networks (Kumar et al., 2021), researchers can detect fraudulent activities such as coordinated efforts to deceive followers or promote fake products.

NATURAL LANGUAGE PROCESSING (NLP):

NLP is also an effective method for identifying fake dealers by analyzing the text in captions, comments, and direct messages. Fraudulent accounts often use persuasive, generic, or suspiciously repetitive language. Techniques like sentiment analysis and topic modeling (Yuan et al., 2022) can help identify irregularities in text patterns.

COLLABORATIVE FILTERING AND RECOMMENDATION SYSTEMS:

Collaborative filtering techniques have been proposed to detect fraudulent dealers by identifying abnormal behavior in user interactions with posts, such as repeated recommendations of certain accounts or products that show patterns**.**

METHODOLOGY :

This section outlines the approach used to identify fake dealers on Instagram, leveraging a combination of machine learning algorithms, image and text analysis, and social network analysis techniques. The methodology is structured into data collection, feature extraction, model training, and evaluation phases.

DATA COLLECTION:

To build a robust dataset for training and testing our models, we collected Instagram account data associated with both legitimate and suspected fake dealers. The data collection process was divided into two phases:

Phase 1: Account Identification

Legitimate Dealers: We manually selected legitimate business accounts based on verified badges,official brand names, and accounts with positive engagement (e.g., product reviews, customer feedback).

Fake Dealers: Fake accounts were identified using a combination of heuristic rules and user reports. These accounts were flagged based on suspicious behaviors, such as:

Frequent usage of generic or low-quality product images.

Poor engagement metrics (e.g., low follower interaction rates).

Phase 2: Data Extraction

For each identified account, we collected data from posts, including images, captions, comments, hashtags, follower/following count, engagement rate, and network connections (e.g., mutual followers).We also scraped publicly available metadata, such as account creation date and frequency of posts.

FEATURE EXTRACTION:

Once the data were collected, we extracted various features from both the image and textual content, as well as social network-related attributes to build comprehensive representations of each account. The feature extraction process includes the following components:

Figure 1.1 : Identify Fake Followers

IMAGE PROPERTIES:

Image Quality: Analyzing the resolution, lighting, and clarity of product images using Convolutional Neural Networks (CNNs). Images of counterfeit products typically exhibit lower

quality or certain patterns associated with low-resolution or stock images.
Facial Recognition: Identifying whether profile images are of real people or if they appear to be stock photos or generated images.

Textual Properties:
Hashtag Analysis: Identifying the most common hashtags used by suspected fake accounts and cross-referencing them with known spam hashtags or scams.
Sentiment Analysis: Analyzing sentiment within comments and captions to detect overly positive or excessively generic feedback, which may be indicative of fake promotions or deceptive marketing.
Follower/Following Ratio: Analyzing the follower-to-following ratio for patterns. Fake accounts often have disproportionately high followings but a low number of followers.
Engagement Metrics: Analyzing likes, comments, and share rates on posts. Fake dealers often exhibit unusually low engagement despite having a significant follower count.

NETWORK PATTERN:
Mapping the account's social network to detect clusters of accounts with similar behavior patterns (e.g., mutual followers, coordinated spamming).
Graph-based techniques and community detection algorithms like Louvain or Girvan-Newman were used to identify potential fake clusters.

MODEL DEVELOPMENT:
We employed a supervised machine learning approach to classify Instagram accounts as legitimate or fake. The process involved the following steps:
Data Perparation:
Normalizing numerical features such as follower count, engagement rate, and image quality.
Tokenizing and vectorizing text data from captions and comments using TF-IDF (Term Frequency-Inverse Document Frequency) and word embeddings (Word2Vec).
Resampling the dataset (using techniques like SMOTE) to address class imbalance between legitimate and fake accounts.
Model Decision:
Random Forest Classifier (RFC): We selected RFC for its ability to handle complex, high-dimensional data and its robustness to overfitting. RFC was trained on the extracted features to predict whether an account is a legitimate dealer or a fake dealer.
support Vector Machine (SVM):
SVM was also employed to compare the performance of different classifiers. It is effective for high-dimensional feature spaces and was used to test the classification accuracy.
Deep Learning Models: For image data, a pre-trained ResNet- 50 model was fine-tuned on our dataset to classify product images based on quality and authenticity.
Training and Validation:
The dataset was split into a training [70%] and testing [30%] sets for dataset.Cross-validation (k-fold) was used to ensure robustness and to avoid overfitting during model training.
Hyperparameter tuning was performed using grid search to optimize model parameters. to ensure robustness and to avoid overfitting during model training.

