

## *Deep Learning-based Sentiment Analysis for Detecting and Addressing Social Media Rumour Spreading*

Analisis Sentimen Berbasis Pembelajaran Mendalam untuk Mendeteksi dan Mengatasi Penyebaran Rumor di Media Sosial

**Jonson Manurung<sup>1</sup>, Poltak Sihombing<sup>2</sup>, Mohammad Andri Budiman<sup>3</sup>, Sawaluddin<sup>4</sup>**

<sup>1,2,3</sup> Faculty of Computer Science and Information Technology, University of Sumatera Utara, Medan, Indonesia

<sup>4</sup> Department of Mathematics and Natural Sciences, University of Sumatera Utara. Medan, Indonesia

---

### **ABSTRACT**

*This research introduces a novel deep learning-based framework for the detection and mitigation of social media rumor spreading through a combination of sentiment analysis and rumor detection. Leveraging advanced neural network architectures, including recurrent neural networks and additional features such as user engagement and source credibility, the model aims to provide a comprehensive understanding of the contextual factors influencing the dissemination of rumors. The numerical example demonstrates the model's capacity to capture nuanced sentiment dynamics in social media posts, enabling the identification of potential rumors by considering both emotional tone and network-related features. The proposed approach holds promise for enhancing the accuracy of rumor detection systems, fostering more informed content moderation strategies, and contributing to the broader efforts aimed at combating misinformation in online communities. Ethical considerations, interpretability, and future directions for model refinement are integral aspects addressed in the research, emphasizing a holistic approach to addressing the complex challenges posed by social media rumor spreading*

**Keywords :** Social Media Rumors, Sentiment Analysis, Deep Learning, Rumor Detection, Information Dissemination

### **ABSTRAK**

*Penelitian ini memperkenalkan kerangka kerja baru berbasis pembelajaran mendalam untuk mendeteksi dan memitigasi penyebaran rumor di media sosial melalui kombinasi analisis sentimen dan deteksi rumor. Memanfaatkan arsitektur jaringan saraf canggih, termasuk jaringan saraf berulang dan fitur tambahan seperti keterlibatan pengguna dan kredibilitas sumber, model ini bertujuan untuk memberikan pemahaman komprehensif tentang faktor kontekstual yang mempengaruhi penyebaran rumor. Contoh numerik menunjukkan kapasitas model untuk menangkap dinamika sentimen yang berbeda dalam postingan media sosial, sehingga memungkinkan identifikasi potensi rumor dengan mempertimbangkan nada emosional dan fitur terkait jaringan. Pendekatan yang diusulkan ini menjanjikan peningkatan akurasi sistem deteksi rumor, mendorong strategi moderasi konten yang lebih terinformasi, dan berkontribusi pada upaya yang lebih luas yang bertujuan memerangi misinformasi di komunitas online. Pertimbangan etis, kemampuan interpretasi, dan arah masa depan untuk penyempurnaan model merupakan aspek integral yang dibahas dalam penelitian ini, yang menekankan pendekatan holistik untuk mengatasi tantangan kompleks yang ditimbulkan oleh penyebaran rumor di media sosial..*

**Kata Kunci :** Rumor Media Sosial, Analisis Sentimen, Pembelajaran Mendalam, Deteksi Rumor, Penyebaran Informasi

---

Info Artikel :

Disubmit: 09 Oktober

Direview: 04 November

Diterima : 20 Desember

## I. INTRODUCTION

In recent years, the pervasive influence of social media has transformed the way information is disseminated and consumed, reshaping public discourse on a global scale [1][2][3]. While these platforms offer unprecedented connectivity, they also amplify the challenges associated with misinformation, rumors, and the rapid spread of unverified content. The unchecked propagation of false narratives not only erodes the trustworthiness of online information but also has tangible consequences for individuals, communities, and even broader societal structures.

The dynamics of rumor spreading on social media are complex, often fueled by emotional reactions and sentiment-laden content. Traditional approaches to sentiment analysis and rumor detection, relying on rule-based systems or simplistic machine learning models, struggle to capture the intricacies of language, context, and evolving online discourse. As a response to these challenges, the emergence of deep learning techniques has opened new avenues for more accurate and nuanced analysis of textual data [4][5][6].

