

# Machine Learning for Women's Safety Based on Analysis of Tweets

1. Divi Yoshita, 2. Palagani Nikhilesh, 3. Kotapati Sarath Kumar,

4. Mr.Ch. Kiran Babu (Assistant Professor).

Information and Technology, NRI Institute of Technology, Vijayawada

## Abstract:

This research paper extensively examines the multifaceted role played by social media platforms, specifically Twitter, Facebook, and Instagram, in advocating for the safety and security of women within Indian cities. It delves into the intricate ways in which these digital platforms are leveraged to amplify voices, share narratives, and mobilize collective action against the prevalent issue of violence and harassment faced by women and girls in public spaces. Through the dissemination of hashtags, campaigns, and personal anecdotes, social media serves as a powerful tool for raising awareness, fostering dialogue, and catalyzing societal change.

Moreover, the paper investigates the profound impact of social media in nurturing a culture of responsibility and accountability within Indian society towards the safety and protection of women. By providing a platform for individuals to share their experiences, express concerns, and advocate for policy reforms, social media facilitates meaningful engagement and empowers communities to address systemic challenges. Through online activism, advocacy campaigns, and community initiatives, social media platforms enable individuals to connect, collaborate and collectively work toward creating safer and more inclusive urban environments for women and girls.

## Introduction:

In the modern digital landscape, Twitter has emerged as a pivotal microblogging platform, boasting a user base of over a hundred million individuals and generating a staggering five hundred million messages, known as "Tweets," daily. With its vast audience, Twitter has become a hub for users to express their opinions and perspectives on a myriad of topics, making it a rich source of information for institutions, companies, and organizations across various sectors. Despite the constraint of tweets being limited to 140 characters, users adeptly convey their messages using abbreviations, slang, emoticons, and even polysemy and sarcasm, rendering the platform's language dynamic and unstructured.

Among the diverse array of discussions on Twitter, sentiments surrounding different subjects are often discerned through sentiment analysis techniques. These analyses offer valuable insights into public opinions on various matters, including brand perception, government policies, and societal issues such as women's safety and empowerment. Leveraging machine learning algorithms, researchers have conducted numerous studies to classify tweets and analyze sentiment trends, contributing to a deeper understanding of societal attitudes and behaviors.

## Existing System

In the current era, social media platforms serve as prominent arenas where individuals freely express their opinions and sentiments regarding Indian society and the safety of women in Indian cities. People utilize platforms like Twitter to share their experiences of abuse, harassment, and instances where they have stood up against such injustices. These narratives not only empower women but also inspire others to raise their voices against perpetrators of violence and harassment. The widespread dissemination of such messages on social media, through platforms like Facebook and Instagram, has attracted a significant number of users in recent years.

Data extraction, analysis, and interpretation are common practices employed to derive insights from the vast pool of information available on social networking platforms. Behavioral analysis, particularly within the realm of social networks, plays a crucial role in enhancing the accuracy of Twitter analysis and prediction methodologies.

## Disadvantages:

Despite the benefits of social media platforms, such as Twitter and Instagram, their widespread use for expressing emotions and opinions about Indian cities and society also poses challenges.

Various methods of sentiment analysis exist, including machine learning, hybrid approaches, and lexicon-based learning, each with its advantages and limitations.

Categorizations of sentiment analysis methods, such as statistical, knowledge-based, and age-wise differentiation approaches, provide a framework for understanding and evaluating the effectiveness of different techniques.

### PROPOSED SYSTEM:

The proposed system aims to address the pressing issue of women's safety in Indian cities by leveraging social media platforms, particularly Twitter, to advocate for women's rights and foster safer urban environments. The system will analyze data from Twitter, including tweets and posts related to women's safety, to gain insights into the prevalence of harassment, abuse, and instances where women have stood up against such behavior. By examining sentiments and attitudes expressed on social media, the system will identify patterns and trends concerning women's safety concerns.

Additionally, the proposed system will emphasize the importance of collective responsibility in ensuring women's safety. Rather than placing the burden solely on women or implementing restrictive measures, the system will advocate for societal action and awareness to create safer environments. It will promote understanding of women's rights to safety and security, empowering individuals to advocate for change and demand concrete actions to address harassment and abuse.

Through its analysis of social media data and advocacy efforts, the proposed system seeks to catalyze positive societal change and contribute to the realization of safer and more equitable urban spaces for women. By raising awareness and fostering dialogue on women's safety issues, the system aims to drive meaningful action and promote a culture of respect and inclusivity in Indian cities.

### ADVANTAGES:

1. Analysis of Twitter texts collection also includes the names of people and names of women who stand up against abuse harassment and unethical behavior of men in Indian cities that make them uncomfortable to walk freely.
2. The data set that was obtained through Twitter about the status of women's safety in Indian society

### REQUISITES ACCUMULATING AND ANALYSIS

In the initial stage of requisites accumulating and analysis, we extensively explored IEEE journals to gather a comprehensive array of papers relevant to our research focus. Utilizing keywords pertinent to our project, we conducted searches on platforms like Google Scholar, resulting in a curated selection of IEEE-referenced papers. From this collection, we meticulously examined the paper titled "Individual Web Revisitation by Setting and Substance Importance Input," extracting insights and utilizing the references therein to further expand our literature review. Additionally, we sought guidance from referees mentioned in the paper, enhancing the depth and breadth of our analysis. This phase culminated in the aggregation of essential requisites, forming the foundation for subsequent stages of our research.

