IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, October 2023

Unmasking Hate: A Multimodal Approach to Hateful Meme Detection

Ishaan Bagul, Roshani Mourya, Khushabu Rindani, Abhijeet Shinde, Jyoti Thakur
Department of Artificial Intelligence & Machine Learning
Loknete Gopinathji Munde Institute of Engineering Education & Research, Nashik, India

Abstract: *Hateful memes are an escalating issue in the digital landscape, demanding innovative solutions for their effective detection and classification. These memes often employ subtlety, sarcasm, and symbolism, presenting formidable challenges for automated detection systems. Moreover, the linguistic and cultural diversity of the internet, transcending geographical and language boundaries, further complicates the task. This research paper presents a comprehensive approach to hateful meme detection, utilizing a Dual Stream Transformer Model, real-world knowledge integration, characteristic detection, and cultural reference understanding. We emphasize the importance of ethics and responsible usage in deploying such technology, underscoring its potential for positive societal impact.*

Keywords: Derogatory, Dual Stream Transformer Model, holistic understanding, convolutional neural networks, multimedia content, Contextual Understanding, Transformer-Based Analysis, visual stream, multimodal.

REFERENCES

[1]. Awan, I., & Blakemore, B. (2019). "Hate speech and co-radicalization processes in the digital spaces: Developing an agenda for research." Studies in Conflict & Terrorism, 42(2), 115-127.
[2]. Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). "Automated hate speech detection and the problem of offensive language." In Proceedings of the Eleventh International Conference on Weblogs and Social Media, 512-515.
[3]. Fortuna, P., Pestian, J., Dehghani, M., & Kamal, N. (2018). "A survey of available corpora for building data-driven dialogue systems." Dialogue & Discourse, 9(1), 12-46.
[4]. Gao, H., Zhang, Y., Xu, Y., Ma, Y., Su, Z., & Cui, L. (2020). "A survey of hate speech detection using natural language processing." Information Processing & Management, 58(2), 102067.
[5]. Hosseinmardi, H., Mattson, S. A., Rafiq, R. I., Han, R., Lv, Q., Mishra, S., ... &Lv, Q. (2015). "Analyzinglabeled cyberbullying incidents on the Instagram social network." In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 229-236.
[6]. Kwok, I., Wang, Y., &Derczynski, L. (2013). "Locate the hate: Detecting tweets against blacks." In Proceedings of the International Workshop on Semantic Evaluation, 497-501.
[7]. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., Chang, Y., & Solorio, T. (2016). "Abusive language detection in online user content." In Proceedings of the 25th International Conference on World Wide Web, 145-153.
[8]. Persing, I., & Nguyen, D. T. (2018). "Model selection in hate speech detection: What works when." In Proceedings of the Third Workshop on Abusive Language Online, 76-86.
[9]. Waseem, Z., &Hovy, D. (2016). "Hateful symbols or hateful people? Predictive features for hate speech detection on Twitter." In Proceedings of the NAACL Student Research Workshop, 88-93.
[10]. Wiegand, M., Ruppenhofer, J., Kleinbauer, T., & Seifert, C. (2018). "Overview of the GermEval 2018 shared task on the identification of offensive language." In Proceedings of the GermEval 2018 Workshop, 35-44.
[11]. Wulczyn, E., Thain, N., & Dixon, L. (2017). "Ex Machina: Personal attacks seen at scale." In Proceedings of the 26th International Conference on World Wide Web, 1391-1399.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, October 2023

[12]. Zhou, P., Zhang, P., Hu, R., & Zhu, F. (2020). "A new survey for hate speech detection: Challenges and solutions." Journal of King Saud University-Computer and Information Sciences.