Deep learning models, including recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and transformer-based architectures, have demonstrated remarkable success in natural language processing tasks [7][8][9]. Leveraging the power of these models for sentiment analysis provides an opportunity to uncover the underlying emotional tone of social media content, thereby aiding in the identification of potential rumors and misinformation [10].

The effective integration of deep learning-based sentiment analysis into a comprehensive framework for rumor detection on social media remains an underexplored area of research [11][12]. Bridging this gap is crucial for developing robust systems that not only discern sentiment polarity but also understand the contextual factors that contribute to the amplification or suppression of information within online communities [13][14][15].

In the contemporary digital landscape, social media platforms serve as ubiquitous sources of information dissemination, enabling rapid communication and connection among users globally [16][17]. However, this interconnectedness also facilitates the swift propagation of rumors, misinformation, and false narratives, posing a significant challenge to the credibility and integrity of online discourse [18][19]. Addressing the rampant spread of social media rumors requires advanced technological solutions that can accurately identify and assess the sentiment behind user-generated content in real-time [20][18][21].

Traditional methods of rumor detection and sentiment analysis often fall short in handling the dynamic and nuanced nature of language on social media platforms [22][23]. The rise of deep learning techniques, such as Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), and Transformer-based models, offers a promising avenue for enhancing the accuracy and efficiency of sentiment analysis [24][25][26][27][28]. However, the integration of these advanced models into a cohesive framework for detecting and addressing social media rumor spreading remains an open research challenge [29][30].

This research builds upon the foundation laid by existing studies in sentiment analysis, rumor detection, and deep learning applications in natural language processing. By delving into the intersection of these domains, the aim is to advance our understanding of the challenges posed by social media rumor spreading and to propose innovative solutions that leverage the capabilities of deep learning models. The insights gained from this research have the potential to inform the development of practical tools and

strategies for mitigating the impact of misinformation, fostering a healthier and more reliable online information ecosystem.

Research by Ma et al [31], with title Rumor Detection on Twitter with Tree-structured Recursive Neural Networks, the study proposed a tree-structured recursive neural network for rumor detection on Twitter. The model demonstrated improved performance compared to traditional methods, effectively capturing the hierarchical structure of Twitter conversations. The research highlighted the importance of considering the conversational context in rumor detection tasks. Research by Zubiaga, et al [32] with the title Detection and Resolution of Rumours in Social Media: A Survey, said The survey provided a comprehensive overview of existing approaches to rumor detection in social media. It highlighted the challenges in distinguishing between rumors and non-rumors and emphasized the role of temporal information. The research called for further exploration of real-time detection techniques and the integration of multiple data sources, as well as research conducted by Ma, et al. [33] with title Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning, in their research they found this study proposed a kernel learning approach to detect rumors in microblog posts by considering the propagation structure. The research demonstrated the significance of capturing the information flow patterns in social networks for rumor detection. The findings suggested that incorporating propagation features enhances the performance of detection models. Research by Ruchansky, et al. [34] with title CSI: A Hybrid Deep Model for Fake News Detection, the research introduced a hybrid deep learning model, CSI (Convolutional neural network, Siamese network, and Inception network), for fake news detection. The study emphasized the importance of combining various neural network architectures to effectively capture different aspects of textual information. The findings showcased the potential of hybrid models in addressing the complexity of misinformation detection.

## II. Research Methods

The conceptual framework for this research on "Deep Learning-based Sentiment Analysis for Detecting and Addressing Social Media Rumor Spreading" encompasses three key components:

### Deep Learning-Based Sentiment Analysis

1. Input: Social media posts (textual data) from diverse sources.
2. Process: Utilize advanced deep learning architectures, such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), or transformer-based models, for sentiment analysis. Incorporate pre-trained word embeddings and attention mechanisms to capture contextual nuances.
3. Output: Predicted sentiment labels (positive, negative, neutral) for each social media post.

### Rumor Detection Framework

1. Input: Sentiment labels, along with additional features like user engagement, source credibility, and network dynamics.
2. Process: Integrate sentiment analysis results into a broader rumor detection model. Explore the correlation between sentiment and rumor propagation. Consider the temporal aspect of information flow to distinguish rumors from non-rumors.
3. Output: Identification of potential rumors, with contextual information on sentiment and other relevant features.

