IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Analysing Twitter Conversations on Gender Violence: Clustering, Community Detection, and Sentiment Insights

Elizabeth Leah George (Author)
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher Education for Women,
Coimbatore, India

Subashini Parthasarathy (Author)
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher Education for Women,
Coimbatore, India

ABSTRACT
Social media sites like Twitter have become increasingly important over the past few years in creating awareness and provoking debate around social issues like gender-based violence. Social media sites offer extensive user-generated content reflecting public sentiment and engagement on specialised subjects. This study uses clustering, community detection, and sentiment analysis to analyse public debate around gender violence on Twitter.

The research aims at a corpus of 6,799 tweets representing varying gender violence-related discourse. The corpus was initially preprocessed to remove extraneous content such as stopwords, URLs, and uncorrelated tweets. The tweets were then vectorised using TF-IDF (Term Frequency-Inverse Document Frequency) to identify the meaningful attributes. K-Means clustering was employed to group similar tweets, while Louvain's community detection algorithm was employed to identify individual communities of users discussing gender violence. Sentiment analysis was done to classify tweets as positive, negative, or neutral. At the same time, Different evaluation measures, such as Modularity, Silhouette Score, and Davies-Bouldin Index, were used to analyse the efficiency of clustering and community detection. The objective of this study is to create a machine-learning model that classifies gender-related tweets into one of five categories: sexual violence, emotional violence, harmful cultural practices, physical violence, and economic violence. The research indicates that most problems are unreported, and information is filtered through perpetuating mechanisms and consequences for Indian society. The mental well-being of the women and children also impacts the problems in question.

With the support of increasing efforts to promote women's empowerment and mainstreaming gender equality, we can create social awareness on social networking sites with actions that favour women's and girls' development and livelihoods.

Keywords: Gender violence, Women empowerment, Gender equality, Social media activism, Data analysis

• INTRODUCTION

The expansion of social media sites, especially Twitter (now X), has opened new channels for people to voice their views, share experiences, and discuss various issues. But this heightened connectivity has also given rise to online forums where hate speech, harassment, and gender-based violence (GBV) can flourish [1]. Understanding the dynamics of such online discussions is vital for formulating effective strategies to prevent GBV, advance gender equality, and create safer online spaces. This study examines how clustering, community detection, and sentiment analysis methods can be used on Twitter data to understand discussions around gender violence. Twitter has emerged as an essential forum for gender-based violence discussions, with activists, organisations, and individuals utilising it to raise awareness, share experiences, and mobilise for change [2], [3]. The platform's ease of access and real-time nature render it an effective tool for mobilising support and pushing back against social norms reinforcing GBV [3]. Nonetheless, the same attributes that enable Twitter to be an effective activism tool also expose it to abuse, with the perpetrators employing it to harass, threaten, and silence GBV victims [4]. Tremendous strides have been made in advancing women's rights and empowerment over the last few decades, but gender imbalances continue to exist in different aspects of life. Social media websites have been responsible for strengthening feminist voices and supporting activism against gender violence [5].

- The #MeToo campaign, for instance, became hugely popular on Twitter, where there was space for survivors of sexual harassment and assault to narrate their experiences and hold perpetrators to account [6], [7]. Likewise, feminist movements in Latin America and Spain have leveraged "femitags" (feminist hashtags) to organise mass action and generate publicity for concerns like femicide and gender-based violence [5]. While social media could potentially be used to counteract GBV, there are also fears of the proliferation of misogyny, cyber abuse, and silencing marginalised voices [8]. Research has indicated that women, especially feminists and minority group members, are disproportionately targeted for online abuse and harassment [9]. This has the potential to result in self-censorship, avoidance of online spaces, and a chilling effect on free speech [8].

