Library Progress International
Vol.44 No.3, Jul-Dec 2024: P.3184-3190

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article

Available online at www.bpasjournals.com

A Machine Learning Approach to Recognising Propaganda on Social Networks

[1]Ramswaroop Reddy Yellu, [2]Sarath Babu Dodda, [3]Brajesh Kumar Sharma, [4]Sohong Dhar [5]Sheetal Temara,

[1]Independent Research,
University of Texas, USA, Email: raparthimohan.scholar@gmail.com
[2]Software Engineer, Central Michigan University, USA, Email:sbabudodda009@gmail.com, ORCID: 0009-0008-2960-2378
[3]Assistant ProfessorComputer Science and Engineering Sir Padmapat Singhania University, Udaipur, Rajasthan, India, brajesh.india@gmail.com
[4] Junior Research Assistant,Library & Information Science, Jadavpur University, West Bengal, India, Email: 25111987sd@gmail.com
[5]Doctoral Scholar, Computer and Information Sciences, University of the Cumberlands, USA, sheetaltemara@gmail.com

How to cite this article: Ramswaroop Reddy Yellu, Sarath Babu Dodda, Brajesh Kumar Sharma, Sohong Dhar Sheetal Temara, (2024). A Machine Learning Approach to Recognising Propaganda on Social Networks. *Library Progress International*, 44(3), 3184-3190.

Abstract
In order to identify and combat misinformation on social networking sites, the present study analyses the use of automated learning algorithms. With the appearance of the internet and its extensive use, social media structures have turn out to be massive venues for statistics dissemination and influence. But, these platforms have additionally been exploited for spreading propaganda and false data, concentrated on people, corporations, and political entities. Our study specializes in analysing diverse machine learning classifiers, which include Support Vector Machines (SVM), Random forest, and deep learning strategies like BERT and Roberta, to differentiate among propaganda and non-propaganda content. We make use of datasets collected from online information resources and Twitter, leveraging the Twitter API for information extraction. The paper presents a complete evaluate of cutting-edge methodologies, highlights the demanding situations faced in propaganda detection, and discusses the effectiveness of different machine learning models based on their accuracy and applicability in real-world eventualities. Our findings indicate that machine learning, especially advanced models like RoBERTa, can appreciably useful resource in figuring out and mitigating the spread of propaganda and extremist content on social networks.

Keywords—*Social Media, False Information, Propaganda, Machine Learning, Detection System.*

Introduction

The internet, one of the most significant innovations, is used by a large variety of people for numerous goals. Through online structures each person can access a variety of social media networks to publish content or share information. Certain persons attempt to spread misleading data because these networks do not validate their identities or the content they provide. These reviews, which can be a disinformation focused on a certain individual, group, organization or political party, are too severe for a human to understand. Machine Learning classifiers with the ability to automatically come across such fake information items are required [1]. Social networks have become a powerful tool for data dispersal, with people sharing information without knowing its authenticity. They are used in various fields, such as elections and advertisements, to control sentiments and spread disinformation as depicted in Fig. 1. An organised and conscious attempt to persuade

people in search of benefits in either religion or politics is known as disinformation or propaganda. Information is collected from online information sources and Twitter through the Twitter API. System learning classifiers categorize textual content into lessons (Propaganda and Non-Propaganda), with Support Vector Machine (SVM) demonstrating advanced performance among conventional algorithms. [2]. Propaganda and extremism that is violent pose a serious threat on a national and international level. Terrorist groups and their adherents have never had it easier to propagate this fanaticism because to the growing popularity and usage of social media. The overwhelming majority of social media interactions by one terrorist group ISIS (the Islamic State of Iraq and Syria) take place on Twitter. To counter the influence and online presence of these extremists on social media sites like Twitter, cyber analytics techniques must be developed [3].