### Response Mechanism

1. Input: Detected rumors and associated contextual information.
2. Process: Develop a response mechanism that may include notifying users, flagging content, or providing informative responses. Consider ethical implications, user privacy, and platform guidelines.
3. Output: Implementation of appropriate responses to mitigate the impact of identified rumors.

What follows is a high-level mathematical formulation for deep learning-based sentiment analysis and rumour detection models.

#### 1. Sentiment Analysis

Inputs:

$X_i$  : represents the input social media post

$W$  : represents the word embeddings matrix.

Model:

$H_i$  : is the hidden representation obtained through the sentiment analysis model.

$\Sigma$  : is the activation function, commonly the sigmoid or softmax function.

The sentiment analysis model can be represented as:

$$H_i = \sigma(W \cdot X_i)$$

#### 2. Rumor Detection

Inputs:

$S_i$  : represents the sentiment score for social media post

$E_i$  : represents additional features (e.g., user engagement, source credibility).

Model:

$R_i$  : represents the rumor score for social media post

$U$  : is the weight matrix for combining sentiment and additional features.

The rumor detection model can be represented as:

$$R_i = \sigma(U \cdot [S_i, E_i])$$

#### 3. Overall Model:

Combining sentiment analysis and rumor detection, the overall model can be represented as:

$$P_i = R_i \cdot H_i$$

Here,  $P_i$  represents the probability of social media post  $i$  being a rumor, considering both sentiment analysis and additional features.

#### 4. Loss Function:

The loss function ( $L$ ) for training the model is typically a binary cross-entropy loss, minimizing the difference between predicted rumor scores and actual labels (rumor or non-rumor).

$$L = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log[P_i] + (1 - y_i) \cdot \log(1 - P_i))$$

Where  $y_i$  is the ground truth label (1 for rumor, 0 for non-rumor), and  $N$  is the number of samples in the training set.

### III. Result And Discussion

A numerical example to illustrate the concepts outlined in the mathematical formulation will consider a dataset with three social media posts ( $i = 1, 2, 3$ ) and three features (sentiment score  $S_i$  and ground truth label  $y_i$ ). We will use randomly generated values for illustration.

Given data

$$X_1=[0.2,0.4,0.6], X_2=[0.52,0.1,0.8], X_3=[0.3,0.7,0.2]$$

$$W = \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.2 & 0.3 & 0.4 \\ 0.3 & 0.4 & 0.5 \end{bmatrix}, U = \begin{bmatrix} 0.4 \\ 0.5 \\ 0.6 \end{bmatrix}$$

Sentiment Analysis

$$H_1 = \sigma(W \cdot X_1) = \sigma([0.48, 0.62, 0.78]) \approx [0.618, 0.650, 0.685]$$

$$H_2 = \sigma(W \cdot X_2) = \sigma([0.76, 0.98, 1.20]) \approx [0.681, 0.727, 0.768]$$

$$H_3 = \sigma(W \cdot X_3) = \sigma([0.62, 0.82, 1.02]) \approx [0.650, 0.694, 0.733]$$

Rumor Detection

$$R_1 = \sigma(U \cdot [S_1, E_1]) = \sigma([0.4 \cdot 0.618 + 0.5 \cdot 0.4, 0.6 \cdot 0.618 + 0.5 \cdot 0.6]) \approx [0.736]$$

$$R_2 = \sigma(U \cdot [S_2, E_2]) = \sigma([0.4 \cdot 0.681 + 0.5 \cdot 0.1, 0.6 \cdot 0.681 + 0.5 \cdot 0.8]) \approx [0.841]$$

$$R_3 = \sigma(U \cdot [S_3, E_3]) = \sigma([0.4 \cdot 0.650 + 0.5 \cdot 0.7, 0.6 \cdot 0.650 + 0.5 \cdot 0.2]) \approx [0.775]$$

Overall Model

$$P_1 = R_1 \cdot H_1 = [0.736] \cdot [0.618, 0.650, 0.685] \approx [0.540, 0.559, 0.581]$$

$$P_2 = R_2 \cdot H_2 = [0.841] \cdot [0.681, 0.727, 0.768] \approx [0.572, 0.598, 0.631]$$

$$P_3 = R_3 \cdot H_3 = [0.775] \cdot [0.650, 0.694, 0.733] \approx [0.503, 0.542, 0.567]$$

