COMPARATIVE ANALYSIS ON DEEP LEARNING METHODS FOR BOT DETECTION

Dr.R.Lalu Naik[1], Likhitha Danda[2],

Sri Lakshmi Boddupalli[3] , HariKrishna Durru[4], Gelli Teja Reddy[4]

[1]*Professor, Department of Computer Science and Engineering, Tirumala Engineering College*

[2,3,4,5]*Student, Department of Computer Science and Engineering, Tirumala Engineering College*

---***---

Abstract - Social bots are automated accounts on social media that are managed by software and overseen by people behind the scenes. While some bots are created for positive reasons, like sharing news updates or offering assistance in crises, others have been misused to spread false information, rumors, or influence political events. There are tools in place to identify and eliminate harmful bots automatically, but creators are continually evolving their techniques to evade detection. Hence, there is a pressing need for improved ways to differentiate between genuine and automated bot accounts. In recent years, several research studies have delved into the realm of social media bot detection, offering an extensive overview of different detection strategies, including advanced approaches such as machine learning (ML) and deep learning (DL) techniques. To the best of our knowledge, this study stands out as the first to solely focus on DL techniques, evaluating their efficacy and rationale in comparison to each other and to conventional ML methods. In this research, we offer a comprehensive classification of the characteristics utilized in deep learning research and provide information on the necessary preprocessing methods for creating appropriate training data for a deep learning model. We highlight the gaps identified in review papers discussing deep learning and machine learning studies, offering insights into future developments in this area. In general, deep learning methods have proven to be efficient in terms of computation and time for detecting social bots, demonstrating superior or comparable performance compared to traditional machine learning techniques.

***Key Words*:** Bot Detection, Social Media, Deep Learning, Computation, Automation

1.INTRODUCTION

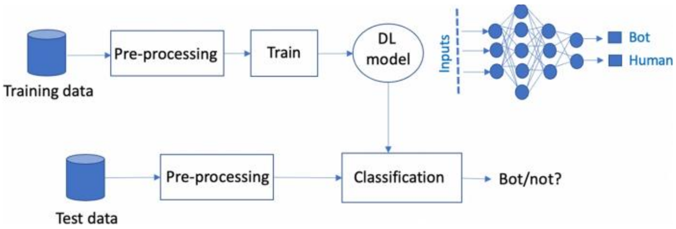
In today's world, social media platforms are growing rapidly in terms of users, data volume, and features. These platforms are essentially online applications that make it easy for users to share their own content. The widespread use of these platforms is transforming the way people communicate. People typically use social media to connect with others through posts, following, and being followed. Platforms like Twitter often feature trending topics that spark daily conversations. Malicious individuals and organizations exploit the flexibility and power of social media to gain influence by creating fake automated accounts, often called social bots or Sybil accounts, and, in this study, social media bots. These accounts can exploit the regular services for malicious purposes by manipulating the discussion and public opinion, spreading rumors and fake news, promoting harmful products and services, defaming other people, or being fake followers of a user to handcraft a fake popularity and spamming, social phishing, profile cloning, and collusion attacks. These attacks can have devastating consequences. Some of the most harmful examples of bot infiltration include interference in the US presidential election, the Russiagate hoax attack, and the spread of rumors during the Boston Marathon bombings on Twitter.

Therefore, there is significant interest in researching social bot detection as a defense against threats. The human-like qualities of these bots and their ever-changing strategies make it challenging to differentiate them from real users. This calls for more advanced countermeasures. Using machine learning to analyze bot behavior patterns is a strategic way to efficiently detect them, utilizing the large amounts of data produced by online platforms.

There have been various machine learning methods suggested for identifying bots on social media platforms like Twitter. These include supervised, unsupervised, and reinforcement learning techniques. While these models generally perform well in detecting social bots and are easy to implement, they can be time-consuming and computationally expensive due to feature extraction. Additionally, they may have slower learning times and reduced effectiveness when dealing with a large number of features. Deep learning is a unique type of machine learning that stands out from traditional approaches because of its layered structure and capability to analyze and identify features from intricate data like images, text, and speech. Several deep learning models have proven to be superior to traditional, shallow machine learning classifiers for tasks like bot detection.