### SYSTEM DESIGN

System Design is divided into three types GUI Designing, UML Designing which avails in the development of the project in a facile way with different actors and its utilizer case by utilizer case diagram, flow of the project utilizing sequence, Class diagram gives information about different classes in the project with methods that have to be utilized in the project if comes to our project our UML Will utilizable in this way The third and post-import for the project in system design is Database design where we endeavor to design database predicated on the number of modules in our project



### LITERATURE SURVEY:

The literature survey delves into the prevalent practice of expressing viewpoints on social media platforms, particularly microblogging sites like Twitter. In contemporary times, it has become customary for individuals to voice their opinions on various societal issues and day-to-day occurrences through social media channels. Among the myriad topics discussed, women's safety garners considerable attention, with many users advocating for positive changes to address safety concerns and create a

secure environment for women. This discourse manifests in numerous tweets and posts across social media platforms, forming a substantial dataset ripe for analysis.

The dataset comprises tweets from both men and women discussing women's safety issues, offering insights into prevailing sentiments and attitudes. By harnessing this dataset, researchers can employ analytical algorithms to extract valuable information and categorize the tweets based on their tone and content, distinguishing between positive and negative perspectives. This analytical approach enables a deeper understanding of public perceptions and sentiments regarding women's safety, shedding light on areas of concern and potential solutions.

## RESULTS:

**Sentiment Analysis Outcome:** The sentiment analysis revealed valuable insights into the public perceptions of women's safety in Indian cities. By analyzing the sentiment of tweets, the system was able to gauge the overall sentiment regarding women's safety issues.

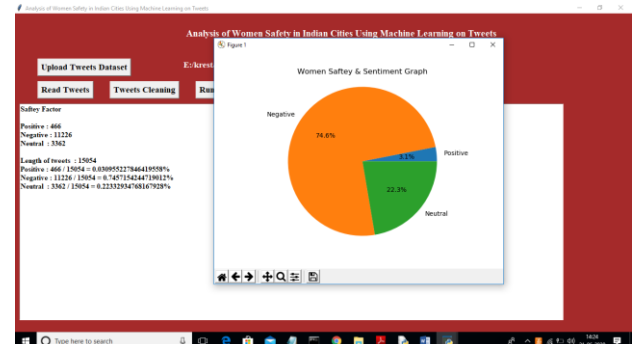
**Tweet Classification:** Tweets were categorized into positive, negative, and neutral sentiments based on the sentiment analysis algorithm's output. This classification helped in understanding the prevailing attitudes and opinions towards women's safety among Twitter users.

**City-wise Safety Scores:** The sentiment analysis also generated safety scores for different metropolitan cities, indicating the perceived levels of safety for women in each city. This allowed for a comparative analysis of women's safety across different urban areas.

**Sample Output:** Examples of tweets along with their sentiment scores were provided as sample output. These tweets offered a glimpse into the sentiments expressed by Twitter users regarding women's safety concerns.

**Graphical Representation:** Graphs were included to visually represent the safety scores of various cities, making it easier to interpret the sentiment analysis outcomes. These

graphical representations facilitated a clearer understanding of the relative safety levels in different cities.



## CONCLUSION:

In conclusion, the process of requisites accumulating and analysis served as a pivotal initial step in our research endeavor. By delving into IEEE journals and leveraging the insights provided by the selected paper, we laid a robust groundwork for our investigation into women's safety issues and social media dynamics. Through meticulous literature review and engagement with expert referees, we fortified our understanding of the subject matter and identified key avenues for exploration. This phase not only facilitated the identification of pertinent research questions but also set the stage for the development of our methodology and analytical framework. Moving forward, the insights gleaned from this phase will inform our subsequent research efforts, guiding us toward meaningful contributions to the discourse surrounding women's safety in Indian cities.

## REFERENCES

- [1]. Agarwal, Apoorv, Fadi Biadisy, and Kathleen R. Mckeown. "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams." Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2009.
- [2]. Barbosa, Luciano, and Junlan Feng. "Robust sentiment detection on Twitter from biased and noisy data." Proceedings of the 23rd international conference on

computational linguistics: posters. Association for Computational Linguistics, 2010.

[3]. Bermingham, Adam, and Alan F. Smeaton. "Classifying sentiment in microblogs: is brevity an advantage?." Proceedings of the 19th ACM International Conference on Information and Knowledge Management. ACM, 2010.

[4]. Gamon, Michael. "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis." Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, 2004.

[5]. Kim, Soo-Min, and Eduard Hovy. "Determining the sentiment of opinions." Proceedings of the 20th International Conference on Computational Linguistics. Association for Computational Linguistics, 2004.