Indian Journal of Engineering Research Networking and Development
Volume: 2 Issue: 02 | February 2025 www.ijernd.com

Intelligent Fake Profile Detection for Instagram Using Machine Learning

Priya Avanthi Institute of Engineering & Technology

Abstract:

The project entails the development of a Python-based system for detecting fake Instagram profiles using machine learning models and image processing techniques. The script employs libraries such as TensorFlow, Keras, Scikit-Learn, Pandas, Matplotlib, Seaborn, and OpenCV, facilitating extensive data handling, statistical analysis, visualization, and machine learning model training. The process begins by importing relevant libraries and datasets, followed by preliminary data exploration and visualization to understand the features and distributions within the data. The training dataset is preprocessed and scaled before being fed into a neural network model structured with multiple dense layers and dropout for regularization, optimized using Adam optimizer for categorical cross-entropy loss. The model's performance is evaluated using metrics like accuracy and loss progression during training/validation phases. Further, the script incorporates image processing techniques to analyze profile pictures using Haar Cascade classifiers for face detection, which contributes to determining the authenticity of profiles. Additionally, OCR (Optical Character Recognition) via Easy OCR is applied to extract textual data from images, providing a comprehensive understanding of profile components such as username, posts, followers, and other descriptive elements. Overall, this project integrates multiple computational techniques to create an automated tool that aids in identifying and analyzing fake profiles on Instagram, showcasing the practical application of AI and machine learning in social media management.

Keywords- Instagram, Fake Profile Detection, Python, Machine Learning, TensorFlow, Keras, Neural Networks, Data Visualization, Image Processing, OpenCV, Haar Cascade, OCR, EasyOCR, Pandas, Matplotlib, Seaborn, Scikit-Learn, Statistical Analysis, Data Preprocessing, Model Evaluation, Text Extraction, Profile Analysis.

Introduction

This research article introduces a sophisticated framework designed to identify fake Instagram profiles using advanced machine learning techniques. The increase in Instagram's popularity has correspondingly led to a surge in fake profiles [1], which often engage in malicious activities such as spreading misinformation, conducting phishing attacks, and disrupting genuine user interactions. The authors developed a neural network model that demonstrates a high degree of precision (93%) and accuracy (91%) in detecting these profiles, significantly contributing to safer social media environments. The framework not only detects but also aims to understand the characteristics that differentiate authentic profiles from fraudulent ones, thereby providing a robust tool against online deceit.

This paper discusses the development of a user-friendly application aimed at detecting fake Instagram profiles through supervised machine learning algorithms [2]. The focus is on enhancing the security measures for companies and individuals against potential frauds by making the detection tool accessible to criminal investigation agencies. The application integrates seamlessly with existing social media frameworks to scrutinize and filter out fake profiles effectively. The approach is grounded in the reality that while social media facilitates connectivity, it also poses substantial risks like online impersonation and fraud, making such detection tools essential for secure digital interactions.

This literature review meticulously examines existing research on the use of machine learning methods for detecting fake accounts across major social media platforms. Highlighting the dual aspects of social media's role—as a tool for global connectivity and a platform for cybercrime—the review addresses the critical threats posed by fake profiles, including their use in scams [3], misinformation campaigns, and as tools in broader cybercriminal strategies. The paper evaluates various machine learning strategies that have been employed to identify and mitigate these threats, offering a comprehensive overview of the state-of-the-art in fake account detection and the ongoing research challenges.

The authors of this study focus on the application of machine learning algorithms to identify fake profiles across multiple social media platforms. They explore various algorithmic approaches that have been developed to detect anomalies and patterns typical of fake accounts [4], which are often hidden behind genuine-looking profiles. The paper discusses the effectiveness of these algorithms, the computational challenges involved, and the potential for these technologies to adapt to evolving cyber threats, making a significant contribution to ensuring the integrity and security of online social interactions.

