



## **ANALYSIS OF WOMEN SAFETY IN INDIAN CITIES USING MACHINE LEARNING ON TWEETS**

**CH. RAJARATNAM, A. AMRITHA, CH. RAMYA**

Student, Department of CSE, Raghu Engineering College, Visakhapatnam, A.P., India.

**k. Priyanka**, Assistant Professor, Department of CSE, Raghu Engineering College, Visakhapatnam, A.P., India.

### **ABSTRACT**

This project centres round the job of online media in advancing the safety of girls in Indian urban communities, given the exceptional relation to the task of web based media sites or applications. Now-a-days most people are using social networking sites to precise their feelings and if any women feel unsafe in any area then she's going to express negative words in her post/tweets/messages and by analysing those messages we will detect which area is more unsafe for women's. This might be very beneficial to extend safety in dangerous areas for ladies travellers.

The main objective of our project is to investigate the assorted types of violence and threats against the ladies by making use of preferred and powerful social media data and by using the concept of machine learning. The quantity of likes, tweets, comments, blogs and post on the actual incident against woman may be used for this analysis. These Social networking sites collectively update the feedback about particular incident and it'll be exhibit under the discussion of the many people. This can give the world picture of assorted crimes against woman and showcase how the intention framed and motivation behind the scenario. This data would be helpful to safeguard the lady from the unlikely violence against them within the society.

**Keywords:** Women safety, Tweets, Sentimental Analysis, NLTK, Natural Language Processing (NLP), Machine Learning, Polarity count.

### **INTRODUCTION**

In India ladies are revered by individuals regarding them as goddesses where as there are expanding number of savagery against ladies. The brutality against ladies has expanded by numerous folds because of the more prominent openness of ladies in each field of life. Wrongdoing against ladies like assault, corrosive tossing, endowment killings, honour killings and constrained prostitution of little youngsters has been accounted for in India. The examination across most mainstream Metropolitan urban areas of India including Delhi, Bangalore and Mumbai shows that ladies feel risky while going out to work or while going out in the open vehicle and so on, true insights show a sensational expansion in the quantity of revealed wrongdoings against ladies.

Nowadays women are experiencing plenty of violence like harassment in places at several cities. This starts from stalking which then results in abusive harassment or also called abuse assault. during this project we mainly concentrate on the role of social media which may be wont to promote the security of girls in India, given the special relevancy the participation of the many social media websites or applications like Twitter, Facebook and Instagram platforms. This project motivation is to develop the responsibilities among the people on the varied parts of Indian cities in order that the protection of girls around them is ensured.

### **LITERATURE REVIEW**

Recently many investigations have been carried out by various researchers in Women safety using Machine learning. As such, few literatures are discussed below:



Apoorv Agarwal, Fadi Biadry, Kathleen R. Mckeown [1] "Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams."

- We present a classifier to predict contextual polarity of subjective phrases in a sentence. Our approach features lexical scoring derived from the Dictionary of Affect in Language (DAL) and extended through WordNet.

- A notable feature of the dictionary is that it has different scores for various inflectional forms of a word (affect and affection) and thus, morphological parsing, and the possibility of resulting errors, is avoided. Moreover, Cowie et al., (2001) showed that the three scores are uncorrelated; this implies that each of the three scores provide complementary information

[2] Luciano Barbosa, Junlan Feng. "Robust sentiment detection on twitter from biased and noisy"

- Data we propose an approach to automatically detect sentiments on Twitter messages (tweets) that explores some characteristics of how tweets are written and meta-information of the words that compose these messages.

### **EXISTING SYSTEM**

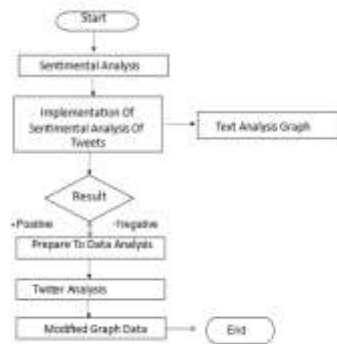
Ladies are encountering a good deal of savagery and badgering go into the open spots in numerous urban communities beginning from following to inappropriate behaviour or attack. Employment opportunities, better education, higher standard of living or simply a craze to measure within the city attract a serious part of the population to migrate to metropolitan cities. With dreadful incidents of rape, acid attacks and assault on women, the protection of girls is at an excellent risk. Women avoid wearing a selected kind of clothes, don't stay late at office, use only crowded public transportation in late hours, avoid visiting certain place at dark or preferred being among a male companion while travelling. This simply postpones the crime, rather than preventing it in terms of safety.