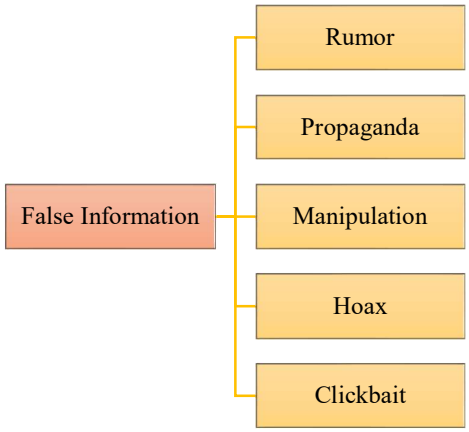

Fig. 1 Types of false information on social media

Given its growing popularity, bots—which have the ability to manipulate user behaviour, disseminate false information, and carry out propaganda—have found great success on Twitter. Machine learning methods were used in [4], such as Random Forest and XGBoost, to estimate the probability that an account will be classified as a bot. Exploratory data analysis, model building, requirement collecting, feature selection, optimisation, hyperparameter tinkering, and algorithm benchmarking were all part of the research. With an accuracy of 0.8908 for XGBoost and 0.8762 for Random Forest, the results demonstrated that XGBoost performs better than Random Forest in the identification of bots. Using Twitter tweets as a source for an extremism dataset, [5] attempts to identify extremism on social media platforms. Texts that promote extremism are categorised by the dataset as non-extremist, radicalization, recruitment, and propaganda. Previous studies don't have a specific dataset and only offer a cursory analysis of texts that promote extremism. Methods based on artificial intelligence such as BERT, DistilBERT, RoBERTa, and Bi-LSTM are used to evaluate the dataset. With a 95% accuracy rate, RoBERTa was determined to be the most appropriate for identifying extremism on social media.

II RELATED WORK

Millions of people are exposed to famous memes on social media, which helps shape collective ideas. While some memes develop naturally, others are maintained by campaigns for information, advertising, or concerted efforts. Certain communication campaigns could be used for evil things like financial market manipulation or terrorist propaganda. Early on, a machine learning framework has been developed in [6] to identify promoted ads and distinguish them from organic ones. The researchers obtained a 95% AUC score in early detection using a dataset of millions of postings linked to popular Twitter hashtags, revealing the distribution patterns inside networks and the alerts provided by using content serve as strong signs for early detection. It's critical to identify when a meme is being promoted maliciously in order to prevent bad intentions. In order to categorise trending memes on Twitter, a machine learning framework was created in [7]. Time-varying characteristics were utilised to record content, sentiment, timing signals, network and diffusion patterns, and user meta-data. For early detection, the framework's AUC score was 75%; after trending, it rose to above 95%. Content cues offer helpful signals, but user traits are more informative for early identification, according to feature selection analysis. Spammers are unwanted users who publish unsolicited or irrelevant content. They are a big problem. A study conducted in [8] suggests methods for identifying spammers on Twitter by utilising learning techniques such as Naive Bayes, Clustering, and Decision Trees. By combining these algorithms in a novel way, the accuracy of spam detection is

increased. According to the study, this system detects non-spammers with 99% accuracy and has an overall accuracy of 87.9%, outperforming classical methods overall. Discussions concerning the possible risks of digital political manipulation have been triggered by the impact of machine learning techniques on public decision-making, notably throughout the Brexit referendum and the presidential elections in Brazil and the USA. Voters' thoughts and decisions can be greatly influenced by disinformation and propaganda thanks to the profiling and targeting tactics employed by social media platforms for advertising and propaganda. An investigation in [9] suggests an approach primarily based on obligation for social media platforms across different political contexts. as a countermeasure. According to the study, social media platforms may reduce risks to society by upholding human rights, promoting education, and being transparent about their algorithmic decision-making. The accuracy rate of suggested strategy is 84.57%, which is very high.

In order to better understand the processes of radicalization through social media communication channels, [10] focuses on the social media behaviour of 25,000 people whose affiliation with ISIS's online radical propaganda has been personally validated. The study examines the dynamics of social influence among ISIS sympathisers using dynamic activity-connectivity maps. The results draw attention to the similarities between the propagation of radical propaganda and epidemics, showing that influential ISIS followers and information broadcasters create extremely contagious information cascades. The results will aid in the creation of potent defences against ISIS and other online extremist groups. Research on identifying phoney identities made by humans is lacking as a result of the increasing number of persons masking their identities on social media platforms (SMPs) for nefarious motives. However, utilising computer-generated attributes like the "friend-to-followers ratio," machine learning models have been able to identify phoney accounts that are manufactured by bots or computers. In order to improve the detection of fictitious identities made by humans on SMPs, [11] attempts to apply these attributes to phoney human accounts. The goal of the project is to enhance SMP fake identity detection. Fake news poses a serious threat to societies and makes it difficult to fight misinformation. Particularly during the COVID-19 epidemic, it has an impact on residents' lives, individual reputations, and democratic elections. In order to counter misinformation, [12] provides a thorough examination of the state of the art in intelligent tool use. In addition to offering expert-based solutions and highlighting the path of intelligent systems in misinformation detection, it also emphasises the historical background and contemporary importance of fake news in the information war. The report also offers a summary of important R&D initiatives pertaining to the identification of false news and helpful resources. Finding difficulties and methodological gaps for further studies were provided in Table 1.

