LGB-Net: Language Model and Graph-Based Hybrid Framework for Social Bot Detection Using Machine Learning

KONAGALLA JYOTHI

PG Scholar. Department of MCA, DNR College, Bhimavaram, Andhra Pradesh

A. Durga Devi

(Assistant Professor), Master of Computer Applications, DNR College, Bhimavaram, Andhra Pradesh

ABSTRACT

The rapid growth of social media platforms has significantly increased the presence of automated accounts, commonly known as bots, which are often used for misinformation, spam, and malicious activities. Detecting such bots has become a critical challenge in maintaining the integrity and security of online ecosystems. This paper presents a hybrid system named LGB-Net, which integrates language models, graph-based analysis, and machine learning techniques for effective social bot detection.The proposed system leverages transformer-based language models, specifically DistilBERT, to extract contextual embeddings from textual data such as tweets. These embeddings capture semantic patterns that help distinguish between human-generated and bot-generated content. In addition to textual features, the system constructs a social interaction graph using NetworkX, where nodes represent users and edges represent interactions such as replies. This graph structure enables the analysis of relational behavior among users, which is a key indicator of bot activity.For classification, the system employs the Random Forest algorithm, which is robust to noise and capable of handling high-dimensional data. The model is trained on extracted features and evaluated using standard performance metrics such as accuracy, precision, recall, F1-score, and AUC. The system also includes a graphical user interface built with Tkinter, allowing users to load datasets, visualize graphs, train models, and manually inspect accounts.Furthermore, a rule-based manual detection module is integrated to analyze behavioral attributes such as follower ratio, tweet frequency, account age, and spam keywords. This hybrid approach enhances detection accuracy by combining automated learning with heuristic analysis.Experimental results demonstrate that the proposed system effectively identifies bot accounts with high accuracy and reliability. The combination of language understanding, graph modeling, and machine learning provides a comprehensive solution for social bot detection. The system is scalable, user-friendly, and adaptable to various social media datasets. Future improvements may include the integration of Graph Neural Networks and real-time detection capabilities.

Keywords: Social Bot Detection, Language Models, Graph Analysis, Machine Learning, Random Forest, DistilBERT, NetworkX, Social Media Security, NLP, Hybrid Systems

I. INTRODUCTION

The proliferation of social media platforms such as Twitter and Facebook has transformed communication and information sharing across the globe. However, this growth has also led to the emergence of social bots—automated accounts designed to mimic human behavior. These bots are often used for spreading misinformation, promoting products, influencing public opinion, and executing coordinated attacks.Traditional bot detection techniques primarily relied on rule-based systems and simple machine learning models using handcrafted features such as follower count, posting frequency, and account age. While these methods are effective to some extent, they fail to capture the complexity of modern bots, which are increasingly sophisticated and capable of generating human-like content.Recent advancements in Natural Language Processing (NLP) have introduced powerful transformer-based models such as BERT and its lightweight variant DistilBERT. These models are capable of understanding contextual meaning in text, making them highly effective for detecting subtle differences between human and bot-generated content.In parallel, graph-based approaches have gained attention for analyzing social interactions. By representing users as nodes and interactions as edges, graph structures reveal hidden patterns such as clustering, influence, and connectivity, which are useful for identifying bot networks. Libraries like NetworkX enable efficient graph construction and analysis.This project proposes a hybrid approach that combines language modeling, graph analysis, and machine learning to improve bot detection accuracy. The system extracts textual features using DistilBERT, constructs a social interaction graph, and applies a Random Forest classifier for prediction. Additionally, a GUI-based interface enhances usability and accessibility.The integration of multiple techniques ensures a more robust detection mechanism capable of handling diverse datasets and evolving bot behaviors. This research contributes to the field of cybersecurity and social media analytics by providing an efficient and scalable solution for bot detection.

