Women Wellbeing Assessment in Indian Metropolises Using Machine Learning models

Bonthala Prabhanjan Yadav¹, M.Sheshikala ², N Swathi³, Kanegonda Ravi Chythanya⁴, Sudarshan E⁵

¹,³,⁵Sumathi Reddy Institute of Technology for Women, Warangal, India.
²,⁴SR Engineering College, Warangal, India.

¹prabhanjanyadav2020@gmail.com

Abstract: In recent times, women working in many metropolises have experienced harassment and nuisance in civic places. In this regard, the research centers taking place the responsibility of social media in endorsing care of these sensitive cases in Indian metropolises with exceptional reference to Twitter, Facebook and Instagram platforms. This paper makes special efforts to understand how the sense of responsibility in Indian society can advance the safety of these women around us.

Keywords: Tweets, Women harassment, working women, safety analysis.

1. Introduction

Some forms of abuse and aggression are so violent that they can be seen and crossed over, and these inappropriate behaviors are frequently perceived as a natural part of metropolitan life. There have been many studies in metropolises in India, and women have made akin comments about sexual harassment and other unidentified men [1–5]. A study conducted in India's largest metropolises including Pune, Mumbai, and Delhi found that 60 percent of women are feel insecure when working or traveling in public transport.

Women are entitled to the metropolises, which means they can go free wherever they want, whether it's education or any other place where women want to go. But because of the embarrassment and abuse of the very unknown eyes of these women, women feel uncomfortable at their workplaces, shopping malls. Lack of security or practical impact on women's lives is the main reason for harassing girls [6].

Too often girls are abused by their neighbors, or the lack of security when going to school is a terror in the minds of little girls, which is an example of their own struggles when it is unfair or forced and sexual harassment by their own neighbors. Safe Metropolises addresses women's health by focusing on women's rights without fear of violence or sexual harassment. It is the society's responsibility to accept the need for women's protection and to recognize that women have the same rights as men in the city, somewhat commanding on women who run the society.

An analysis of the collection of twitter scripts shows that the names of men and women in Indian metropolises are against erotic nuisance and the immoral behavior of males who are painful to walk liberally. Data set by Twitter on the security prestige of women in Indian society Developing data
analysis techniques to extract nil values and smooth data using Laplace and Porter's theory Processing and retweeting through machine learning algorithms and collecting unnecessary data from data sets to provide consistent and accurate data set [7].

2. Literature Review

Social media users openly share their views about Indian civilization and representatives who claim Indian metropolises are harmless for women [8-15] on social media. General public can express themselves on social media platforms and post their experiences wherever women are against sexual harassment or sexual harassment. Females’ data on the identical social networking platform or app, such as Facebook, have been inspired by tweets about women’s health and growing sexual harassment. Womenfolk post these posts and tweets, empowering 5 men or 10 women and voicing those who have made Indian metropolises an insecure place for womenfolk. Social media sites like Twitter, Instagram and Facebook attract large numbers of individuals and many use it to share their thoughts and ideas in Indian metropolises and Indian civilization. There are many sentimental approaches that can be classified as hybrid machine learning and lexicon-based learning [16-21]

In addition, the public has another classification presented with numerical, awareness-based and age-wise discriminatory methods. It is conjoint exercise to collect data from social networking through information withdrawal, information analysis and data interpretation techniques. Using behavioral analysis based on social networks, the accuracy and predictability of Twitter analysis can be assessed.

3. Twitter Analysis

With people actively communicating and sharing their thoughts on societal media, including Twitter and Facebook, social networking is well-thought-out as the ideal platform to get people’s thoughts and feelings about various events. There are many opinion-based systems for gathering information and analysis, which are intended to draw people's opinions on various topics. There are short messages on Twitter provided, people want to use different words and short words. Therefore, numerous researchers have used machine learning methods to capture and dig for the convergence of sentences.

4. Execution of emotional investigation of Tweets

We report in this practical document taken from the Twitter API, which is supported by Twitter. Because there is a Twitter API, there are numerous methods accessible for emotional examination of social media data. The project uses a collection of available libraries. The way to know the feelings from the tweet is given bellow:

i. Start by transferring the Emotional Vocabulary.
ii. Then take information sets for checking on twitter and augment them to the program as input.
iii. Clean up tweets by eliminating words and sounds like repeating letters from a stop.
iv. Bring every word into the token and power the word and nourish it into the program.
v. Compare separately each word with a word list of positive feelings and negative feelings, then increase positive or negative calculation of the whole expression.
vi. Lastly, based on positive and negative calculations, we obtain a percentage of the feedback result, to determine the polarization classified as positive, negative, and neutral.