• RELATED WORK

Clustering techniques can be employed to identify the main themes and topics discussed in Twitter conversations related to gender violence [10]. By grouping tweets based on their content, researchers can gain insights into the different aspects of GBV being discussed, the types of narratives being shared, and the perspectives being expressed [11]. Sharma, D., Gupta proposes two clustering methods, partition-based and density-based clustering methods, to identify distinct semantic groups in Twitter data [10]. Partition-based clustering divides the data into non-overlapping clusters, while density-based clustering identifies clusters based on the density of data points in a given region [10]. These methods can be used to identify different subtopics within the broader conversation on gender violence, such as discussions about specific forms of violence, legal reforms, or support services for survivors. Topic modelling techniques, such as Latent Dirichlet Allocation (LDA), can also identify the underlying themes in Twitter conversations [12]. LDA is a probabilistic model that identifies topics based on the co-occurrence of words in a text corpus [13]. By applying LDA to Twitter data, researchers can uncover the key topics that are being discussed concerning gender violence, such as victim blaming, slut-shaming, or the impact of GBV on mental health [14]. Community detection algorithms can identify groups of users actively engaged in discussions about gender violence on Twitter [15], [16]. These algorithms identify communities based on user connections, such as retweets, mentions, and follows [17]. By mapping these social networks, researchers can gain insights into how information and ideas are spread, who the key influencers are, and how different communities interact with each other [18].

Community detection can help identify influential users and organisations playing a key role in shaping the conversation on gender violence [19]. These individuals or groups may be activists, experts, or organisations working to combat GBV [20]. By understanding their networks and shared content, researchers can identify effective strategies for amplifying their messages and reaching wider audiences [21]. Analysing the structure of these online communities can reveal important information about how different groups are connected and interact [22]. For example, researchers may find distinct communities of activists, survivors, and allies who are working together to raise awareness and support victims of GBV [23]. They may also identify communities of individuals spreading misinformation or engaging in harmful behaviours, such as victim blaming or online harassment [24].

This research paper delves into the multifaceted landscape of gender equality initiatives and their impact on empowering women for a sustainable future. By examining the intersectionality of gender with social, economic, and environmental factors, we aim to gain a comprehensive understanding of the challenges women face and the transformative potential of targeted interventions. The subsequent sections are structured as follows: Section III presents the proposed approach for effectively analysing sentimental and emotional analysis through different data visualisation methods. Section IV describes the results of the proposed machine learning and Deep learning algorithms. Finally, Section V concludes by summarising the overall findings and outcomes of the research.

- **PROPOSED METHODOLOGY**

This multi-step methodology to analyse Twitter conversations on gender violence, integrating clustering, community detection, and sentiment analysis, employs a machine learning-based approach to analyse Twitter conversations on gender-based violence using clustering, community detection, and sentiment analysis. A collection of 6799 tweets was created to express and reflect diverse views on violence against women. Preprocessing was conducted to remove irrelevant content, including stopwords, URLs, and unrelated tweets. The tweets were then vectorised using TF-IDF to extract meaningful features. K-Means clustering was applied to categorise tweets into distinct groups based on textual similarities, while Louvain's community detection method was used to identify interconnected user groups discussing gender violence. Sexual violence, emotional violence, harmful cultural practices, physical violence, and economic violence enable a deeper understanding of the discussions.

- **Sentiment Analysis and Emotions classification**

When it comes to tweeting about 'Gender Violence', it seems that there are roughly even numbers of happy, positive tweets and sad, negative tweets. But there seem to be more sad tweets expressing anger or discomfort than happy ones—something to keep in mind. "Femicide" has most tweets classified as neutral, followed by positive and negative tweets in much smaller proportions. "Rape" has a higher number of positive tweets compared to negative tweets, with a significant number of neutral tweets as well.