### Loss Calculation

Assuming ground truth labels  $y_1=1$ ,  $y_2=0$ ,  $y_3=1$ ,

$$L = -\frac{1}{3}(1 \cdot \log(P_1) + 0 \cdot \log(P_2) + 1 \cdot \log(P_3))$$

### Discussion

The results demonstrate the interplay between sentiment analysis and rumor detection in assessing the likelihood of a social media post being a rumor. Post 2, with a more positive sentiment, received a higher rumor score and overall probability of being a rumor. This aligns with the hypothesis that highly positive sentiment may contribute to rumor propagation. Post 3, with sentiment similar to Post 1 but slightly less positive, received a moderate overall probability of being a rumor. This emphasizes the importance of considering additional features beyond sentiment in rumor detection.

## IV conclusion

This research endeavors to address the challenge of detecting and addressing social media rumor spreading through the development of a deep learning-based sentiment analysis and rumor detection model. The integration of sentiment analysis provides valuable context to the model. Posts with more positive sentiment may have a higher likelihood of being associated with rumors, suggesting a potential correlation between sentiment and misinformation propagation. The rumor detection mechanism successfully combines sentiment scores with additional features to assess the likelihood of a post being a rumor. This approach acknowledges the multifaceted nature of rumor spreading, considering not only the emotional tone but also user engagement and source credibility. The numerical example highlights the nuanced interplay of factors contributing to the model's predictions. The combination of sentiment analysis and rumor detection allows for a more comprehensive evaluation of social media content, capturing both emotional context and contextual features. The model's outputs, represented as probabilities, offer a level of interpretability. By examining these probabilities, users and stakeholders can gain insights into the likelihood of a post being a rumor, empowering them to make informed decisions about content moderation and user engagement. It is crucial to acknowledge the limitations of the presented model, including the simplicity of the numerical example. In a real-world scenario, the model's performance would depend on the quality and diversity of the training data, the sophistication of the chosen deep learning architecture, and the effectiveness of the integrated features. This research lays the groundwork for future exploration and refinement. Further research could focus on enhancing the model's accuracy by incorporating more advanced deep learning architectures, experimenting with different feature combinations, and considering the temporal dynamics of social media conversations. The ethical considerations related to privacy, bias, and responsible AI are paramount. As the model is deployed in real-world applications, it is essential to continuously evaluate and address any potential biases, ensuring fairness and equity in its predictions.

## References

- [1] A. Hermida, "Social media and the news," *SAGE Handb. Digit. Journal.*, pp. 81–94, 2016.
- [2] A. Jungherr, G. Rivero, and D. Gayo-Avello, *Retooling politics: How digital media are shaping democracy*. Cambridge University Press, 2020.

