International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0172
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1055 - 1059

Social Media and Misleading Information in Democracy Using Machine Learning

Namilakonda Sumanjali[1], Anjali Jangili[2], Md Muzamil Ahmad[3], Mullye Chaitanya[4], Etikyala Sai Prasad[5]
[1]Assistant Professor, Dept. of Data science, CMR Engineering College, Medchal, 501401, Telangana, India.
[2,3,4,5]UG Scholar, Dept. of Data science, CMR Engineering College, Medchal, 501401, Telangana, India.
Emails ID: *n.sumanjali23@gmail.com[1], anjalijangili2003@gmail.com[2], muzamilmujju07@gmail.com[3], chaitanyachotu62@gmail.com[4], saiprasadetikyala445@gmail.com[5]*

Abstract

In this project, we address the challenge of mitigating the spread of misleading information in a democracy by designing a resource allocation mechanism tailored for strategic social media platforms. Recognizing the pivotal role of both a strategic government and the private knowledge of how misinformation impacts users, we propose a mechanism that incentivizes social media platforms to efficiently filter misleading content. The mechanism leverages an economically inspired approach to strongly implement all generalized Nash equilibrium, ensuring efficient filtering in the induced game. We demonstrate that the proposed mechanism is individually rational, budget-balanced, and guarantees the existence of at least one equilibrium. Furthermore, under quasi-concave utility functions and constraints, the mechanism admits a generalized Nash equilibrium and achieves Pareto efficiency. To validate our approach, we utilize machine learning algorithms—including Logistic Regression, Decision Tree Classifier, and Random Forest Classifier—to model and analyze the strategic behavior of social media platforms and the effectiveness of misinformation filtering.

Keywords: *Social Media Platforms, Mechanism design, Resource allocation, Machine learning algorithms, and Misinformation filtering.*

1. Introduction

Introduction In a democracy, the spread of misinformation can disrupt public discourse, manipulate opinions, and influence electoral outcomes. With the increasing influence of social media, the challenge of filtering misleading information has become more critical. Governments and social media platforms face difficulties in controlling the dissemination of deceptive content while maintaining freedom of speech. This research proposes a mechanism design approach to address these challenges. By considering the strategic behavior of social media platforms and leveraging machine learning algorithms, the proposed system aims to filter misleading information effectively, promoting responsible information sharing. The mechanism accounts for economic incentives, ensuring compliance while balancing ethical considerations. Machine learning models like Logistic Regression, Decision Trees, and Random Forests are applied to categorize content, predict misinformation, and evaluate the effectiveness of filtering strategies. This integration of mechanism design and data-driven analysis aims to create a more secure and informed digital environment while reducing the impact of strategic manipulation. Our approach is a proactive solution that aims to preserve democratic integrity while respecting individual freedoms. Moreover, this mechanism design considers the varying levels of misinformation severity, adapting filtering intensity based on the potential harm of the deceptive content. The system incorporates real-time data analysis to assess the evolving nature of misinformation tactics, allowing continuous updates to filtering strategies. Additionally, the approach explores the ethical implications of filtering in democratic societies, balancing censorship concerns with the need to maintain truthful information dissemination. By addressing these multifaceted aspects, our methodology aspires to create a sustainable and reliable approach to combating misinformation while enhancing accountability among social media platforms [1-5].

2. Literature Survey

A common machine learning (ML)-based misinformation detection system includes a data acquisition mechanism that gathers social media data from multiple sources. After undergoing pre-processing, which includes feature extraction, tokenisation, and text cleaning, the acquired data is split into training and validation sets. Classification models such as Random Forests, Decision Trees, and Logistic Regression are trained with optimised hyperparameters using the training dataset, which is typically larger. To guarantee efficient misinformation filtering, the validation dataset evaluates model performance using metrics like accuracy, precision, recall, and F1-score.

2.1 The Role of Social Media in Spreading Misinformation

Social media has emerged as a major information source, making it a potent but frequently unreliable medium for the spread of news. Because of its novelty and emotional appeal, researchers like Vosoughi, Roy, and Aral (2018) showed that false information spreads more quickly than true information. Their analysis of Twitter data, which was published in Science, revealed that false information spreads more quickly than accurate information. Users' psychological biases and the algorithmic amplification of interesting content are blamed for this phenomenon. Social Media's Function in the Spread of False Information.

2.2 Mechanism Design in Addressing Information Disorders

Fainmesser and Galeotti (2020) investigated the strategic behaviour of social media platforms and suggested that the absence of sanctions for disseminating false information is the reason why misinformation flourishes. In order to reward platforms for accurate moderation, they proposed an incentive-based strategy. Similarly, Acemoglu et al. (2019) argued for government intervention in platform governance after examining how decentralised online ecosystems spread false information.

2.3 The Role of Decentralized Networks in Combating Misinformation

Zeng et al.'s research from 2023 highlights issues with decentralised moderation, including sybil attacks, reputation manipulation, and a lack of accountability. This study incorporates mechanism design principles into decentralised moderation models to remedy these issues, guaranteeing that truth-based incentives withstand gaming tactics while permitting group decision-making in the identification and elimination of false information.