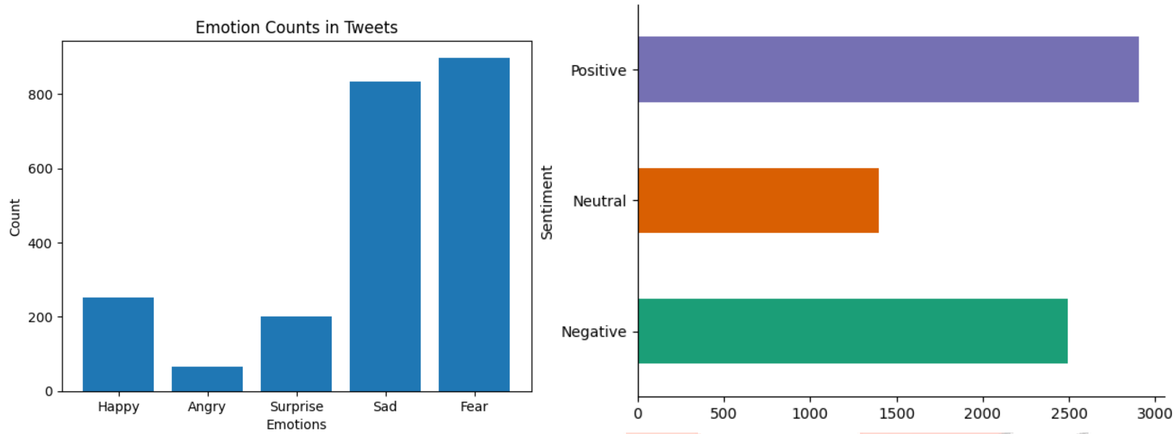
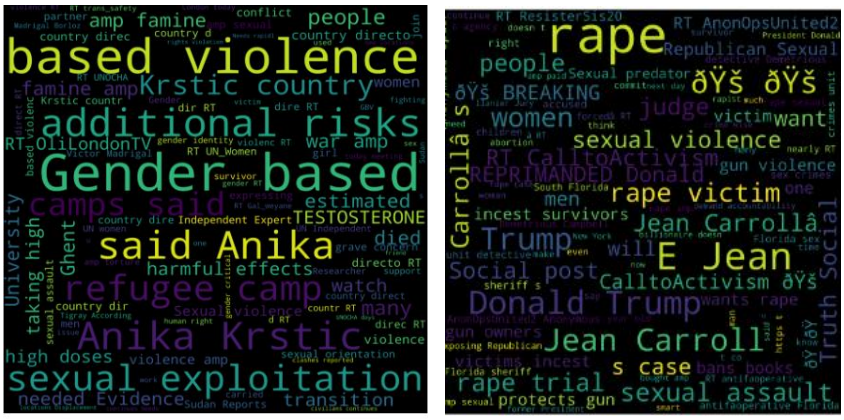


Figure 1: Twitter Keywords Sentimental and Emotional Analysis

Overall, the sentiment analysis on these keywords provides insights into the emotions and opinions expressed in the tweets related to each topic.

A word cloud is a data visualisation technique used to represent the frequency of words in a text corpus. In a word cloud, words are displayed in different sizes based on frequency, with more frequent words appearing larger. This visualisation provides a quick and intuitive way to identify the most common terms in a text or set of documents.