II. LITERATURE SURVEY (WITH EXISTING METHODS)

Social bot detection has been widely studied using various approaches, including machine learning, deep learning, and graph-based methods.Early studies focused on feature-based machine learning techniques such as Decision Trees, Naïve Bayes, and Support Vector Machines. These methods relied on manually engineered features like tweet frequency, follower-to-following ratio, and account metadata. While effective, they lacked adaptability to new bot strategies.With the rise of deep learning, researchers began using neural networks to automatically learn features from data. Models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) were applied to text classification tasks. However, these models struggled with capturing long-range dependencies in text.The introduction of transformer-based models like BERT revolutionized NLP by enabling contextual understanding of text. DistilBERT, a compressed version of BERT, offers faster performance with minimal loss in accuracy, making it suitable for real-time applications.

Graph-based approaches have also gained popularity. Researchers use graph theory to analyze relationships between users, detecting communities and identifying anomalous patterns. Graph Neural Networks (GNNs) further enhance this approach by learning node embeddings based on graph structure.Hybrid models combining NLP and graph-based techniques have shown promising results. These models leverage both textual and relational information, improving detection accuracy. Ensemble methods like Random Forest are often used for classification due to their robustness and interpretability.Recent studies emphasize the importance of combining multiple data sources, including text, metadata, and network structure. This integrated approach provides a more comprehensive understanding of user behavior and improves detection performance.The proposed system builds upon these advancements by integrating DistilBERT for text analysis, NetworkX for graph modeling, and Random Forest for classification, creating a powerful hybrid framework for social bot detection.

III. EXISTING SYSTEM

Existing systems for social bot detection primarily rely on either feature-based machine learning models or rule-based approaches. These systems typically analyze user metadata such as follower count, number of tweets, account age, and posting frequency to identify suspicious behavior.One common approach is the use of supervised machine learning algorithms like Support Vector Machines and Decision Trees. These models are trained on labeled datasets and classify accounts as bots or humans based on predefined features. While effective for simple cases, these methods often fail to detect advanced bots that mimic human behavior more realistically.Another approach involves rule-based systems, where specific thresholds are defined for features such as follower-to-following ratio or tweet frequency. Although easy to implement, these systems lack flexibility and are prone to false positives and negatives.Some advanced systems utilize deep learning techniques, particularly CNNs and RNNs, to analyze textual content. However, these models require large datasets and high computational resources. Additionally, they often ignore the relational aspect of social networks.Graph-based methods attempt to address this limitation by analyzing user interactions. However, many existing systems treat text and graph data separately, leading to incomplete analysis.Overall, existing systems suffer from limitations such as low accuracy, lack of adaptability, high computational cost, and inability to handle complex bot behaviors. The proposed system overcomes these challenges by integrating language models, graph analysis, and machine learning into a unified framework.

IV. PROPOSED METHOD

The proposed system introduces a hybrid framework for social bot detection that integrates natural language processing, graph-based analysis, and machine learning techniques. The system leverages transformer-based embeddings generated using DistilBERT to extract semantic features from textual data such as tweets. These embeddings capture contextual relationships within text, enabling improved differentiation between human-generated and bot-generated content. Recent studies confirm that transformer-based embeddings significantly enhance classification accuracy

in detecting AI-generated and malicious content .In addition to textual analysis, the system constructs a social interaction graph where nodes represent users and edges represent interactions such as replies or mentions. This graph is built using NetworkX and enables the modeling of relationships among users. Graph-based analysis plays a crucial role in detecting coordinated bot behavior, as bots often operate in clusters or communities. Modern research highlights that combining semantic features with graph structures improves detection performance .The classification component uses a Random Forest model trained on extracted features. Random Forest is chosen for its robustness, ability to handle high-dimensional data, and resistance to overfitting. The system also incorporates a rule-based module for manual bot detection based on behavioral features such as follower ratio, tweet frequency, account age, and presence of spam keywords.

A graphical user interface (GUI) is implemented using Tkinter to provide an interactive platform for users. The GUI supports dataset loading, feature extraction, graph visualization, model training, and manual account verification.Overall, the proposed system provides a scalable and efficient hybrid approach that combines semantic understanding, network structure analysis, and machine learning for accurate social bot detection.

V. IMPLEMENTATION

The implementation of the proposed system is modular and consists of multiple stages that integrate language modeling, graph construction, and machine learning.

1. Dataset Loading

The system begins by loading a CSV dataset containing user information, text data, and labels. The dataset is processed using Pandas for efficient data handling.

