IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10834

AN ADAPTIVE SOCIAL SPAMMER DETECTION MODEL WITH SEMI-SUPERVISED BROAD LEARNING

Apoorva.R[1], Hemanth Kumar B.N[2]

PG Scholar, Dept. of MCA, Vidya Vikas Institute of Engineering and Technology, Mysuru, Karnataka, India[1]

Associate professor, Dept. of MCA, Vidya Vikas Institute of Engineering and Technology, Mysuru, Karnataka, India[2]

Abstract-This approach is highly effective in creating predictions, especially when compared to certain traditional forms of supervised learning. When using ASSD, identifying individuals within a group becomes simpler and requires less effort. To update the spammer detection model without requiring retraining of the model's users, incremental learning is used as a technique because social scammers often modify their behavior to deceive the spammer detection model. When benchmarked against other controlled and semi-supervised machine learning algorithms, the Social Honeypot Dataset is used to compare ASSD's performance. The study's findings suggest that the proposed model outperforms baseline approaches in terms of memory capacity and accuracy. Additionally, ASSD maintains its high accuracy in identifying spammers by continuously updating its model with newly collected data from social media.

I. INTROUDCTION

Mobile social networks are becoming increasingly popular among social media users who collaborate to share messages with each other. However, these networks are often targeted by con artists who post links to malicious software and advertisements or follow a large number of people, flooding the network with deceptive messages. In light of this, we propose a model called ASSD, which can identify adaptive social spammers. To create a spam predictor, we use a limited set of labeled patterns along with unknown patterns.

This approach is highly effective in creating predictions, especially when compared to classic forms of supervised learning. Using ASSD, it is much simpler to find out who is in a group and requires less effort. Social scammers often change their behavior to evade spam detection, so we use "incremental learning" to update the spammer detection model without requiring retraining of users. The Social Honeypot Dataset is used as a benchmark to compare ASSD with other controlled and semi-supervised machine learning algorithms. The study's findings indicate that the proposed model performs better than baseline approaches in terms of memory capacity and accuracy. Furthermore, ASSD maintains its high accuracy by continuously upgrading its model with newly collected data from social media.

Problem Statement

Social media platforms have become a significant part of our daily lives, providing a platform for communication, information sharing, and networking. However, they are also susceptible to the presence of spammers who engage in various malicious activities such as disseminating fake information, promoting scams, and disrupting genuine conversations. Detecting and mitigating social spammers is crucial to maintaining a healthy and trustworthy online environment.

Aim

The aim of this project is to develop an adaptive social spammer detection model that leverages semi-supervised broad learning techniques. Traditional spam detection methods often rely solely on labeled data, which may not be sufficient to identify evolving and adaptive spamming behaviors.

II. LITERATURE SURVEY

1. "A Survey of Machine Learning Techniques for Spam Detection" Authors: S. Aljawarneh, et al

This survey paper provides an extensive overview of machine learning techniques employed for spam detection across various domains, with a particular focus on their application to social media platforms. It discusses the challenges and trends in spam detection and highlights the significance of machine learning in combating spam.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10834

2. "Semi-Supervised Learning" Author: Olivier Chapelle, et al
This seminal paper introduces the concept of semi-supervised learning, presenting a comprehensive understanding of the principles and methodologies involved. It covers key techniques and explores the benefits of leveraging unlabeled data alongside labeled data for machine learning tasks.

3. "Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach" Authors: S. Kumar and A. Jaiswal
This research paper addresses the detection of spam bots in online social networks through the application of machine learning techniques. It presents a detailed investigation into the characteristics and behaviors of spam bots, proposing machine learning-based solutions for their identification and mitigation.

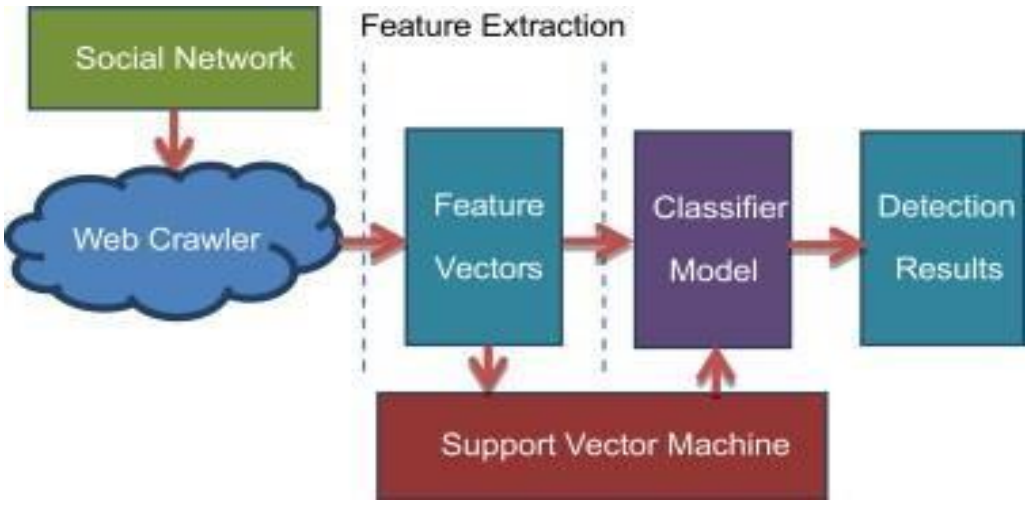
4. "Deep Learning for Hate Speech Detection in Tweets" Authors: ZeerakWaseem, et al.
Focusing on the identification of hate speech in Twitter data, this paper explores the use of deep learning methodologies. It discusses the development and evaluation of deep learning models tailored for detecting hate speech in tweets, contributing to the broader field of content moderation on social media.