The reason for conducting this systematic review focusing solely on deep learning research for social media bot detection stems from two key factors. Firstly, deep learning approaches have demonstrated significant promise in identifying both benign and malicious bots and adapting quickly to their constantly changing features. This capability is crucial in the current landscape. Secondly, previous systematic reviews have not delved into the effectiveness, limitations, and obstacles specifically concerning these techniques in social media bot detection.

Furthermore, to tackle the problem of cyborgs displaying human-like behavior, deep learning methods, especially generative adversarial networks, can prove to be very efficient.



2. LITERATURE SURVEY

A recent study on the use of IOT devices in the United States from 2018 to 2019 found that these devices are susceptible to attacks in 57% of cases, with severity levels ranging from moderate to severe. Additionally, 41% of attackers exploit these vulnerabilities. Furthermore, a social media bot suggests that 71% of Twitter accounts talking about U.S. stocks are actually social bots, although only 37% of them have been recognized by Twitter. Another study in 2018 revealed the activities of spam bots in Twitter microblogs discussing stock market trends. In a study, it was found that 71% of Twitter accounts that shared stock updates were actually bots, and 37% of these accounts were eventually banned. In 2019, 11% of Facebook accounts were also identified as bots. Garcia, Zunio, and Campo conducted a survey in 2014 on methods for detecting botnets in network communication and security systems. Various types of attacks, such as phishing and password theft, are becoming more common. Bots are targeting individual computers by stealing user IPs and monitoring system traffic in order to launch attacks.

3. SYSTEM ANALYSIS

- **Existing System**

  In the current system, numerous researchers are exploring various methods for identifying bots, with some utilizing advanced deep learning techniques. A study introduced a model combining CNN and LSTM that achieved an accuracy of approximately 87%, which is regarded as effective in detecting attacks on IoT devices. However, the model has limitations as it is not suitable for a wide range of datasets and cannot determine the best algorithm. Additionally, it may not always deliver highly precise results.

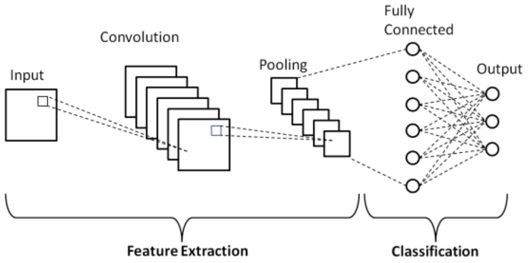
- **Proposed System**

  In the current system, numerous researchers are exploring various methods for identifying bots, with some utilizing advanced deep learning techniques. A study introduced a model combining CNN and LSTM that achieved an accuracy of approximately 87%, which is regarded as effective in detecting attacks on IoT devices. However, the model has limitations as it is not suitable for a wide range of datasets and cannot determine the best algorithm. Additionally, it may not always deliver highly precise results.