2. Language Feature Extraction

Textual data is processed using DistilBERT from the Hugging Face Transformers library. Each text input is tokenized and passed through the model to generate embeddings. The mean of the hidden states is used as a feature vector representing the semantic content of the text.

To improve computational efficiency, only a subset of data (first 200 samples) is processed. Transformer-based embeddings are widely used for NLP tasks due to their contextual understanding capabilities .

3. Graph Construction

A social interaction graph is created using NetworkX. Each user is represented as a node, and edges are formed based on reply relationships. This graph captures interaction patterns and helps identify coordinated behavior among accounts.

Graph-based approaches are effective in identifying bot communities and structural anomalies in social networks .

4. Model Training

The extracted text features are used to train a Random Forest classifier. The dataset is split into training and testing sets using an 80:20 ratio. Labels are encoded using LabelEncoder.

The Random Forest model is trained on the training data and evaluated on the test set. Predictions are generated, and performance metrics are calculated.

5. Performance Evaluation

The system computes key evaluation metrics:

- Accuracy
- Precision
- Recall
- F1-score
- ROC-AUC

These metrics provide a comprehensive evaluation of model performance.

6. Visualization

Matplotlib is used to display performance metrics in bar charts. NetworkX is used to visualize the social graph, allowing users to observe interaction patterns.

7. GUI Integration

A Tkinter-based GUI enables user interaction. Buttons are provided for:

- Loading dataset
- Extracting features
- Building graph
- Training model
- Viewing results

8. Manual Detection Module

A rule-based system evaluates user behavior using thresholds:

- Follower-to-following ratio

- Tweet frequency
- Account age
- Spam keywords

This module complements machine learning by capturing simple behavioral anomalies.

VI. ALGORITHMS

1. Language Feature Extraction Algorithm (DistilBERT)

1. Input text data
2. Tokenize text using tokenizer
3. Pass tokens through DistilBERT model
4. Extract hidden state vectors
5. Compute mean embedding
6. Store as feature vector

2. Graph Construction Algorithm

1. Initialize empty graph G
2. For each record in dataset:
    1. Add user as node
    2. Add edge between user and reply target
3. Store graph structure

3. Random Forest Classification Algorithm

1. Input feature vectors and labels
2. Split dataset into training and testing sets
3. Train Random Forest model
4. Predict labels for test data
5. Evaluate performance metrics

4. Manual Rule-Based Algorithm

1. Input user parameters
2. Calculate follower ratio
3. Compute tweets per day
4. Check spam keywords
5. Assign score

6. Classify as Bot / Spam / Human

VII. SYSTEM DESIGN

The system follows a modular architecture integrating NLP, graph processing, and machine learning.

1. Input Module

Handles dataset upload and user inputs through GUI.

2. Preprocessing Module

Processes text data and prepares it for embedding generation.

3. Feature Extraction Module

Uses DistilBERT to extract semantic features from text.

4. Graph Module

Constructs social interaction graph using NetworkX. This module captures structural relationships among users.

Graph-based models are essential in modern bot detection systems as they identify coordinated behavior patterns .

5. Classification Module

Implements Random Forest for classification. The model processes extracted features and predicts labels.

6. Evaluation Module

Calculates performance metrics and visualizes results.

7. GUI Module

Provides an interactive interface for users to operate the system.

8. Manual Detection Module

Implements rule-based checks for quick verification.