The developers have analyzed various sentiments on Twitter for various devotions and instantaneous twitter sentiment examination for current events in the world such as polls, murders, movies and more. Figure 1 shows a high level of emotional analytical algorithm.

We have diverse procedures for linking to the twitter API and receiving tweets. The next step is to clean the tweet or delete the words that need to be deleted, and to categorize that tweet, which is the validity of the tweet, and finally the result.
Using the python tools, we performed sentiment examination. Some of the packages comprising Tweepie and Textblob have been used. The libraries for this task are established by the subsequent instructions:

i. pip install tweepy
ii. pip install textblob

The next step is to download the dictionary with the succeeding command:

- Python -m textblob.download_corpora.

Text blow is a natural language processing Python library, we need to feed the tweet analysis to the Corpora in a ordered set of texts / words.

4.2 Linking to the twitter API

We require making an entry on twitter and generating app or software to link to the twitter API and to enquiry the up-to-date tweets to collection in the database. We should have visited apps.twitter.com/app/new, generated the API bases needed to provide for platform. Figure 2 shows the application settings.
4.3 Results
Based on the latest 300 tweets, the resulting sample program shows the output for the query termed as 'rape'.
• Percentage of positive tweets: 16.39 percent
• Percentage of negative tweets: 72.13 percent
• Percentage of neutral tweets: 11.47 percent
The numerous tweets as a sample are taken from database are as follows:
= "@BinaNepram Sexual assaults against women # Manipur # NortheastIndia since the armed conflict began in the 1960s ~In a region where hundreds of thousands of armed forces operate with impunity under # AFSPA.
Parliament, NONE suspected of # rape disciplined in # uniform
50 years #MeTooIndia #MeToo" ="@ Stateless1 Orphans, Refugee Camp Victims Demand 'Global Security and Justice' # Refugees # Rape # Myanmar # Bangladesh
= "@ So devastating to hear prominent named figures in
The # MeToo movement After a few big names like # NanaPatekar and # AlokNath, now # KailashKher?
As per our implementation fig.2, The following graph shows the sentimental ratings.
Metropolitan wellbeing chart for womenfolk based on tweets (percentage)

Table 1: Wellbeing aspect of population

| Metropolity | Number of tweets | Wellbeing factor (tweet vs population) |
|-------------|------------------|---------------------------------------|
| Delhi       | 175,949          | 77%                                   |
| Mumbai      | 44,740           | 94.6%                                 |
| Kolkata     | 26,990           | 98.5%                                 |
| Chennai     | 17,442           | 98.7%                                 |

4.4 Out Come Information

If we execute the program at different times and with little change, we can get different results in for each one of the tweet label case from the dictionary. To test this, the program was executed for three times and the output of this test is an average of a series of these outputs.

When the unbiased tweets are huge that it means people are less interested in the topic and there is no positive / negative side to it. It should also be noted that we get very different results based on the data of the experiment, which is the public opinion based on conditions such as rape news, in the year 2017, that news has became so famous. The neutral tweet range of views is over 60 parentages. From this analysis, we conclude that Chennai is a safe city, Delhi is an insecure city.

5. Conclusion

We discussed various machine learning algorithms throughout the research paper, which enable us to manage and analyze large amounts of Twitter data that share daily tweets and text messages. Such ML algorithms are precise and convenient for analyzing large volumes of data, such as the SPC algorithm and linear factor modeling methodologies that support classify data into real categories. Support vector machines (SVM) are another form useful algorithms for collecting data from the social newtorking websites and make predictions about the wellbeing of women in Indian metropolises.