### **PROPOSED SYSTEM**

The proposed system introduces the model so as to eradicate harassment within the society. The key idea behind this scenario is to gather opinions and interests of individuals through multiple social networking sites, due to the actual fact that social networking sites are powerful and spread information everywhere globe during a very less time also, they contain huge data regarding social issues. Hence data regarding this idea is found through social networking where opinions, comments, likes and dislikes of individuals are found. This paper examines essential centers round the function of web-based media in advancing the safety of girls in several areas with exceptional relevancy the a part of online media sites and applications including Twitter stages. Tweets on Twitter which usually contains pictures and text and furthermore composed messages and statements which centres round the security of girls in numerous urban areas will be utilized to peruse a message among the generation and instruct individuals to create exacting move and rebuff the individuals who disturb the women. This will give the world picture of assorted crimes against woman and showcase how the intention framed and motivation behind the scenario.



Flow Chart



## IMPLEMENTATION

In this project, we will propose an idea for evaluating women's safety using social networking messages and identifying which areas are safe and which areas are risky for women using machine learning algorithms. We'll be using Python and NLTK (Natural Language Tool Kit) to download tweets concerning women's safety via MEETOO. Using NLP Algorithm And then check the tweets with Polarity value. This project might potentially be put as an add-on to social networking apps so that females can identify whether conversing with a certain individual is safe or unsafe for them. By this project, we can identify risky places in metropolitan cities and increase safety within these areas as a result of our research, allowing women to feel safe to go wherever they want.

## ALGORITHM

### NLP Algorithm:

Natural language processing (NLP) is the ability of a computer program to understand human language as it is spoken and written referred to as natural language. It is a component of artificial intelligence (AI).

### Modules:

#### Data Pre-processing:

Data pre-processing is an essential step in building a Machine Learning model and depending on how well the data has been pre-processed; the results are seen as Tokenization.

#### Data Cleaning:

Text cleaning is the process of preparing raw text for NLP (Natural Language Processing) so that machines can understand human language.

#### Polarity count:

Polarity refers to the degree of positivity or negativity in a given text. In NLP, polarity analysis is used to determine the sentiment of a text, whether it is positive, negative, or neutral.

### NLP Tools:

#### Natural Language Tool Kit(NLTK) :

The learning curve of Python is very fast and NLTK is written in Python so NLTK is also having very good learning kit. NLTK has incorporated most of the tasks like tokenization, stemming, Lemmatization, Punctuation, Character Count, and Word count. It is very elegant and easy to work with.

#### TextBlob corpora package :

TextBlob is a Python library for processing textual data. It provides a simple API for diving into common natural language processing tasks.





## CONCLUSION

In this paper, our study will present a global picture of various crimes against women and demonstrate how the scenario's goal and motivation are structured. We can identify risky areas in metropolitan cities and increase safety in those areas as a result of our project, allowing women to feel free to go wherever they want. As a result, we present recommendations in order to reduce the possibility of unsecured problems appearing in society, which will undoubtedly aid in lowering the risk of these devious actions affecting society, and by rigorously adhering to the rules, we will be able to reach a greater level of security.

## FUTURE SCOPE FOR FURTHER DEVELOPMENT

When it comes to handling enormous amounts of data from social media sites, the machine learning algorithms utilised are quite effective and perform well on a variety of platforms. These algorithms have the potential to make a significant difference in women's safety by extracting data and creating a diverse range of datasets to work with. In the near future, we intend to work on this project more and tweak it to make it even more efficient.

## REFERENCES

- TEJASHWINI, M., PRASANNA, V. L., PRAVALLIKA, S., AFROZ, S., YADAV, R. R., & DHANUNJAYA, M. Analysis of Women Safety in Indian cities using Machine Learning on Tweets.
- Reddy, P. S., Ramasubbaiah, B., & Chandra, G. R. Analysis of Women Safety in Indian Cities Using Machine Learning on Tweets.
- Riyazuddin, Y. M., Sriram, G. J., Vaibhav, P. M., & Vikranth, I. (2020). Utilization Of Support Vector Machine For Analyzing Women Safety In Indian States. *International Journal of Grid and Distributed Computing*, 13(1), 2244-2251.
- Kumar, D., & Aggarwal, S. (2019, February). Analysis of women safety in Indian cities using machine learning on tweets. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 159-162). IEEE.
- TEJASHWINI, M., PRASANNA, V. L., PRAVALLIKA, S., AFROZ, S., YADAV, R. R., & DHANUNJAYA, M. Analysis of Women Safety in Indian cities using Machine Learning on Tweets.
- Reddy, P. S., Ramasubbaiah, B., & Chandra, G. R. Analysis of Women Safety in Indian Cities Using Machine Learning on Tweets.
- Riyazuddin, Y. M., Sriram, G. J., Vaibhav, P. M., & Vikranth, I. (2020). Utilization Of Support Vector Machine For Analyzing Women Safety In Indian States. *International Journal of Grid and Distributed Computing*, 13(1), 2244-2251.
- Swapna, D., Ashrita, J., Ashwini, K., & Bhargavi, T. B. Analysis of Women Safety in Indian Cities Using Twitter Data.
- Khan, R., Siddiqui, S., & Rastogi, A. (2021). Crime Detection Using Sentiment Analysis. *Intelligent system for women's safety using data science. International Journal of Adaptive and Innovative Systems*, 3(1), 25-42.
- Kohli, P., & Singh, K. (2021).