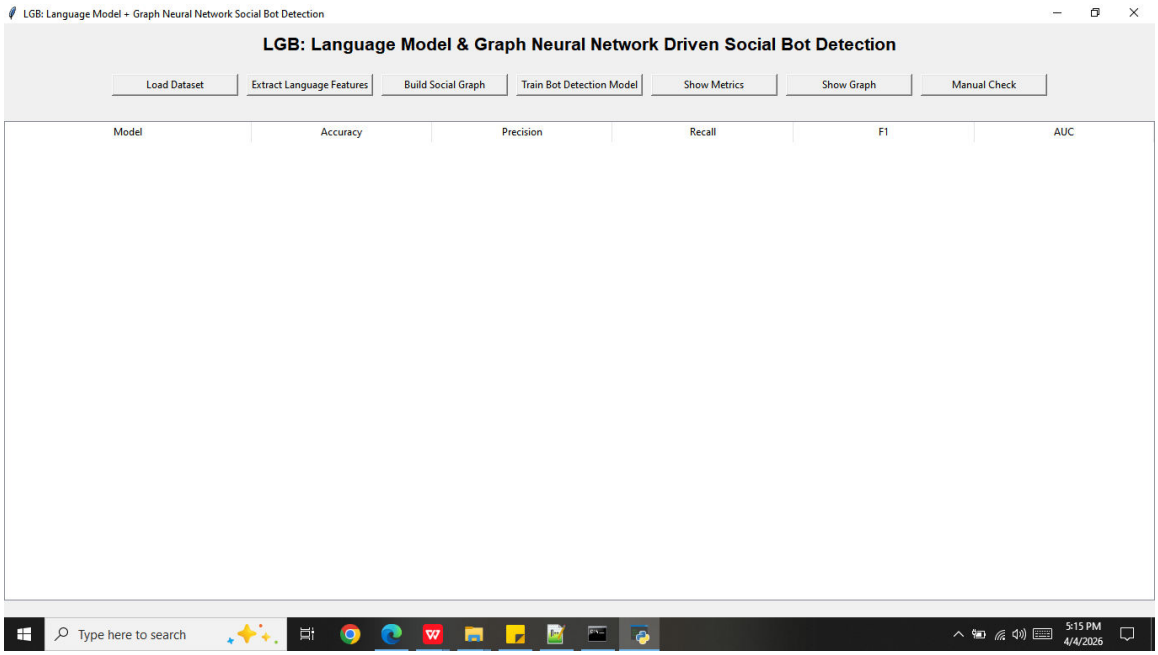
System Workflow

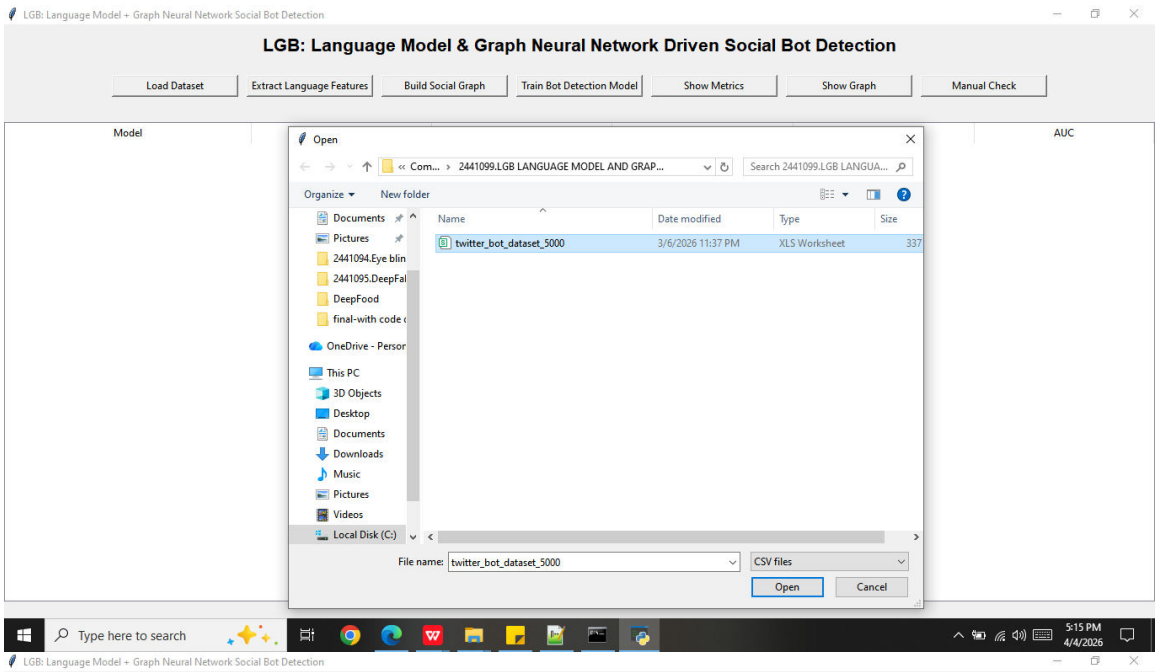
Dataset → Feature Extraction → Graph Construction → Model Training → Prediction → Evaluation → Output