6. References

[1] Gupta, B, Negi M, Vishwakarma K., Rawat, G, & Badhani, P. (2017). Study of Twitter sentiment analysis using machine learning algorithms on Python. *International Journal of Computer Applications*, 165(9) 0975-8887

[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] Sheshikala, M., Kothandaraman, D., Vijaya Prakash, R. & Roopa, G. 2019, "Natural language processing and machine learning classifier used for detecting the author of the sentence", *International Journal of Recent Technology and Engineering*, vol. 8 no. 3 pp 936-939

[4] Mamgain, N., Mehta, E., Mittal, A., & Bhatt, G. (2016, March). Sentiment analysis of top colleges in India using Twitter data. in Computational Techniques in Information and
[5] Kumar, R.R., Reddy, M.B. & Praveen, P. 2019, "Text classification performance analysis on machine learning", International Journal of Advanced Science and Technology, vol 28 no 20 pp 691-697

[6] Ravi Kumar, R., Babu Reddy, M. & Praveen, P. 2019, "An evaluation of feature selection algorithms in machine learning", International Journal of Scientific and Technology Research vol. 8 no 12 pp 2071-2074

[7] Kumaresan, Y., Kim, H., Pak, Y., Poola, P.K., Lee, R., Lim, N., Ko, H.C., Jung, G.Y. and Dahiya, R., 2020. Omnidirectional Stretchable inorganic-Material-Based Electronics with Enhanced Performance. Advanced Electronic Materials, p.2000058.

[8] Ahamed, S.A.A., Devaraju, A. & Rao, K.V.N. 2019, "Impact of finer granules on tensile and micrograph characterization of solid welded AA2014", Materials Today: Proceedings, pp. 2688.

[9] Abdy Sayyed, M.A.H., Gupta, R. & Tanyimboh, T. 2019, "Combined flow and pressure deficit-based penalty in GA for optimal design of water distribution network", SH Journal of Hydraulic Engineering.

[10] Poola, P.K. and John, R., 2017. Label-free nanoscale characterization of red blood cell structure and dynamics using single-shot transport of intensity equation. Journal of biomedical optics, 22(10), p.106001.

[11] Agrawal, R., Jana, D., Upadhyay, R.K. & Rao, V.S.H. 2018, "Dynamic relationship between the mutual Interference and gestation delays of a hybrid tritrophic food chain model", ANZIAM Journal, vol 59 no 3 pp 370-401

[12] Ahmed, S.M. 2017, "Message from the Organizing Chair", Proceedings - 2017 International Conference on Recent Trends n Electrical, Electronics and Computing Technologies, CRTEECT 2017, vol. 2017-December, pp. xi.

[13] Ahmed, S.M., Madhuri, G., Reddy, M.S. & Condoor, S.S. 2018, "Skill development n freshmen by adopting project based learning" Introduction to engineering” course", Journal of Engineering Education Transformations, vol 2018 no Special issue.

[14] Poola, P.K., Jayaraman, V., Chaithanya, K., Rao, D. and John, R., 2018. Quantitative label-free technique for morphological evaluation of human sperm—a promising tool n semen evaluation. OSA Continuum, 1(4) pp1215-1225

[15] Kollem, S, Reddy, K R L & Rao, D S 2019, "A review of image denoising and segmentation methods based on medical images", International Journal of Machine Learning and Computing, vol 9 no 3 pp 288-295

[16] Seena Naik K and Sudarshan E 2019 Smart healthcare monitoring system using raspberry Pi on IoT platform ARPN Journal of Engineering and Applied Sciences 14(4) 872-876

[17] Ch Vinay Kumar Reddy, Dr I Rajasri, Dr Sammaiah and ARajesh 2020 Flatten The Curve And Raise The Line To Reduce The Risk Of Covid-19 Journal Of Mechanics Of Continua And Mathematical Sciences 15(8) https://doi.org/10.26782/jmcms.2020.08.00001

[18] Mahender K, Ramesh KS and Kumar TA 2017 An efficient ofdm system with reduced papr for combating multipath fading Journal of Advanced Research in Dynamical and Control Systems 9(Special issue 14) 1939-1948

[19] Sudarshan, E, Satyanarayana, C and Bindu, C.S., 2017, September. A Parallel RLE Entropy Coding Technique for DICOM Images on GPGPU. In 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC) (pp 963-966) IEEE

[20] Sudarshan E, Naik K.S, Kumar P.P 2020 Parallel approach for backward coding of wavelet trees with CUDA. ARPN Journal of Engineering and Applied Sciences 15(9), pp.1094-1100
[21] Sallauddin M, Ramesh D, Harshavardhan A, Pasha SN and Shabana 2019 A comprehensive study on traditional AI and ANN architecture International Journal of Advanced Science and Technology 28(17) 479-487