4. DEEP LEARNING TECHNIQUES FOR BOT DETECTION

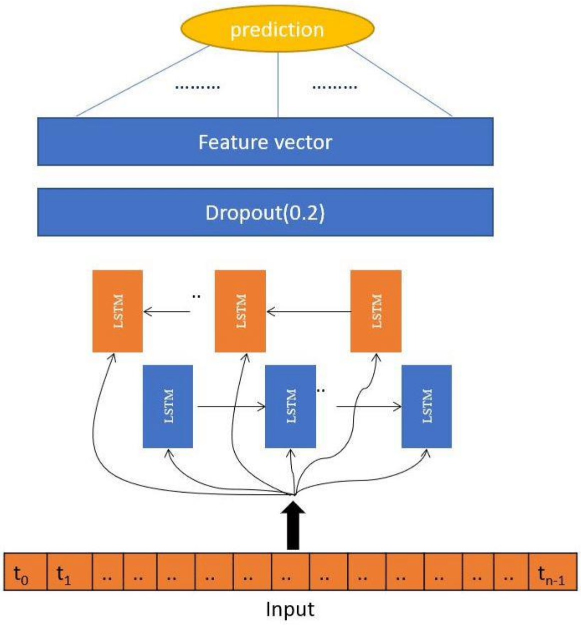
- **Convolutional Neural Networks**

  Convolutional Neural Networks have shown promise in detecting bots by analyzing structured and unstructured data like user behavior logs, text content, and network traffic. With convolutional layers, CNNs can extract intricate features from input data, effectively identifying spatial patterns that signal bot activity. CNNs are widely applied in various bot detection tasks, such as sentiment analysis of user comments, categorizing social media posts, and examining network traffic patterns. They excel in scenarios where the spatial relationships between features are critical for distinguishing between content created by bots and content created by humans.

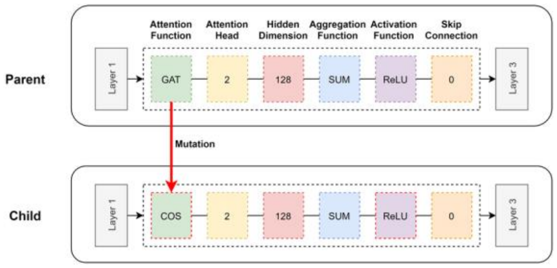


- **Recurrent Neural Networks**

  Recurrent Neural Networks are great for understanding sequences of data and relationships over time, which is useful for identifying bots by analyzing how users interact and behave over time. Some versions like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) help with retaining information over long periods and solving issues like the vanishing gradient problem, which is important for spotting subtle patterns in bot behavior. Uses: RNNs can be used for tasks like analyzing user sessions, finding unusual behavior in sequences of actions, and recognizing bot-generated content by looking at patterns in language. They are especially skilled in situations where it is important to accurately identify the sequential patterns in bot behaviors.



- **Graph Neural Networks**
  Graph Neural Networks excel at capturing intricate relationships and interdependencies within graph-structured data, making them particularly useful for identifying bots in social networks and online communities. By pooling information from nearby nodes in the graph, GNNs can create representations that encompass both local details and overarching structural patterns, facilitating the detection of coordinated bot behaviors and irregularities on a community-wide scale. Potential use cases for GNNs include community identification, predicting connections between nodes, and spotting anomalies within social network graphs.



- **Transformer models**
  Transformer models such as BERT and GPT are popular in natural language processing because they can understand relationships between words in text. They are often used to detect bots by looking at text and user interactions. This model is used for tasks such as classifying text, analyzing thoughts, and creating dialogue. These uses are important for identifying bots in online chats and user-generated content.

5. EVALUATION METRICS

Key metrics often used when comparing detection tools are:

- **Accuracy:**
  Accuracy tells us how many times each event was correctly classified and gives a rough idea of how well the model is performing.
- **Precision:**
  Measure the accuracy of identifying real bots based on bot examples.
- **Sensitivity:**
  This demonstrates the model's ability to accurately detect robots.
- **F1-Score:**
  Average precision and recall that fairly evaluates the model's performance, taking into account false positives and biases.
- **Area Under the ROC Curve (AUC-ROC):**
  It evaluates the model's ability to distinguish robotic and non-robotic situations based on different parameters. A higher value indicates that the model has better separation.

6. FUTURE ENHANCEMENTS

When it comes to the future potential, we have only completed the initial phase of the project in line with the ultimate objective, which involves evaluating existing models in the field to identify any shortcomings. This can be expanded upon by introducing a tailored model to yield more precise outcomes. Additionally, we have only utilized the first version of the datasets so far, but exploring version 2 may lead to more accurate results due to the additional features it offers. As for the model itself, we have currently categorized the data into only two groups: whether it is a bot or not, specifically benign or malicious. There are also numerous sub-categories within these types of attacks. We attempted to implement sub-class identification but were unsuccessful due to the limited features and computational power needed for building a suitable model for such classification. This is an area where improvements can be made or further development can take place.