5. "Adaptive Machine Learning for Detecting Online Fake News" Authors: Diego F. Silva, et al.
This paper investigates the problem of detecting online fake news using adaptive machine learning approaches. It discusses the challenges posed by the spread of misinformation and presents techniques for continually adapting machine learning models to identify fake news articles on the internet.

III. BACKGROUND STUDY

With the increasing popularity of social media, it has become crucial to detect and prevent spam from spreading across these platforms. There are currently three primary approaches to identifying spam: network topology-based, user behavior-based, and content analysis-based.

Topology-based methods leverage the structural characteristics of social networks, such as node degree, community structure, and clustering coefficient, to differentiate between legitimate users and spammers. Machine learning algorithms are often used to address this problem, as it can be framed as a binary classification task.



IV. METHODOLOGY

We have proposed a semi-supervised broad learning method for detecting spammers. The method builds a spammer detection model with high accuracy, using a small number of labeled social patterns and a large number of unlabeled user information. To learn the changing distributions of social features adaptively, we have designed an incremental learning method. This allows the spammer detection model to be updated dynamically without retraining.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10834

We have evaluated the proposed model on real Twitter datasets, and its accuracy has been found to be higher than the baseline algorithms. Moreover, our model maintains a high detection accuracy by adaptively updating itself with newly generated social media data. Compared to conventional supervised learning methods, our model has a high prediction accuracy.

By applying ASSD, the time and energy required to label the identity of social members are reduced. Additionally, the incremental learning method designed to update the spammer detection model adaptively without retraining is useful because social spammers frequently change their behavior to deceive the spammer detection model.

To construct additional labeled training data, the most confident predictions of each classifier are applied. First, the labeled dataset is used to train a predictive model. Next, the unlabelled data is classified by the training model. After evaluating the classification results, a small amount of unlabelled data with a high confidence level and predicted labels are selected as new input data.

Data Set

The dataset is regarding the Instagram fake accounts and comments. The dataset includes the data like the data like the username, profile picture, posts, number of followers, number of followings and more. And the dataset also includes the data regarding the comments on which they are categorized into fake or real.

V. CONCLUSION

In conclusion, the study presented an adaptive social spammer detection model utilizing semi-supervised Broad learning. The model demonstrated promising results in identifying socialspammers by combining deep and shallow learning techniques and leveraging both labeled and Unlabeled data.

FUTURE ENEHANCEMENT

Although the proposed model already combinesdeep and shallow learning techniques, further exploration of advanced deep learningarchitectures, such as recurrent neural networks (RNNs), long short-term memory (LSTM)networks, or transformers, can be considered.

REFERENCES

[1]. C. Cao and J. Caverlee's "Detecting spam URLs in social media via behavioral analysis" was presented at the European Conference on Information Retrieval in 2015.
[2]. F. Wu, J. Shu, Y. Huang, and Z. Yuan's "Co-detecting social spammers and spam messages in microblogging via exploiting social contexts" was published in Neurocomputing in 2016.
[3]. S. Ghosh, B. Viswanath, F. Kooti, N. K. Sharma, G. Korlam, F. Benevenuto, et al.'s "Understanding and combating link farming in the Twitter social network" was presented at the 21st International Conference on World Wide Web in 2012.
[4]. H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao's "Detecting and characterizing social spam campaigns" was presented at the 10th ACM SIGCOMM conference on Internet measurement in 2010.
[5]. E. Tan, L. Guo, S. Chen, X. Zhang, and Y. Zhao's "Unik: Unsupervised social network spam detection" was presented at the 22nd ACM International Conference on Information & Knowledge Management in 2013.
[6]. T. Hong, C. Choi, and J. Shin's "CNN-based malicious user detection in social networks" was published in Concurrency and Computation: Practice and Experience in 2018.
[7]. G. Stringhini, C. Kruegel, and G. Vigna's "Detecting spammers on social networks" was presented at the 26th Annual Computer Security Applications Conference in 2010.
[8]. Y. Zhang, T. Wang, M. Qiao, A. Zhu, C. Li, and H. Snoussi's "Detection of abnormal events in complex situations using strong classifier based on BP adaboost" was presented at the International Conference on Intelligent Computing in 2016.
[9]. C. Li, S. Wang, L. He, S. Y. Philip, Y. Liang, and Z. Li's "SSDMV: Semi-supervised deep social spammer detection by multi-view data fusion" was presented at the 2018 IEEE International Conference on Data Mining (ICDM).
[10]. J. Xue, Z. Yang, X. Yang, X. Wang, L. Chen, and Y. Dai's "VoteTrust: Leveraging friend invitation graph to defend against social network sybils" was presented at the 2013 Proceedings IEEE INFOCOM.