This paper delves into various machine learning techniques to address the proliferation of fake profiles on social media [5]. It presents an analysis of how these profiles undermine the security and integrity of online networks and discusses different machine learning approaches that have been tested to detect and counteract these threats. The study highlights the critical need for robust security measures in the rapidly expanding digital landscape and assesses the effectiveness of current methods while suggesting areas for future research. In this research, the authors propose a hybrid machine learning model that integrates various techniques to detect fake profiles on social media. The model addresses the sophisticated strategies employed by operators of fake accounts to mimic genuine interactions [6], thereby avoiding detection. The paper emphasizes the adaptive nature of these threats and the necessity for equally dynamic detection systems. The hybrid model is evaluated for its effectiveness in real-world scenarios, providing a promising solution to one of the most pressing issues facing social media platforms today.

Literature Survey

Meshram et al., 2021 This paper highlights the significant increase in cybercrimes associated with fake profiles on social media platforms like Instagram. The authors emphasize the simplicity with which impostors can create fake profiles due to minimal account creation requirements, which only need an email ID. The study uses Logistic Regression and Random Forest algorithms to distinguish fake profiles from real ones, providing a methodological approach to addressing this cybersecurity threat.**Harish et al., 2023**This research focuses on distinguishing between fake and genuine Twitter profiles using a variety of machine learning techniques, including neural networks, LSTM, XG Boost, and Random Forest. It investigates the effectiveness of these models in classifying profiles based on characteristics such as follower and friend counts, status updates, and other behavioural metrics. The study presents a comprehensive approach to improve cybersecurity measures on social media platforms.**Gaikwad et al., 2024**The authors propose a novel machine learning-based methodology to identify and classify fake Instagram profiles. This approach is aimed at enhancing user privacy and security by leveraging detailed feature analysis and machine learning techniques to detect deceptive profiles effectively. The study suggests extending these methodologies to other platforms to maintain a secure and trustworthy online environment.**Sajja et al., 2024**Addressing the issue of fake interactions on Instagram, this paper employs algorithms such as Naive Bayes, logistic regression, SVMs, and neural networks. It discusses the challenges posed by automated and phony accounts, which lead to business losses and the spread of misinformation. The study highlights the use of a cost-sensitive evolutionary algorithm and the Smote Nc technique to handle data imbalance, improving the accuracy of fake account detection.**Sathvika et al., 2024**Utilizing Python and machine learning models like the Random Forest and Decision Tree Classifiers, this research identifies characteristics of fake Instagram accounts such as profile pictures, privacy settings, and activity metrics. The study underscores the role of machine learning in ensuring the integrity and security of social media platforms by effectively detecting and classifying fake profiles.**Goyal et al., 2024** This study explores sophisticated machine learning techniques, including DistilBERT for text processing and Random Forest for classification, to detect fake profiles based on user biography length and other features. It presents detailed performance metrics such as accuracy, precision, recall, and F1-Score, demonstrating the effectiveness of the proposed model in enhancing online safety and user trust.**Saranya Shree et al.**The literature survey in this paper reviews past studies that utilize machine learning and natural language processing to enhance the detection of fake profiles on social media. It discusses the importance of integrating user social engagements as auxiliary data to improve detection accuracy, reflecting on various methods and their applications in the social media context.**Meshram et al., 2021** This survey paper reviews different machine learning strategies for combating fake Instagram profiles. It assesses the effectiveness of Logistic Regression and Random Forest in detecting these profiles, considering the role of minimal account creation requirements that facilitate the proliferation of impostors on social media platforms.**Ezarfelix et al., 2022** This review focuses on combining image detection and natural language processing to identify fake accounts on Instagram, highlighting the role of machine learning in streamlining the detection process. It discusses the integration of these technologies into a cohesive system that can efficiently and effectively reduce the prevalence of fake accounts.**Bharti Goyal et al., 2024** This literature review addresses various approaches and methodologies proposed in past studies for detecting fake profiles on Instagram. It highlights the use of advanced machine learning techniques to deal with challenges such as data imbalance and the need for accurate classification models to maintain the integrity and trustworthiness of social media platforms.