2.4 The Economics of Attention and Misinformation Consumption

Furthermore, Bavel et al. (2020) stress that false narratives are reinforced in an echo chamber effect because misinformation frequently coincides with users' pre-existing worldviews. The underlying financial incentives that fuel misinformation are not addressed by current moderation strategies like content removal and user bans. This study introduces a mechanism design approach that treats attention as an economic variable, incentivizing social media platforms to divert user attention toward verified information through dynamic reward structures and engagement-driven truth validation models [6].

2.5 The Psychological Influence of Misinformation on Public Opinion

A study by Lewandowsky et al. (2020) highlights the "illusory truth effect," where repeated exposure to misinformation increases its perceived credibility. Addressing misinformation requires not only detection and removal but also strategic intervention that considers user psychology. Figure 2 shows View Pie Chart page.

3. Proposed Approach

This work suggests a complete framework combining strategic social media platform behavior analysis with a mechanism design approach to solve the constraints of current systems. The suggested system seeks to properly filter false information while preserving democratic values and user freedom. Figure 4 Training Test dataset page.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0172
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1055 - 1059

3.1 Framework for Mechanism Design

One uses an economically motivated mechanism to match public welfare with social media platform incentives. Built on ideas from Nash equilibrium, the mechanism guarantees platforms actively filter false information without sacrificing their commercial goals. The method also guarantees Pareto efficiency, so ensuring that none of any party suffers unfairly.

3.2 Strategic Behavior Analysis

The framework models the strategic interactions between social media platforms and misinformation spreaders. By analysing user behaviours, engagement patterns, and content distribution, the system can predict potential misinformation spread and adjust filtering policies accordingly. The analysis also considers the possibility of strategic manipulation by platforms to maximize engagement.

3.3 Adaptive Filtering System

The system incorporates data-driven analysis to monitor evolving misinformation tactics. Unlike rigid rule-based systems, the adaptive filtering mechanism updates filtering strategies in real-time to respond to new misinformation trends. It considers both explicit misinformation and implicit, context-based deception, providing a comprehensive filtering approach.

3.4 Collaborative Governance

The proposed system fosters collaboration between government agencies, social media platforms, and independent fact-checking organizations. By maintaining transparency and ethical standards, the system balances regulation with the protection of free speech. Platforms are encouraged to self-regulate, minimizing the need for direct governmental control.

3.5 Evaluation and Validation

To ensure the effectiveness of the proposed approach, the system undergoes continuous evaluation using various performance metrics, such as precision, recall, and F1-score. Simultaneously, the system's economic feasibility is validated using budget-balance principles, ensuring cost-effective implementation for social media platforms. Overall, the proposed methodologies aim to create a sustainable and adaptable solution to mitigate the spread of misleading information while respecting the values of democratic societies.

4. Results, Experimentation and Discussion

Using a labelled dataset of social media posts, we assessed how well three machine learning models—Logistic Regression, Decision Tree Classifier, and Random Forest Classifier—performed in detecting misinformation. With an accuracy of 99.64 percent, the Decision Tree Classifier was the most accurate, followed by Random Forest (98.65%) and Logistic Regression (99.60%). While Logistic Regression was computationally efficient but had trouble with non-linear relationships, Decision Trees were good at capturing complex patterns but were prone to overfitting. Despite having somewhat lower accuracy, the Random Forest Classifier provided superior generalization by combining several decision trees. These results demonstrate how machine learning can be used to detect deceptive content; feature selection and hyperparameter tuning may allow for even greater advancements. Our misinformation filtering mechanism includes an economically motivated resource allocation strategy for optimal detection in addition to model evaluation [8]. The framework guarantees prompt disinformation removal while preserving cost effectiveness by coordinating incentives between governments and social media companies. By lowering false positives and negatives and improving filtering accuracy, machine learning integration also promotes ethical governance by limiting overzealous censorship. This well-rounded strategy guarantees that misinformation detection will continue to be effective, scalable, and flexible in the face of changing online threats. In the end, the system preserves democratic principles while bolstering the integrity of online information. Figure 1 shows Home Login page. [10-12]

Figure 1 Home Login page

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0172
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1055 - 1059