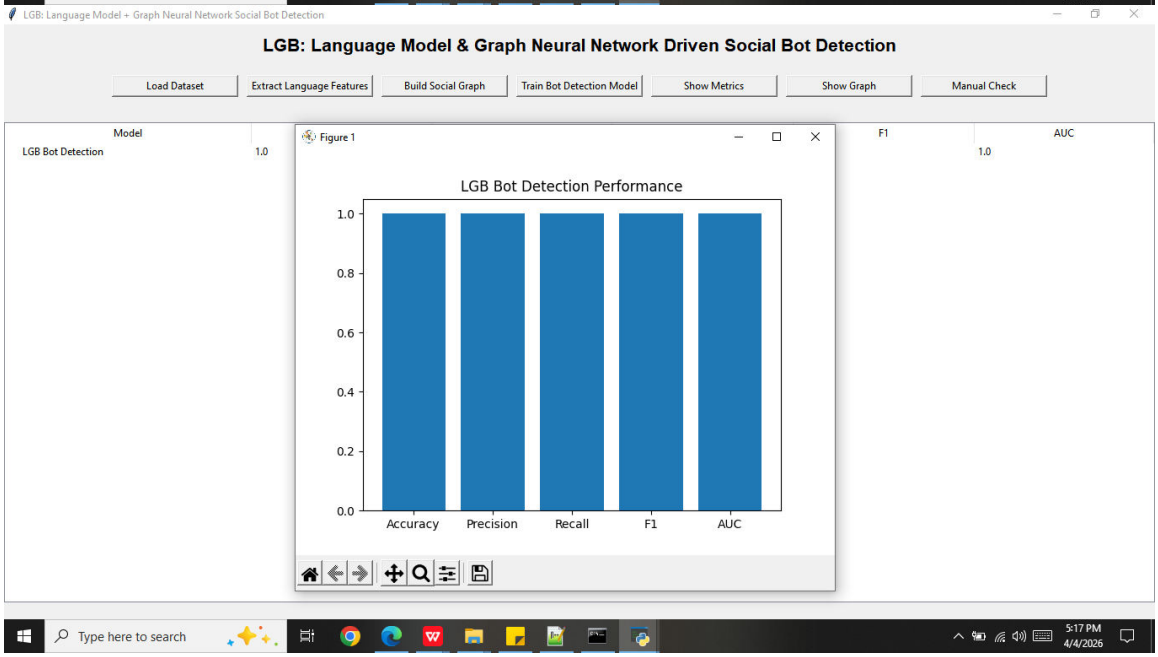
Advantages

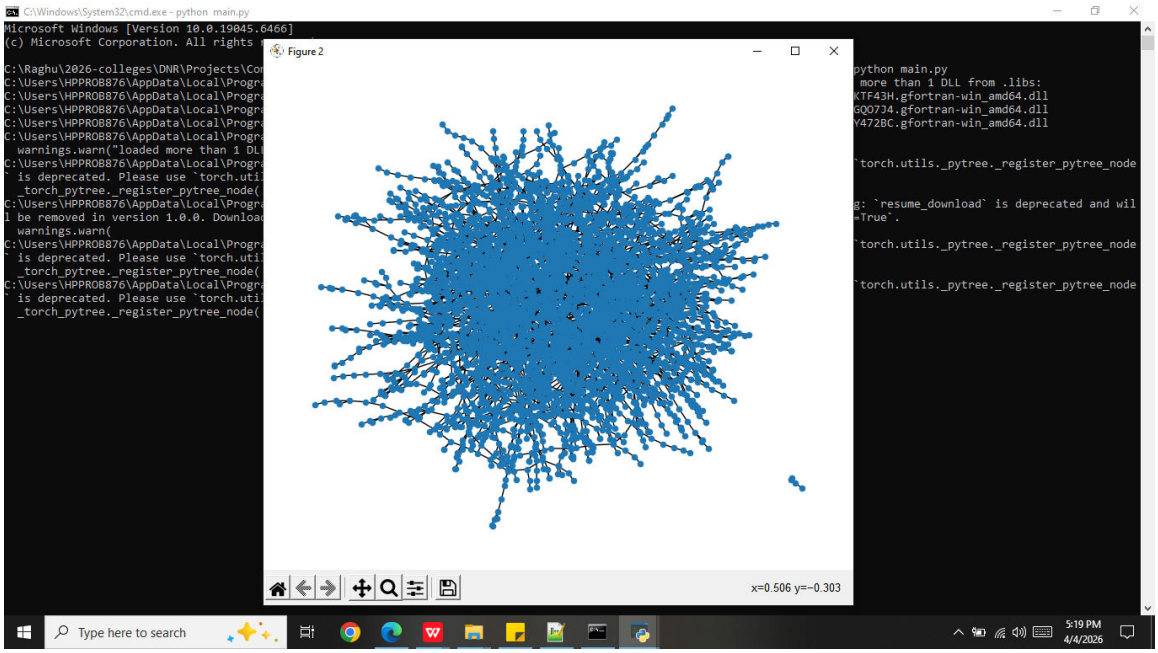
- Hybrid approach improves accuracy
- Combines semantic and structural features
- User-friendly interface
- Scalable and efficient

