Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

Sentiment Analysis of Women's Crime News Tweets Using NLP Techniques

Shraddha Suresh Kale[1], Yashoda Raghubar Gautam[2]

[1] *Computer Science and Information Technology, Marathwada Institute of Technology, Cidco, Chhatrapati Sambhajinagar*
Email: shraddhakale190@gmail.com

[2] *Computer Science and Information Technology, Marathwada Institute of Technology, Cidco, Chhatrapati Sambhajinagar*
Email: gautamanu7411@gmail.com

Abstract – Crimes against women remain a significant social issue, and social media platforms like Twitter provide real-time insights into public opinion and emotional responses. This research focuses on analyzing

sentiments expressed in tweets related to women's crime news using Natural Language Processing (NLP) techniques. The dataset is collected from Kaggle and includes tweets related to harassment, domestic violence, and gender-based crimes. The data is preprocessed using techniques such as tokenization, stop-word removal, and text cleaning. A rule-based sentiment analysis approach is applied to classify tweets into positive, negative, and neutral categories. The results indicate that the majority of tweets express

negative sentiment, reflecting public concern, fear, and anger regarding women's safety. This study highlights the importance of sentiment analysis in understanding public perception and supports policymakers and organizations in improving safety measures.

Keywords- Sentiment Analysis, NLP, Twitter Data, Women Safety, Machine Learning

1. Introduction

Social media platforms, particularly Twitter, have become powerful tools for sharing information and expressing opinions. Crimes against women such as harassment, domestic violence, and sexual assault are widely discussed on these platforms. Unlike traditional media, Twitter captures real-time public emotions and reactions.

Sentiment analysis, a key application of NLP, helps in identifying and classifying emotions in textual data. This study aims to analyse public sentiment regarding women-related crime news using Twitter data and machine learning techniques.

2. Problem Statement

Although large volumes of tweets discuss crimes against women, there is limited structured analysis of public sentiment. Traditional reporting focuses on facts but does not capture emotional responses. The challenge is to process noisy and unstructured Twitter data to understand public opinion effectively.

3. Objectives

- To collect tweets related to crimes against women
- To preprocess and clean textual data
- To classify sentiments into positive, negative, and neutral
- To analyze public opinion and emotional trends
- To visualize sentiment patterns

4. Literature Review

Previous studies have applied sentiment analysis on social media data to understand public opinion. Research shows that machine learning techniques such as Naïve Bayes and Support Vector Machines are effective in sentiment classification. However, limited work has been done specifically on women-related crime analysis using Twitter data.

Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 04 Issue: 4 | April-2026
ISSN (O) 3107-6696

- In the paper Quantitative Sentiment Analysis of Women's Safety Using Twitter Data: An NLP Approach, researchers use Twitter data to examine public opinion on women's safety. This study highlights a gap: although large volumes of relevant social media data exist, they are often underutilized to generate actionable insights for safety and policy interventions.

- Saha, P., et al. (2019). Hate Speech and Abusive Language Detection in Social Media.

This work is relevant to women crime analysis as discussions often include abusive language, victim-blaming, and misogynistic content, which standard sentiment models may fail to capture

- Similarly, Analysis of Twitter Sentiment in Cases Of Domestic Violence: Comparison of Lexicon-Based and Naïve Bayes focuses on domestic-violence cases and compares different sentiment-analysis approaches. The findings suggest that machine-learning (Naïve Bayes, in this case) can effectively classify sentiment — implying the feasibility of automated sentiment detection in contexts of gender-based crimes.

5. Methodology

5.1 Data Collection

The dataset is obtained from Kaggle and consists of tweets related to women's safety and crime using keywords such as harassment, rape, and domestic violence.