7. CONCLUSION

In our study, we found that each type of deep learning method has its own strengths and weaknesses when it comes to detecting bots. CNNs are great at capturing patterns and features in data, making them ideal for analyzing all types of data. RNNs are experts at understanding relationships over time, especially when it comes to user behavior sequences. GNNs are useful for uncovering coordinated bot actions on social networks by looking at network structures and node characteristics. Transformer models are top performers at identifying patterns and context in text data, resulting in precise identification of bot-generated content.

Our examination also brings to light various obstacles and factors to keep in mind when using deep learning techniques for detecting bots. These hurdles encompass concerns regarding the adequacy of the dataset, imbalance in data, the ability to understand the model, available computing resources, and the capacity for growth. Overcoming these obstacles necessitates thoughtful deliberation on data manipulation methods, model structures, fine-tuning of parameters, and approaches to assessment. The online platform landscape is constantly changing, so it's important to keep researching and innovating in deep learning methods for detecting bots. By tackling the challenges we've identified and building on the lessons learned from our analysis, we can create better bot detection systems that protect the integrity and security of online communities.

8. ACKNOWLEDGEMENT

We are extremely thankful to Dr. R. Lalu Naik, our guide, for his unwavering support and guidance during the project. His expertise and motivation were instrumental in our success, and we greatly appreciate his dedication and mentorship. We also extend our gratitude to the Computer Science and Engineering Department faculty at Tirumala Engineering College for allowing us to participate in this research endeavor, which has been a valuable learning opportunity.

9. REFERENCES

1. "Research about cyber attacks." [Online]. Available: https://unit42. paloaltonetworks.com/iot-threat-report-2020/
2. "Different types of attacks." [Online]. Available: https://www.csoonline.com/ article/3295877/what-is-malware-viruses-worms-trojans-and-beyond.html
3. Kadhim Hayawi, Susmita Saha, Mohammad Mehedy Masud, Sujith Samuel Mathew & Mohammed Kaosar Social media bot detection with deep learning methods: a systematic review Published: 06 March 2023 Volume 35, pages 8903–8918, (2023)
4. H. Alkahtani and T. H. Aldhyani, "Botnet attack detection by using cnn-lstmmodelfor internet of things applications," Security and Communication Networks, vol. 2021, 2021.
5. https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909
6. https://www.researchgate.net/figure/The-architecture-of-the-recurrent-neural-network-RNN-model-In-addition-we-have_fig4_346263746
7. https://www.researchgate.net/figure/Graph-neural-network-architecture-evolution-example-The-GNN-architecture-can-be-encoded_fig1_346144034
8. A. Derhab, A. Aldweesh, A. Z. Emam, and F. A. Khan, "Intrusion detection system for internet of things based on temporal convolution neural network and efficient feature engineering," Wireless Communications and Mobile Computing, vol. 2020,2020.
9. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. doi:10.1038/nature14539
10. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735
11. Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive Representation Learning on Large Graphs. Advances in Neural Information Processing Systems, 30, 1024–1034. Retrieved from https://papers.nips.cc/paper/6703-inductive-representation-learning-on-large-graphs.pdf
12. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805 [cs]. Retrieved from https://arxiv.org/abs/1810.04805
13. Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global Vectors for Word Representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543. doi:10.3115/v1/d14-1162
14. Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep Learning (Adaptive Computation and Machine Learning series). MIT Press.
15. "Details about bot and botnets." [Online]. Available: https://www.netscout.com/ what-is/bot
16. "Bot net." [Online]. Available: https://datadome.co/learning-center/ how-to-detect-mitigate-botnets/