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: 7.53
Volume 4, Issue 6, April 2024

Anti-Semitic Speech Detection and Classification in Online Social Network using Deep Learning

Mrs. P. Elakkiya[1], Abiya M[2], Divya Bharathi V[3], Nandhini R[4]
Assistant Professor, Department of Computer Science and Engineering[1]
Student, Department of Computer Science and Engineering[2,3,4]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamil Nadu, India

Abstract: *Every individual possesses the entitlement to freedom of speech. However, in the guise of free expression, this privilege is being abused to discriminate against and harm other people. This prejudice is referred to as hate speech. A clear definition of hate speech is language that expresses hatred for an individual or a group of individuals based on traits including race, religion, ethnicity, gender, nationality, handicap, and sexual orientation. Hate speech has become increasingly widespread, both in physical spaces and on the internet, in recent years. Thus, recent studies used a range of machine learning and deep learning techniques with text mining method to automatically recognise the hate speech messages on real-time datasets in order to address this developing issue in social media sites. This project's goal is to examine comments on social networks using Natural Language Processing (NLP) and a Deep Learning method called VADER method. In order to identify the text as positive or negative, VADER neural networks are used to extract the keywords from user generated content. If it's negative, immediately block the comments in accordance with the user's preferences and block the friends in accordance with pre-established threshold values. The proposed framework was deployed in a real-time social networking site with an improved notification system, according to experimental findings*

Keywords: NLP (Natural Language Processing), Deep Learning Models, Sentiment Analysis, Online Social Networks Keyword Extraction, Text Mining, VADER (Valence Aware Dictionary and Sentiment Reasoner) Algorithms

REFERENCES

[1] Singh, A., Kumar, S., & Gupta, R. (2023). Application of Deep Learning Models for Detecting Hate Speech in Online Social Networks. Presented at the 2023 IEEE International Conference on Big Data (Big Data) (pp. 210-218).
[2] Gupta, A., Varma, V., & Gupta, M. (2022). Deep Learning Approaches for Hate Speech Detection on Social Media: A Comparative Study. In Proceedings of the 2022 IEEE International Conference on Data Mining (ICDM) (pp. 102-110).
[3] Ranasinghe, T., & Meedeniya, D. (2021). Hate Speech Detection and Classification in Online Social Networks using Deep Learning Models. In Proceedings of the 2021 IEEE International Conference on Big Data (Big Data) (pp. 320-328).
[4] Chakraborty, A., Mondal, M., & Saha, S. (2020). Hate Speech Detection in Online Social Networks Using Deep Learning Techniques. In Proceedings of the 2020 International Conference on Data Science and Machine Learning (pp. 87-95).
[5] Basile, V., Caputo, A., Castellucci, G., Patti, V., & Rosso, P. (2019). Grasping abuse: An arrangement of subtasks for detecting abusive language. Journal of experimental & theoretical artificial intelligence.
[6] Waseem, Z., & Hovy, D. (2018). Exploring abuse: Categorizing subtasks for detecting abusive language. In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers).
[7] Founta, A. M., Djouvas, C., Chatzakou, D., Leontiadis, I., Blackburn, J., Stringhini, G., and Kourtellis, N. (2018). Extensive crowdsourcing and examination of abusive conduct on Twitter. Delivered at the Twelfth International AAAI Conference on Web and Social Media.

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: 7.53
Volume 4, Issue 6, April 2024

[8] Fortuna, P., Nunes, S., & Rodrigues, F. (2018). Assessing automated techniques for identifying hate speech in textual content. ACM Computing Surveys (CSUR).
[9] Burnap, P., & Williams, M. L. (2017). Differentiating cyber hate on Twitter based on multiple protected characteristics. EPJ Data Science, 6(1), 1-20.
[10] Zhang, L., & Wang, Y. (2017). "Deep Learning for Detecting Cyberbullying Across Multiple Social Media Platforms." In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining .
[11] Deutsch, D., Freitas, A., D'Amico, D., & Santamaría, J. J. (2017). Detecting Hate Speech on Twitter using a Convolution-GRU Based Deep Neural Network.
[12] Razavi, A. H., Marcus, A., & Rus, D. (2016). Handling Multimodal Hate Speech with Deep Learning: A Case Study of Facebook.
[13] Ribeiro, M. T., Calais, P. H. R., Santos, I. S., & Almeida, V. A. F. (2016). Application of Convolutional Neural Networks for Hate-Speech Classification.
[14] Burnap, P., & Williams, M. L. (2015). Identifying cyber hate speech on Twitter: Applying machine classification and statistical modeling for policy and decision-making purposes. Policy & Internet, 7(2), 223-242.
[15] Burnap, P., & Williams, M. L. (2012). Cyber hate speech on web 2.0: An empirical study of UK universities. In SocialCom/PASSAT (pp. 134-139).