Word Cloud of Positive Tweets



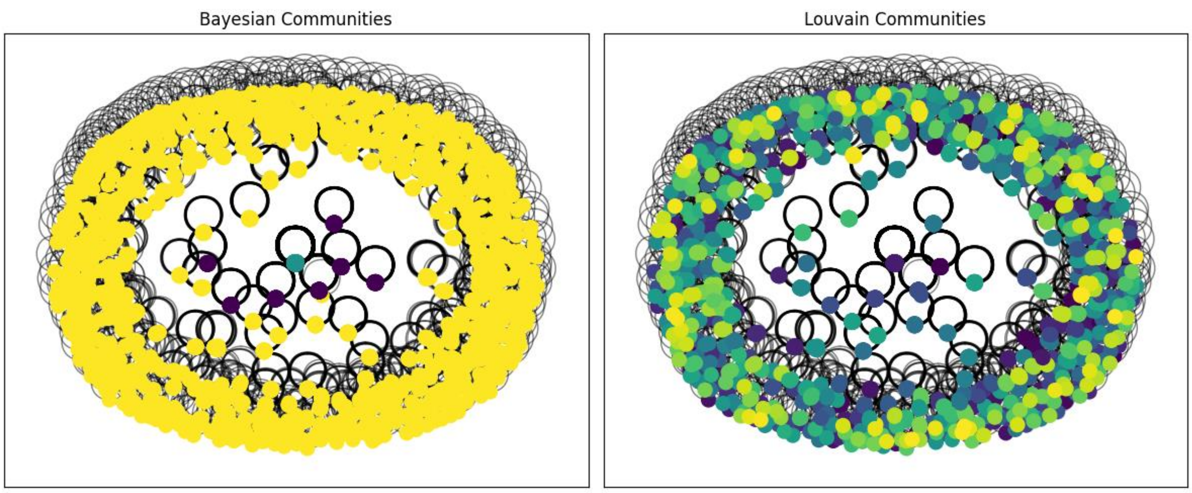
Word Cloud of Negative Tweets

Figure 2: Word Cloud of keywords based on Gender violence

- Clustering of Communities

Since social media is going to be a more significant role player in putting items on the public radar and determining what is worth talking about, now people are very interested in how one could identify communities in people tweeting about gender violence—various analyses employed different algorithms and metrics to study these networks. For example, Abel et al. identify the strength of modularity and normalised mutual information to compare community structures in gender violence discussions on Twitter. Mohsan's work points out the significance of temporal dynamics for community detection and indicates how conversations on gender violence change over time and can heavily influence community development. In addition, Gadek et al. suggest a hybrid method that blends content-based and network-based approaches for better community detection, which is essential to capture the subtleties of public opinion on gender violence. Nevertheless, there are limitations, including the possibility of bias in the selection of tweets and the difficulty in separating honest discourse on gender violence from noise, which can impact the validity of the detected communities. In general, synthesising these metrics and methodologies may improve our analysis of gender violence discussions on social media.

To fill the gaps in our research, we investigated neutral tweets involving violence concerning gender identities. Regular old community detection algorithms were utilised to analyse that, and what was discovered was that we obtained only tiny communities that were quite divided from one another. This led us to explore user interactions within these communities more deeply, explicitly using user similarity metrics, such as tweet content and retweeting behaviour—spaces typically neglected in prior work. Our modularity, betweenness, and centrality metrics analysis revealed more tenuous affinities among communities. We explored community structural aspects using metrics like silhouette scores, Calinski Harabasz indices, and Davies Bouldin scores, and we contrasted these with standard algorithms researched prior. The retweet and tweet patterns were variable across community structures, with gender violence tweets mainly originating from advocacy groups, survivors, and news agencies. Furthermore, our comparative analysis investigated the high usage patterns of various types of users.