Preliminaries Overview

The code involves several crucial preliminary steps that form the foundation for machine learning model development and evaluation:

1. **Importing Libraries:**
   - **Pandas**: Used for loading and manipulating datasets in a tabular format.
   - **Matplotlib** and **Seaborn**: Visualization libraries for creating various types of plots to analyze data and model performance.
   - **NumPy**: Provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays.
   - **TensorFlow and Keras**: Frameworks for building and training machine learning models, particularly deep learning models.
   - **Scikit-Learn**: Offers tools for preprocessing data, splitting datasets, and evaluating model performance using metrics like accuracy, confusion matrix, and ROC curves.
2. **Data Loading:**
   - The datasets for training and testing the model (train.csv and test.csv) are loaded into Pandas DataFrames. This step is crucial for preparing the

data for further processing and analysis.

3. **Data Exploration and Visualization:**
   - Basic data exploration includes checking the shape of the dataset, the types of data, and basic statistics.
   - Visualizations help in understanding the distribution of different features and the target variable (fake). For example, using count plots to view the distribution of categorical features like 'profile pic' and histograms for numerical features.
4. **Data Preprocessing:**
   - **Handling Missing Values**: Before training the model, it's important to handle any missing data either by filling them with appropriate values or removing the rows/columns containing them.
   - **Feature Encoding**: Non-numeric features are converted to a numeric format using encoding techniques because machine learning models generally work with numerical input.
   - **Feature Scaling**: Features are often scaled to have a similar range of values; this helps certain algorithms converge faster and perform better. Common scaling techniques include standardization and normalization.
5. **Model Building and Compilation:**
   - **Defining the Model Architecture**: Using Keras, a Sequential model is built with Dense layers that are fully connected and Dropout layers to prevent overfitting. Activation functions like ReLU (for hidden layers) and Softmax (for output layer) are specified.
   - **Compilation**: The model is compiled with an optimizer (e.g., Adam), a loss function (e.g., categorical_crossentropy for classification tasks), and

metrics (like accuracy) to evaluate the model.

6. **Model Training:**
   - The model is trained on preprocessed training data, with validation data being used to monitor the model's performance on unseen data during training. This helps in tuning the hyperparameters like the number of epochs and the batch size.

7. **Model Evaluation:**
   - After training, the model's performance is evaluated on a separate test set. Metrics such as accuracy, precision, recall, and F1-score provide insights into how well the model is performing, especially in distinguishing between the classes (fake vs. not fake profiles).

References

1. **"Fake Instagram Profile Identification and Classification using Machine Learning"**
   Authors: Not specified
   Published in: International Journal of Novel Research and Development (IJNRD)
   Summary: This research introduces a machine learning-based approach to detect and categorize fake Instagram profiles, aiming to enhance user privacy and security.
   ijnrd.org

2. **"Prediction of Fake Instagram Profiles Using Machine Learning"**
   Authors: I. Anupriya, V. Sowmiya, Dr. G. Devika
   Published in: Journal of Emerging Technologies and Innovative Research (JETIR)
   Summary: This study proposes the use of machine learning and natural language processing techniques to improve the accuracy of fake profile detection

on social networks, specifically Instagram.

jetir.org

3. **"Supervised Machine Learning Algorithms to Detect Instagram Fake Users"**
   Authors: Not specified
   Published in: IEEE Xplore
   Summary: This research aims to detect fake Instagram users based on user profiles using supervised machine learning algorithms.

ieeexplore.ieee.org

4. **"Securing Social Spaces: Machine Learning Techniques for Fake Profile Detection on Instagram"**
   Authors: Not specified
   Published in: Springer
   Summary: This paper proposes an effective machine learning model to detect fake accounts on Instagram, utilizing a dataset of real and fake accounts.

link.springer.com

5. **"Impersonation on Social Media: A Deep Neural Approach to Identify Ingenuine Content"**
   Authors: Koosha Zarei, Reza Farahbakhsh, Noel Crespi, Gareth Tyson
   Published in: arXiv
   Summary: This study focuses on identifying impersonators on social media, particularly Instagram, by proposing a deep neural network architecture that analyzes profile characteristics and user behaviors to detect fake content.

arxiv.org