EVALUATION MEASURES :
The performance of the classifiers was evaluated using the following metrics:

Accuracy: Proportion of correctly identified fake and legitimate accounts.
Precision and Recall: Precision measures the accuracy of fake account predictions, while recall measures how many fake accounts were successfully identified.
F1-Score: The harmonic mean of precision and to recall offers a balanced measure of performance and combining to the strengths of both metrics.
Area Under the Receiver Operating Characteristic (ROC) Curve (AUC- ROC): To assess the model's ability to distinguish between fake and legitimate accounts.

Figure 1.2 : Protecting Your Account

ETHICAL CONSIDERATIONS

Throughout the research, ethical considerations were prioritized, particularly in relation to privacy and user data. All data used in this study were publicly available, and Instagram's terms of service were followed. No private user data were accessed, and the results of the research were anonymized to avoid any personal privacy violations**.**

Avoid Over-Automation:

While AI can be effective in identifying fake dealers, over- reliance on automated systems without human oversight can lead to errors orinjustices. It's important to have human reviewers involved in validating whether an account is truly a fake dealer, especially in borderline cases.
False Positives and Consequences: If legitimate dealers are falsely flagged as fake, it could harm their reputation and business. Ethical considerations include ensuring that the algorithm minimizes false positives (i.e., legitimate dealers being incorrectly marked as fake). Clear procedures should be in place to allow businesses to appeal or correct mistakes.

INFORMED CONSENT AND USER AWARENESS :

In the context of identifying fake dealers on Instagram, Informed Consent and User Awareness are key ethical principles to ensure that the data used for the identification process is handled in a responsible, transparent, and lawful manner. These principles focus on ensuring that users are fully aware of how their data is being used, and that they consent to it in a clear and explicit way

User Awareness: Instagram users, including both legitimate dealers and consumers, should be informed about the technologies being used to detect fake accounts and how their data is processed. While Instagram's Terms of Service often cover these points, it's important to make sure that users understand

how their data might be analyzed for fraud detection.
Consent for Data Use: If any private data (e.g., direct messages, private profiles) is to be used for detection, explicit consent must be obtained from the account owner. This might not be a concern if only publicly available
data is analyzed, but it's still an important consideration.
Informed consent refers to the process of obtaining permission from users before collecting or analyzing their data. Users should be informed in simple and understandable terms about what data will be collected (e.g., profile information, posts, comments, interactions).
It's important that users understand the scope and purpose of data collection, such as identifying fake dealers or fraudulent activities.
The consent should be specific about what types of data will be processed. For instance, if image data or text captions are analyzed, the user should be explicitly told what kind of analysis is being performed.

If a machine learning model is being used to detect fake accounts, users should be aware of how the model works, what features are analyzed, and the purpose behind the analysis.

CONCLUSION :

Ethical considerations are essential when developing and deploying systems to identify fake dealers on Instagram. By ensuring privacy, fairness, transparency, and accountability, these systems can be made more effective and respectful to all users. Striking a balance between consumer protection, business fairness, and individual privacy is key to creating responsible and reliable fraud detection mechanisms in the social media space. This methodology combines multiple approaches— machine learning, image and text analysis, and social network analysis— to identify fake dealers on Instagram. By analyzing both account and network behaviors and leveraging advanced AI techniques, this framework aims to improve the accuracy and effectiveness of detecting fraudulent accounts, providing better protection for consumers and brands.

REFERENCE:

- Zafar, H., s Khan, S. [2020]. the main Detecting Fake Accounts on Social Media Using Machine Learning. International Journal of Computer Science and Information Security 12G-136.
- Binnss, R., Veale, M., Van Kleek, M., Shadbolt, N., s Shilton, K. [2018]. 'It's Reducing My Flexibility': The Ethics of Consent in the Context of Data Sharing. Proceed to the 2018 CHI Conference on Human Factors in Computing Systems.
- Abdullah, M., s Qadir, J. [2020]. Social Media Fraud Detection: An Overview of Techniques for Fake Account Detection."International Symposium on Cyber Threats and Security.
- Ghosh, A., s Lerman, K. [202l]. An Empirical Study of Fake News Spread on Social