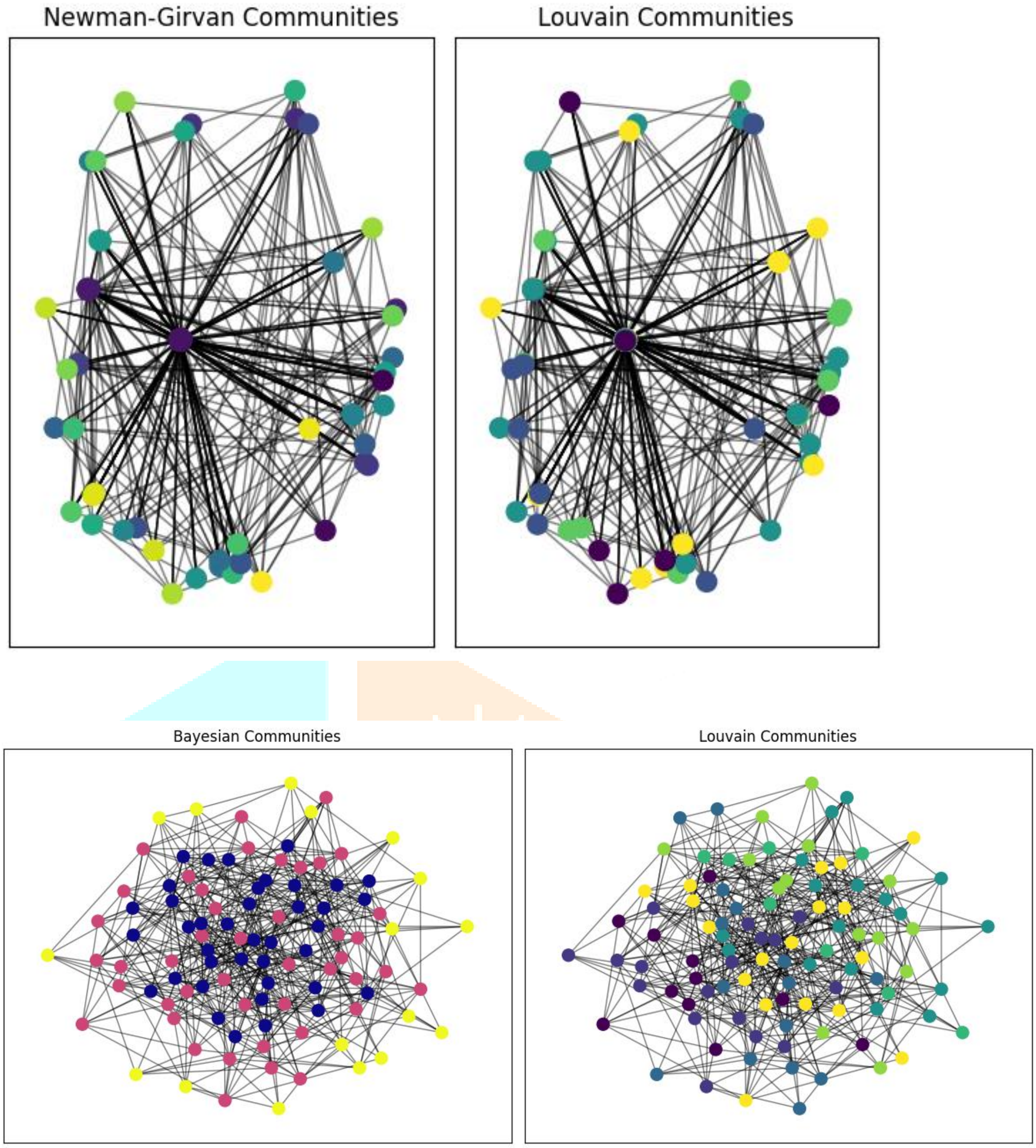


Figure 3: Accuracy Measure with Deep Learning Algorithms

Utilising the Louvain algorithm for community detection revealed 696 unique communities, achieving a modularity score of 0.1088, suggesting that these communities were delineated. The clustering metrics demonstrated remarkable outcomes:

Silhouette Score: 0.7916, signifying well-defined clusters.
Homogeneity: 0.9872, indicating that most of the cluster's members exhibit high similarity.
Completeness: 0.9872, signifying that most relevant tweets were encompassed within the identified communities.
V-Measure: 0.9872, validating the precision and reliability of community identification.
Modularity (Bayesian): 0.9003, indicating the robustness of community structure.

- **RESULTS AND ANALYSIS**

The clustering outcome showed that K-Means could efficiently group tweets into meaningful clusters, with the number of clusters optimised using the silhouette score. Community detection by Louvain identified separate user communities and the importance of key influencers in determining conversations about gender violence. Sentiment analysis showed that a high percentage of tweets had negative sentiments, expressing public concerns and frustrations about gender-based violence. The classification model was able to classify tweets into five specific types of violence, offering insights into the prevailing themes in the discourse. The research also established that most gender violence cases are underreported because of socio-cultural constraints and that Twitter discussions are subject to psychological, societal, and systemic influences. In addition, the mental health effect on women and children became a theme that recurred throughout the analysis.

- Analysis of Deep Learning Classifiers

The accuracy measurement chart shows how different machine learning classifiers perform in classifying gender violence-related tweets. Conventional ensemble-based classifiers like Extra Trees Classifier and Random Forest Classifier perform highly accurately throughout, ranging between 0.91-0.92 for categories such as Femicide and Gender Violence. Boosting classifiers like AdaBoost and Gradient Boosting also perform accurately, ranging between 0.85-0.89 for many categories. Conversely, deep learning models like CNN, LSTM, and RNN have a steep drop in accuracy, with CNN as low as 0.26-0.5 for certain classifications, indicating difficulty in coping with textual subtleties in this dataset. Between various categories, "Rape" and "Gender Violence" have higher accuracy, while "Women Violence" and "Domestic Abuse" have comparatively lower scores, between 0.71-0.75, reflecting the difficulty in differentiating these tweets. The findings indicate that tree-based classifiers perform better than neural network-based models in this case, possibly because of the structured nature of the dataset and the success of conventional feature-based classification techniques. The findings point to the need to choose the right models when examining social media discourse on sensitive issues such as gender-based violence.