Feature/Study	**Objective**	**Methodology**	**Key Findings**	**AUC Score/ Accuracy**	**Additional Notes**
[6] Machine Learning Framework for Ad Detection	Identify promoted ads vs. organic ones on social media	millions of tweets with hashtags posted in one dataset	Effective early detection of promoted ads, highlighting the role of network diffusion patterns and content cues	95% AUC score in early detection	-
[7] Machine Learning Framework for Trending Memes	Categorize trending memes on Twitter	Time-varying characteristics including content, sentiment, timing signals, network and diffusion patterns,	User traits found more informative for early detection; accuracy improves significantly after trending	75% AUC score early detection, >95% after trending	Content cues useful but user traits are more informative for early detection

		user meta-data			
[8] Spam Detection on Twitter	Identify spammers	Combined Naive Bayes, Clustering, and Decision Trees algorithms	High accuracy in detecting non-spammers and overall spam detection, outperforming classical methods	99% accuracy for non-spammers, 87.9% overall accuracy	Novel combination of algorithms increases detection accuracy
[9] Countermeasures to Digital Political Manipulation	Address risks of machine learning in political contexts	Social Media Strategy: Responsibility-Based	Suggests promoting education, transparency in algorithmic decision-making, and upholding human rights	84.57% accuracy rate of the suggested strategy	Focuses on reducing societal risks through platform responsibility
[10] Radicalization Dynamics on Social Media	Understand radicalization through ISIS's social media activity	Analyzed social media behaviour of 25,000 users associated with ISIS using dynamic activity-connectivity maps	Highlights parallels between radical propaganda and epidemics, identifying highly infectious information cascades	-	Aids in creating defenses against online extremism
[11] Fake Identity Detection on SMPs	Improve detection of fake human accounts on social media	Applying Designed Features to Construct False Human Accounts	Improving Phoney Identities Identification on Social Media	-	Addresses the challenge of detecting human-created fake accounts
[12] Fighting Misinformation with Machine Learning	Counter misinformation and fake news	Comprehensive review of intelligent tools, expert-based solutions, and R&D projects related to fake news detection	Emphasizes the role of intelligent systems in detecting misinformation sources	-	Aims to identify challenges and gaps for future research

III CONCEPTUALIZING PROPAGANDA IN THE DIGITAL AGE

By giving news objects on social media a social context to be defined and described, context-based fake news detection makes it easier to identify false news using data analysis and pattern recognition. But while machine learning and natural language processing (NLP) techniques have made content-based fake news identification more popular, the context-based approach has received very little exploitation. As a result, it is now important to thoroughly investigate and incorporate alternative solutions for contextualized social media identification of fake news [13]. The study conducted in [14] examines the use of various machine learning algorithms to classify dark data, which can be valid or bogus, and propaganda, which is widely used for decision-making. The study uses Logistic Regression, Multinomial Naive Bayes, Support Vector Machine, and Decision Tree algorithms. The results show that Support Vector Machine (SVM) has the highest accuracy at 69.84%, while Logistic Regression has a higher accuracy at 68.76%. The study also found that propagandistic sentences have longer lengths. Nowadays, propaganda serves many purposes than merely swaying public opinion. Propaganda is now a means to instantly participate in political disputes from the comfort and safety of your living room chair in our digitally mediated environment. Ironically, it can also be used to quickly sever ties between friends and family members whose viewpoints diverge. Participatory propaganda facilitates the socialisation of conflict and helps to normalise it [15]. This expanding scope of engagement may also result in an internalisation of conflict. This means that participatory propaganda seeks to alter your cognitive filters and the way you interact with your surroundings, rather than promoting the use of alternative information sources. The internet and social media have become a platform for organized and misinformation campaigns, often maligning businesses, government policies, individuals, and political parties. This has led to the abuse of the democratic right of free speech, discussion, and debate. Fake news requires three essential ingredients: tools and services for manipulation, a relevant social media network, and motivation for action as depicted in Fig. 2. Online resources are accessible for both free and charged use, providing resources and tools for fabricating and disseminating false information.