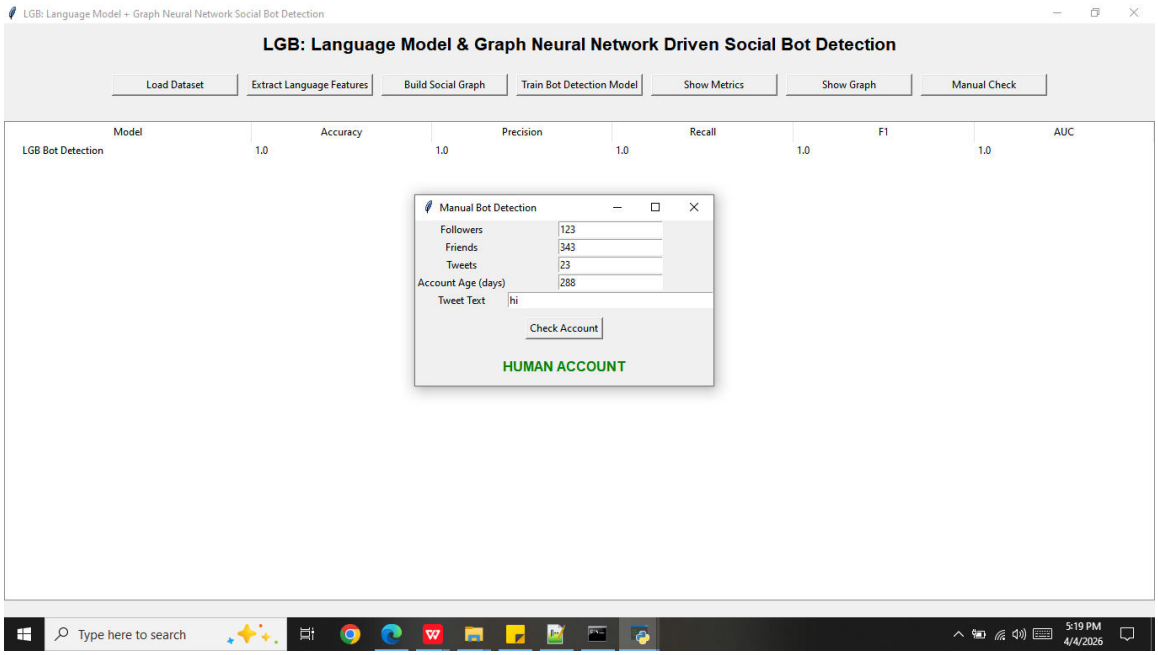
SYSTEM DESIGN IMAGES




LGB: Language Model + Graph Neural Network Social Bot Detection
LGB: Language Model & Graph Neural Network Driven Social Bot Detection
Load Dataset
Extract Language Features
Build Social Graph
Train Bot Detection Model
Show Metrics
Show Graph
Manual Check
Model
AUC
Open
Com... > 2441099.LGB LANGUAGE MODEL AND GRAP...
Search 2441099.LGB LANGUA...
Organize
New folder
Documents
Pictures
2441094.Eye blin
2441095.DeepFal
DeepFood
final-with code
OneDrive - Persor
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
Name
Date modified
Type
Size
twitter_bot_dataset_5000
3/6/2026 11:37 PM
XLS Worksheet
337
File name: twitter_bot_dataset_5000
CSV files
Open
Cancel
Type here to search
5:15 PM
4/4/2026