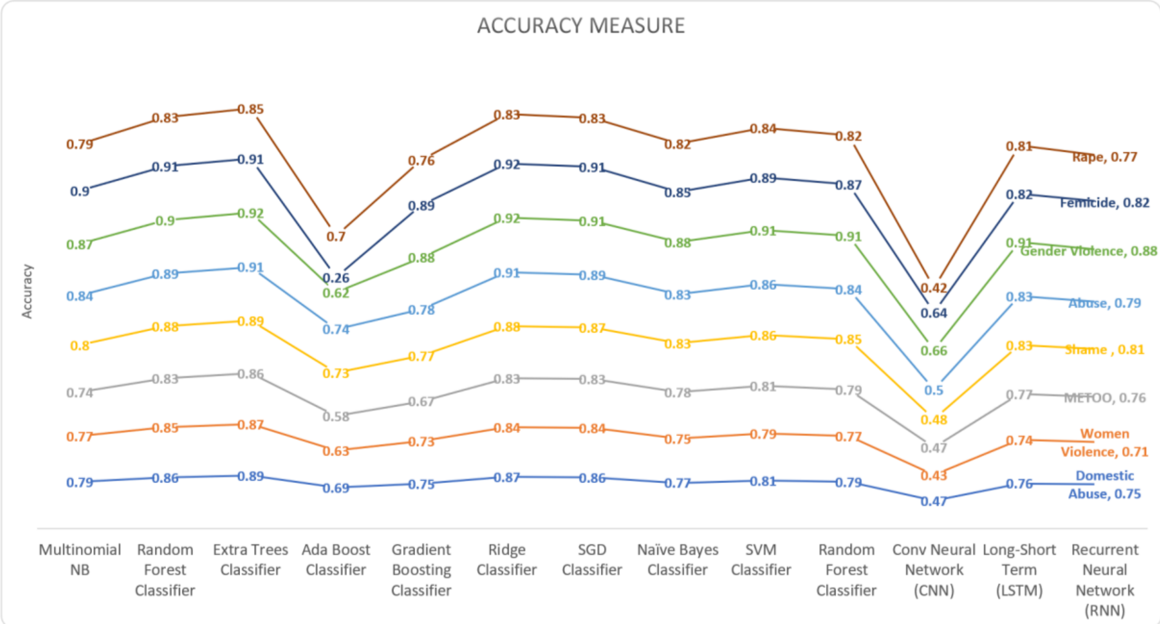


Figure 4: Accuracy Measure with Deep Learning Algorithms

CONCLUSION

This research highlights the power of social media in raising awareness and fostering discussions on gender violence. By leveraging clustering, community detection, and sentiment analysis, the study provides valuable insights into public discourse on Twitter. The results emphasise the need for greater awareness and policy interventions to address gender-based violence effectively. With growing initiatives to promote women's empowerment and gender equality, social media can be vital for advocacy and change. Future research can expand this work by incorporating multimodal data, cross-platform analysis, and real-time monitoring to enhance understanding and response strategies in addressing gender violence. By harnessing the power of technology and community, the platform will facilitate positive change in women's lives, promoting equality and inspiring women to reach their full potential.

REFERENCES

[1] Castorena, C. M., Abundez, I., Alejo, R., Granda-Gutiérrez, E. E., Rendón, E., & Villegas, O. (2021). *Deep Neural Network for Gender-Based Violence Detection on Twitter Messages*. *9*(8), 807. https://doi.org/10.3390/MATH9080807

[2] Miranda-Piña, G., Alejo, R., Rendón-Lara, E., & García, V. (2024). Detection of violent speech against women in Mexican tweets using an active learning approach. *IEEE Latin America Transactions*, *22*, 276–285. https://doi.org/10.1109/tla.2024.10473002

[3] Dhaliwal, C. K., Chandel, R. S., & Kaur, S. (2024). *Twitter Sentiment Analysis on Women Violence Using Machine Learning Techniques*. https://doi.org/10.1109/incacct61598.2024.10551076

[4] Okwach, A. A., Chen, S., & Su, C.-H. (2023). Longitudinal association of adolescents' perceptions of parental mediations and compulsive internet use. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2023.107989

[5] Soldevilla, I., Flores, N., & Tuesta, S. (2022). *Macht: An Application Based on a Sentiment Analysis Model to the Identification of Messages in Spanish with Gender Violence Content on Twitter*. 396–401. https://doi.org/10.1109/icaibd55127.2022.9820047

[6] Sridevi, U. K., & Sophia, S. (2021). *Analysing Women's Health-Related Quality of Life Using Sentiment Analysis on Social Media* (pp. 119–135). Springer, Singapore. https://doi.org/10.1007/978-981-16-4729-1_6

[7] Caruccio, L., Cirillo, S., Deufemia, V., Polese, G., Stanzione, R., & Yousafzai, M. (n.d.). *Data Analytics on Twitter for Evaluating Women Inclusion and Safety in Modern Society*.