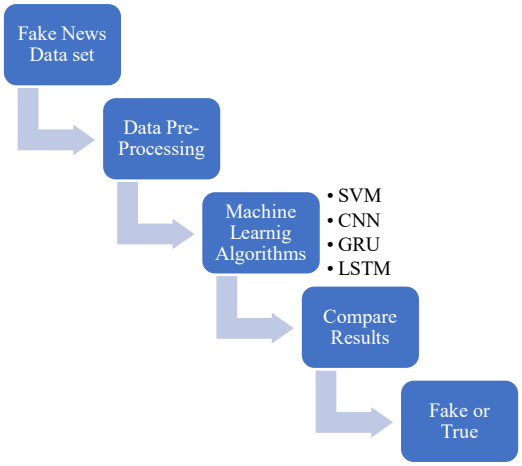

Fig. 2 Process of fake news detection system

The spectrum of tools consists of paid likes and followers, carrying out online polls, and even pressuring website owners to get rid of testimonies. Social media platforms have to remain operational and engaged to disseminate propaganda, with users investing significant time on these platforms to remain informed. Manipulation efforts that use misleading information as their main driving force are always motivated by the issue of "why." Whether it is for monetary gain through advertising and marketing or for criminal or political reasons, the ultimate effect of any fake news hinges on its real-world outcomes. The Russian invasion of Ukraine in February 2022 led to information warfare practices, with pro-Russian messages reaching around 14.4 million users on social media. A study in [17] found. The study discovered that pro-Russian messages garnered 251,000 retweets, extensively amplifying their preliminary spread. Bots were significantly influential in disseminating these messages, specially in nations that abstained from vote casting on UN resolution ES-11/1. In total, 20.28% of the individuals spreading the messages were identified as bots, many of which have been set up on the outset of the invasion. These findings indicate the presence of a significant Russian propaganda campaign on social media. and suggest curbing bots as an effective strategy. The purpose of [18] was to identify features of visual propaganda in the era of social media and online social networking. Images from the Gaza conflict in November 2012 that were posted to Twitter by the Israel Defence Forces and Hamas' Alqassam Brigades were examined. To find recurring themes and

frames in all 243 Twitter photos that the two sides shared over the course of two months, content analysis was done. Themes of resistance and togetherness dominated the photos that Israel uploaded, while resistance and civilian casualties dominated the photos that Hamas shared. While the emotional propaganda frame predominated in Hamas photos, the analytical propaganda frame was present in the bulk of Israeli photographs. In discussing ISIS's media tactics, [19] emphasises the group's active use of social media, global audience targeting, and reliance on religious themes for legitimacy. In order to support counterterrorism activities, it recommends that Western governments take into account contested internet platforms, create social media specialised teams, and enact comprehensive legislation. The report also issues a warning that an intentional media campaign may undermine ISIS's theological underpinnings and religious legitimacy, undermining the organization's reputation as a global defender of Muslims. In reaction to cyberattacks and internet misinformation, the Ukrainian government restricted access to Russian websites, including VKontakte. Even though the majority of users could technically and legally get around the prohibition, the abrupt filtering policy decreased activity on VKontakte. Users who had significant social and political ties to Russia were equally likely to be impacted [20]. The primary factor influencing users' reaction to the restriction was the availability of access to online media, indicating that this realistic assessment of censorship's impact is applicable even in the highly politicised Russia-Ukraine conflic

IV. CONCLUSION

In conclusion, our research into the machine learning method for recognizing propaganda on social networks has revealed that those strategies can substantially decorate our potential to detect and mitigate the spread of harmful content material. By using integrating diverse machine learning models, which include SVM, Random woodland, and advanced deep getting to know methods like BERT and RoBERTa, we've proven their capability in distinguishing among true and propagandistic content material. Our research also underscores the significance of continuous information series and model training to conform to the evolving nature of propaganda strategies. Although challenges which includes information bias, moral concerns, and the dynamic nature of social media content persist, the advancements in system mastering provide a promising avenue for preventing propaganda. We propose for a collaborative approach, related to social media structures, policymakers, and the studies community, to refine those detection mechanisms and enforce them efficiently. The fight against propaganda on social networks is ongoing, and machine learning offers a valuable device in this essential endeavor.