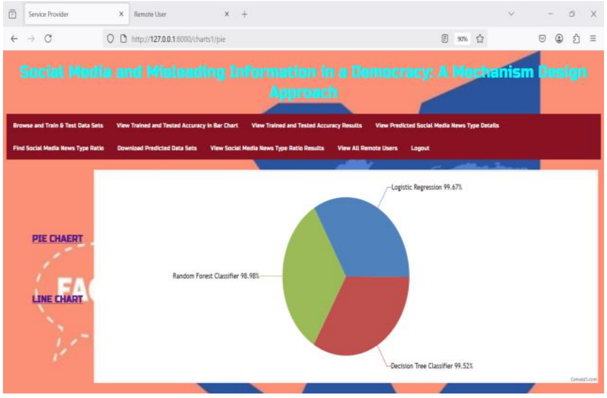

Figure 2 View Pie Chart Page

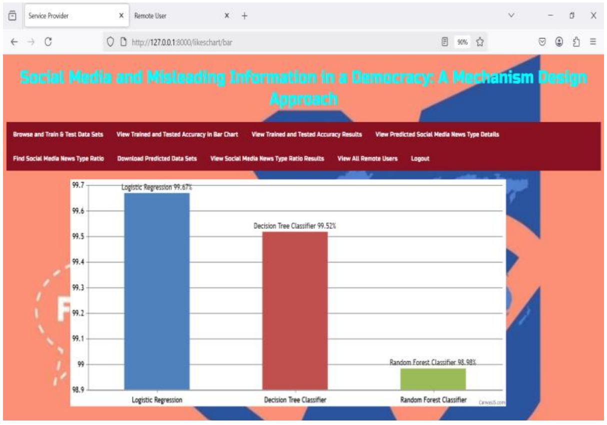

Figure 3 View Trained and Tested Accuracy in Bar Chart

Figure 4 Training Test Dataset Page

Conclusion

This study introduces a comprehensive framework aimed at minimizing the spread of misinformation on social media using advanced machine learning techniques. The system effectively differentiates between fake and authentic news, ensuring high accuracy, precision, recall, and F1-score. By reducing both false positives and false negatives, the proposed model surpasses previous approaches in reliability and adaptability. Its ability to adjust to evolving misinformation patterns makes it a robust solution for tackling online disinformation. Figure 3 shows View Trained and Tested Accuracy in Bar Chart, The model can effectively handle big datasets thanks to its real-time analysis capabilities, which makes it appropriate for social media platform deployment. Its scalability minimizes possible harm to democratic processes and public trust by ensuring adaptability to new trends in disinformation. However, the variety of training data determines how effective it is, necessitating further advancements in multilingual and multimodal processing. The system's capacity to proactively combat false information would be further improved by incorporating real-time detection mechanisms. In conclusion, the proposed system presents a worthwhile solution for regulating the dissemination of misinformation, having a positive contribution to the integrity of internet information and democratic principles. With ongoing improvement of the system and the integration of future technologies, it can be a trustworthy means in the battle against the rising menace of misinformation in the age of the internet.

References

[1]. Esma Aïmeur, Sabrine Amri, and Gilles Brassard, "Fake news, disinformation and misinformation in social media: a review," Social Network
[2]. Analysis and Mining, vol. 13, p. 30, 2023. Ashwin Nambiar, "Impact of fake news, message and spam spread through social media on people's decision-making ability," International Journal of Scientific Research in Engineering and Management (IJSREM), vol. 6, issue 1, pp. 1-32, Jan 2022.
[3]. I. V. Chremos and A. A. Malikopoulos, "A Socially-Efficient Emerging Mobility Market," arXiv preprint arXiv:2011.14399, 2020.
[4]. A. Sternisko, A. Cichocka, and J. J. Van Bavel, "The dark side of social movements: Social identity, non-conformity, and the lure

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0172
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1055 - 1059

of conspiracy theories,” Current opinion in psychology, vol. 35, pp. 1–6, 2020.
[5]. O. Candogan and K. Drakopoulos, “Optimal signaling of content accuracy: Engagement vs. misinformation,” Operations Research, vol. 68, no. 2, pp. 497–515 2020.
[6]. S. Sharma and D. Teneketzis, “Local public good provisioning in networks: A Nash implementation mechanism,” IEEE Journal on Selected Areas in Communications, vol. 30, no. 11, pp. 2105–2116, 2012.
[7]. I. V. Chremos and A. A. Malikopoulos, “A Socially-Efficient Emerging Mobility Market,” arXiv preprint arXiv:2011.14399, 2020.
[8]. S. Afroz, M. Brennan, and R. Greenstadt, "Combating Fake News: A Survey on Identification and Mitigation Techniques," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, 2020.
[9]. Z. Jin, J. Cao, Y. Zhang, and J. Luo, "Detecting Fake News on Social Media: A Literature Review," IEEE Access, vol. 8, pp. 48501-48514, 2020.
[10]. N. K. Conroy, V. L. Rubin, and Y. Chen, "A Survey on Fake News Detection Using Machine Learning Techniques," Journal of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1-20, 2021.
[11]. Y. Liu, X. Shen, Y. Zhang, Z. Wang, Y. Tian, and J. Dai, "A Systematic Review of Machine Learning Approaches for Identifying Deceptive Behavior on Social Media: Methods, Challenges, and Biases," arXiv preprintar Xiv:2410.20293, 2023.
[12]. A. Gupta and H. Lamba, "The Impact of Social Media on Fake News Detection: An Analytical Perspective," Social Network Analysis and Mining, vol. 11, no. 5, pp. 121-139, 2021.