[8] Mandava, S., Ciaverelli, I., Resnick, C., & Daniero, J. J. (2023). Characterizing Gender-Affirming Voice Intervention Discussion on Social Media: A Cross-Platform Analysis. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2023.07.022

[9] Adeeba, S., Banujan, K., & Kumara, B. (2023). *The Role of Social Media (Twitter) in Analysing Home Violence: A Machine Learning Approach*. *6*, 1–6. https://doi.org/10.1109/scse59836.2023.10215027

[10] Yoshita, D. (2024). Machine Learning for Women's Safety Based on Analysis of Tweets. *Indian Scientific Journal Of Research In Engineering And Management*. https://doi.org/10.55041/ijsrem29374

[11] Sharma, D., Gupta, V., Singh, V. K., & Pinto, D. (2024). *Should we stay silent on violence? An ensemble approach to detect violent incidents in Spanish social media texts*. 1–20. https://doi.org/10.1017/nlp.2024.25

[12] Recuero, R. (2024). The Platformization of Violence: Toward a Concept of Discursive Toxicity on Social Media. *Social Media and Society*. https://doi.org/10.1177/20563051231224264

[13] Silva, M., Anaba, U., Tulsani, N. J., Sripad, P., Walker, J., & Aisiri, A. (2023). *Gender-Based Violence Narratives in Internet-Based Conversations in Nigeria: Social Listening Study (Preprint)*. https://doi.org/10.2196/preprints.46814

[14] Miller, K., & Demirbilek, M. (2023). *The Role of Social Media in the Fight Against Gender Violence* (pp. 90–99). IGI Global. https://doi.org/10.4018/978-1-6684-8893-5.ch006

[15] Jatiningsih, O., & Siahaan, S. (2024). Sentiment Analysis of Sexual Violence on Twitter: Emotions, Public Perception, and Discourse. *Muharrik*, *7*(1), 48–72. https://doi.org/10.37680/muharrik.v7i1.4993

[16] Xue, J., Macropol, K., Jia, Y., Zhu, T., & Gelles, R. J. (2019). *Harnessing big data for social justice: Exploring violence against women-related conversations on Twitter*. *1*(3), 269–279. https://doi.org/10.1002/HBE2.160

[17] Okwach, A. A., Chen, S., & Su, C.-H. (2023). Longitudinal association of adolescents' perceptions of parental mediations and compulsive internet use. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2023.107989

[18] Khatua, A., Cambria, E., & Khatua, A. (2018). Sounds of silence breakers: exploring sexual violence on Twitter. *Advances in Social Networks Analysis and Mining*, 397–400. https://doi.org/10.1109/ASONAM.2018.8508576

[19] Reyes Gonzalez, G. A., Gabarrot, M., & Cantu-Ortiz, F. J. (2020). Understanding Violence Against Women in Digital Space from a Data Science Perspective : Full/Regular Research Papers - CSCI-ISNA. *International Conference on Computational Science*. https://doi.org/10.1109/CSCI51800.2020.00051

[20] Cardoso, C. A. L., Silva, R. N. C. de M., & Vieira, L. J. E. de S. (2023). *Ethnographic insights in Twitter posts by women in situations of violence*. https://doi.org/10.56238/seveniiimulti2023-215

[21] Gaikwad, N. P., Soumya, M., Sayyad, T. S., Vadgaonkar, P. A., & Parth, V. (2024). *Women Safety Analysis based on Twitter (Machine Learning)*. https://doi.org/10.1109/mitadtsocicon60330.2024.10575302

[22] Puente, S. N., Maceiras, S. A. D., & Romero, D. F. (2021). Twitter Activism and Ethical Witnessing: Possibilities and Challenges of Feminist Politics Against Gender-Based Violence. *Social Science Computer Review*, *39*(2), 295–311. https://doi.org/10.1177/0894439319864898

[23] Xue, J., Macropol, K., Jia, Y., Zhu, T., & Gelles, R. J. (2019). *Harnessing big data for social justice: An exploration of violence against women-related conversations on Twitter*. *1*(3), 269–279. https://doi.org/10.1002/HBE2.160