REFERENCES

[1] Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., & Choi, M. S. “Detecting fake news using machine learning: A systematic literature review”. arXiv preprint arXiv:2102.04458, 2021.

[2] M. Ud Din Khanday, Q. Rayees Khan and S. T. Rabani, "Analysing and Predicting Propaganda on Social Media using Machine Learning Techniques," *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, Greater Noida, India, 2020, pp. 122-127, doi: 10.1109/ICACCCN51052.2020.9362838.

[3] J. Thee, I. Alsmadi and S. Al-khateeb, "Pro-ISIS Tweets Analysis Using Machine Learning Techniques," 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 4351-4358, doi: 10.1109/BigData50022.2020.9378356.

[4] F. N. Pramitha, R. B. Hadiprakoso, N. Qomariasih and Girinoto, "Twitter Bot Account Detection Using Supervised Machine Learning," 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2021, pp. 379-383, doi: 10.1109/ISRITI54043.2021.9702789.

[5] Rajendran et al., "Detecting Extremism on Twitter During U.S. Capitol Riot Using Deep Learning Techniques," in IEEE Access, vol. 10, pp. 133052-133077, 2022, doi: 10.1109/ACCESS.2022.3227962.

[6] Ferrara, E., Varol, O., Menczer, F., &Flammini, A. (2016). Detection of promoted social media campaigns. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 10, No. 1, pp. 563-566).

[7] Varol, O., Ferrara, E., Menczer, F., &Flammini, A. (2017). Early detection of promoted campaigns on social media. EPJ data science, 6, 1-19

[8] Gupta, A., &Kaushal, R. (2015, March). Improving spam detection in online social networks. In 2015 International conference on cognitive computing and information processing (CCIP) (pp. 1-6). IEEE.

[9] Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. European journal of operational research, 291(3), 906-917.

[10] Ferrara, E. (2017). Contagion dynamics of extremist propaganda in social networks. Information Sciences, 418, 1-12.

[11] Van Der Walt, E., &Eloff, J. (2018). Using machine learning to detect fake identities: bots vs humans. IEEE access, 6, 6540-6549.

[12] Choraś, M., Demestichas, K., Giełczyk, A., Herrero, Á.,Ksieniewicz, P., Remoundou, K., ... &Woźniak, M. (2021). Advanced Machine Learning techniques for fake news (online disinformation) detection: A systematic mapping study. Applied Soft Computing, vol. 101, pp. 107050, 2020.

[13] Bani-Hani, O. Adedugbe, E. Benkhelifa, M. Majdalawieh and F. Al-Obeidat, "A Semantic Model for Context-Based Fake News Detection on Social Media," 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA), Antalya, Turkey, pp. 1-7, .020, doi: 10.1109/AICCSA50499.2020.9316504.

[14] M. Ud Din Khanday, Q. Rayees Khan and S. T. Rabani, "Analysing and Predicting Propaganda on Social Media using Machine Learning Techniques," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, pp. 122-127, 2020. doi: 10.1109/ICACCCN51052.2020.9362838.

[15] Asmolov, G. "The effects of participatory propaganda: From socialization to internalization of conflicts".2019.

[16] Goswami, ManashPratim. "Fake News and Cyber Propaganda: A study of manipulation and abuses on Social Media." Mediascape in 21st Century: Emerging Perspectives pp. 535-544, 2018.

[17] Geissler, D., Bär, D., Pröllochs, N., &Feuerriegel, S. "Russian propaganda on social media during the 2022 invasion of Ukraine". EPJ Data Science, vol. 12, no. 1, pp. 35, 2023.

[18] Seo, H. "Visual propaganda in the age of social media: An empirical analysis of Twitter images during the 2012 Israeli–Hamas conflict". Visual Communication Quarterly, vol. 21, no. 3, pp. 150-161, 2014.

[19] Yeung, J. C. "A Critical analysis on ISIS propaganda and social media strategies". Terrorism and Security Studies, 2015.

[20] Boyte, K. J. "An analysis of the social-media technology, tactics, and narratives used to control perception in the propaganda war over Ukraine". Journal of Information Warfare, vol. 16, no. 1, pp. 88-111, 2017.