- [3] D. Patrikarakos, *War in 140 characters: how social media is reshaping conflict in the twenty-first century*. Hachette UK, 2017.
- [4] R. Debnath, S. Darby, R. Bardhan, K. Mohaddes, and M. Sunikka-Blank, "Grounded reality meets machine learning: A deep-narrative analysis framework for energy policy research," *Energy Res. Soc. Sci.*, vol. 69, p. 101704, 2020.
- [5] K. Kheiri and H. Karimi, "Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning," *arXiv Prepr. arXiv2307.10234*, 2023.
- [6] Z. Amiri, A. Heidari, N. J. Navimipour, M. Unal, and A. Mousavi, "Adventures in data analysis: A systematic review of Deep Learning techniques for pattern recognition in cyber-physical-social systems," *Multimed. Tools Appl.*, pp. 1–65, 2023.
- [7] A. Malte and P. Ratadiya, "Evolution of transfer learning in natural language processing," *arXiv Prepr. arXiv1910.07370*, 2019.
- [8] T. Shao, Y. Guo, H. Chen, and Z. Hao, "Transformer-based neural network for answer selection in question answering," *IEEE Access*, vol. 7, pp. 26146–26156, 2019.
- [9] S. Chaudhury and K. Sau, "A BERT encoding with Recurrent Neural Network and Long-Short Term Memory for breast cancer image classification," *Decis. Anal. J.*, vol. 6, p. 100177, 2023.
- [10] P. Meel and D. K. Vishwakarma, "Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities," *Expert Syst. Appl.*, vol. 153, p. 112986, 2020.
- [11] A. Choudhary and A. Arora, "Linguistic feature based learning model for fake news detection and classification," *Expert Syst. Appl.*, vol. 169, p. 114171, 2021.
- [12] J. An, W. M. N. W. Zainon, and B. Ding, "Leveraging Vision-Language Pre-Trained Model and Contrastive Learning for Enhanced Multimodal Sentiment Analysis," *Intell. Autom. Soft Comput.*, vol. 37, no. 2, 2023.
- [13] H. Rheingold, "A slice of life in my virtual community," *Glob. networks Comput. Int. Commun.*, pp. 57–80, 1993.
- [14] I. C. Anderson and J. W. G. Cairney, "Diversity and ecology of soil fungal communities: increased understanding through the application of molecular techniques," *Environ. Microbiol.*, vol. 6, no. 8, pp. 769–779, 2004.
- [15] M. A. Smith and P. Kollock, *Communities in cyberspace*. Psychology Press, 1999.
- [16] A. H. Ali, "The power of social media in developing nations: New tools for closing the global digital divide and beyond," *Harv. Hum. Rts. J.*, vol. 24, p. 185, 2011.
- [17] M. Vuori, "Exploring uses of social media in a global corporation," *J. Syst. Inf. Technol.*, vol. 14, no. 2, pp. 155–170, 2012.
- [18] P. Iosifidis and N. Nicoli, *Digital democracy, social media and disinformation*. Routledge, 2020.
- [19] S. Tufchi, A. Yadav, and T. Ahmed, "A comprehensive survey of multimodal fake news detection techniques: advances, challenges, and opportunities," *Int. J. Multimed. Inf. Retr.*, vol. 12, no. 2, p. 28, 2023.
- [20] X. Liu *et al.*, "Reuters tracer: A large scale system of detecting & verifying real-time news events from twitter," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, 2016, pp. 207–216.
- [21] J. Van Dijk and D. Alinejad, "Social media and trust in scientific expertise: Debating the Covid-19 pandemic in the Netherlands," *Soc. Media+ Soc.*, vol. 6, no. 4, p. 2056305120981057, 2020.
- [22] T. K. Balaji, C. S. R. Annavarapu, and A. Bablani, "Machine learning algorithms for social media analysis: A survey," *Comput. Sci. Rev.*, vol. 40, p. 100395, 2021.
- [23] M. Ghiassi and S. Lee, "A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach," *Expert Syst. Appl.*, vol. 106, pp. 197–216, 2018.
- [24] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," *arXiv Prepr. arXiv1801.01078*, 2017.
- [25] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, pp. 650–655, 2021.
- [26] W. Xu, Y. Jiang, X. Zhang, Y. Li, R. Zhang, and G. Fu, "Using long short-term memory networks for river flow prediction," *Hydrol. Res.*, vol. 51, no. 6, pp. 1358–1376, 2020.
- [27] A. Gillioz, J. Casas, E. Mugellini, and O. Abou Khaled, "Overview of the Transformer-based Models for NLP Tasks," in *2020 15th Conference on Computer Science and Information Systems (FedCSIS)*, 2020, pp. 179–183.
- [28] P. Ganesh *et al.*, "Compressing large-scale transformer-based models: A case study on bert," *Trans. Assoc. Comput. Linguist.*, vol. 9, pp. 1061–1080, 2021.

- [29] B. Abboodi, S. F. Pileggi, and G. Bharathy, "Social Networks in Crisis Management: A Literature Review to Address the Criticality of the Challenge," *Encyclopedia*, vol. 3, no. 3, pp. 1157–1177, 2023.
- [30] B. Abboodi, S. F. Pileggi, and G. Bharathy, "Social Networks in Crisis Management: a Concise Literature Review," 2023.
- [31] J. Ma, W. Gao, and K.-F. Wong, "Rumor detection on twitter with tree-structured recursive neural networks," 2018.
- [32] A. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, and R. Procter, "Detection and resolution of rumours in social media: A survey," *ACM Comput. Surv.*, vol. 51, no. 2, pp. 1–36, 2018.
- [33] J. Ma, W. Gao, and K.-F. Wong, "Detect rumors in microblog posts using propagation structure via kernel learning," 2017.
- [34] N. Ruchansky, S. Seo, and Y. Liu, "Csi: A hybrid deep model for fake news detection," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 797–806.