5.2 Data Preprocessing

- Lowercasing text
- Removing URLs, hashtags, and mentions
- Tokenization
- Stop-word removal
- Stemming/Lemmatization

5.3 Model Use

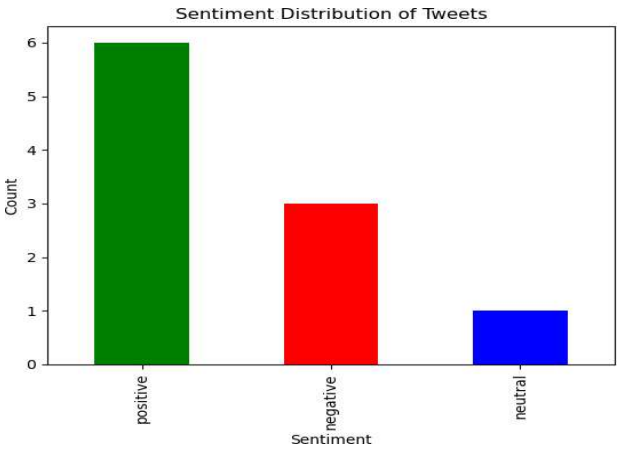
A rule-based sentiment analysis approach is used to classify tweets based on predefined positive and negative keywords.

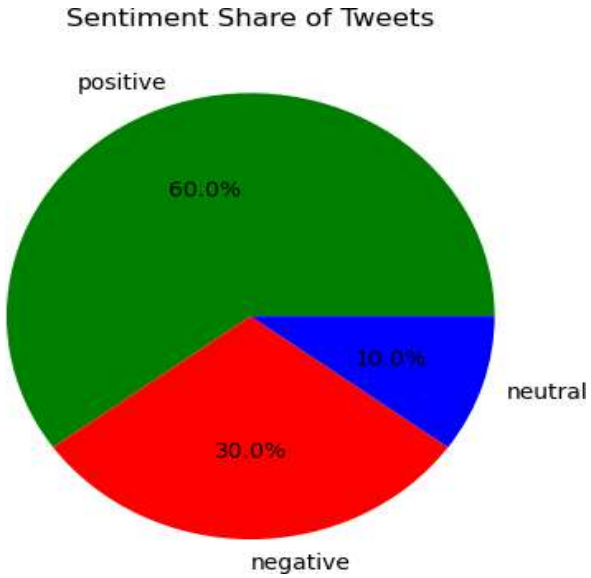
6 Tools and Technologies

- Python
- Pandas
- Matplotlib
- Jupyter Notebook

7. Results and Analysis

The analysis shows that most tweets express negative sentiment, indicating public concern, anger, and fear. Positive tweets include support and awareness messages, while neutral tweets are minimal.





8. Conclusion

This study demonstrates that sentiment analysis is effective in understanding public opinion on women's safety issues. The dominance of negative sentiment highlights the seriousness of the problem and the need for improved safety measures.

9. Future Scope

- Use deep learning models such as LSTM and BERT
- Real-time tweet analysis
- Multilingual sentiment analysis
- Integration with dashboards

10. References

1. Liu, B., Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, 2012.
2. Pak, A., & Paroubek, P., Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Proceedings of LREC, 2010.
3. Go, A., Bhayani, R., & Huang, L., Twitter Sentiment Classification Using Distant Supervision, Stanford University Technical Report, 2009.
4. Kaur, P., & Kaur, G., Sentiment Analysis on Social Media Data Using Machine Learning Techniques, International Journal of Computer Applications, 2021.
5. Bhat, S., & Sharma, R., Public Sentiment Analysis on Women Safety Issues Using Twitter Data, International Journal of Data Science Research, 2022.
6. Bird, S., Klein, E., & Loper, E., Natural Language Processing with Python, O'Reilly Media, 2009.
7. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL-HLT Conference, 2019.
8. Kaggle, Twitter Sentiment Datasets, Kaggle Database, n.d.
9. Twitter Inc., Twitter API Documentation: Developer Guide to Data Collection, 2023.
10. National Crime Records Bureau (NCRB), Crime in India Report, Government of India, 2023.