LGB: Language Model + Graph Neural Network Social Bot Detection
LGB: Language Model & Graph Neural Network Driven Social Bot Detection
Load Dataset
Extract Language Features
Build Social Graph
Train Bot Detection Model
Show Metrics
Show Graph
Manual Check
Model
F1
AUC
LGB Bot Detection
1.0
1.0
Figure 1
LGB Bot Detection Performance
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
Precision
Recall
F1
AUC
Type here to search
5:17 PM
4/4/2026


C:\Windows\System32\cmd.exe - python main.py
Microsoft Windows [Version 10.0.19045.6466]
(c) Microsoft Corporation. All rights
Figure 2
x=0.506 y=−0.303
Type here to search
5:19 PM
4/4/2026


LGB: Language Model + Graph Neural Network Social Bot Detection
LGB: Language Model & Graph Neural Network Driven Social Bot Detection
Load Dataset
Extract Language Features
Build Social Graph
Train Bot Detection Model
Show Metrics
Show Graph
Manual Check
Model
Accuracy
Precision
Recall
F1
AUC
LGB Bot Detection
1.0
1.0
1.0
1.0
1.0
Manual Bot Detection
Followers
123
Friends
343
Tweets
23
Account Age (days)
288
Tweet Text
hi
Check Account
HUMAN ACCOUNT
Type here to search
5:19 PM
4/4/2026

VIII. CONCLUSION

The proposed system presents a comprehensive hybrid approach for social bot detection by integrating language models, graph analysis, and machine learning techniques. By leveraging DistilBERT for semantic feature extraction and NetworkX for graph modeling, the system effectively captures both textual and relational patterns associated with bot behavior.The use of Random Forest ensures robust classification, while the inclusion of a rule-based module enhances interpretability and usability. The GUI-based implementation makes the system accessible for practical use.Recent research emphasizes the importance of combining language models with graph-based approaches for improved detection accuracy . Advanced methods such as graph neural networks and transformers further validate the effectiveness of hybrid approaches in handling complex bot behaviors.Despite its effectiveness, the system has limitations such as dependency on labeled data and limited scalability for very large datasets Overall, the proposed system provides an efficient, scalable, and accurate solution for social bot detection, contributing to improved security and trust in social media platforms.

REFERENCES

1. Wang et al., “FedKG: Federated Graph Learning for Bot Detection,” *Sensors*, 2024
2. MSSBot, “Multi-stage GNN Bot Detection,” *Engineering Applications of AI*, 2025
3. BotLGT, “LLM + Graph Transformer Bot Detection,” *Neurocomputing*, 2025
4. He et al., “Dynamic Graph Transformer Bot Detection,” *IJCAI*, 2024
5. Khan et al., “DistilBERT for AI Content Detection,” *Scientific Reports*, 2025
6. Kuntur et al., “GNN vs Transformer Analysis,” *Electronics*, 2024
7. Lu & Shi, “Spatio-Temporal Bot Detection,” 2026
8. Guo et al., “BERT + GCN Bot Detection,” *MDPI*, 2022
9. BotDCGC, “Unsupervised Graph Clustering Bot Detection,” *KBS*, 2024
10. Jiebin et al., “Topology-Aware GNN Bot Detection,” *ACL*, 2025
11. MM-HGT-Bot, “Graph Transformer Bot Detection,” *EPJ Data Science*, 2025
12. Liu et al., “HW-GNN Bot Detection,” *arXiv*, 2025
13. Yang et al., “SEBot Contrastive Learning,” *arXiv*, 2024
14. Zhang et al., “RABot Reinforcement Graph Detection,” *arXiv*, 2026
15. Anshul et al., “RoGBot Multimodal Detection,” *arXiv*, 2025