International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Predicting women's safety based on Sentimental analysis

[1]Mr. B. Rajesh Kumar, [2]Mr. C. Magesh Kumar,
[1,2] Assistant Professor,Dept. of CSE,
Malla Reddy Engineering College (Autonomous), Secunderabad, Telangana State

Abstract
This globally people are using social media platforms to share their ideas and information related to different topics every day. Twitter is one of the most popular social media platforms to send and read posts to communicate with others known as “tweets”. Peoples share their ideas, reviews, experiences, and post their opinions on a particular topic or issue. This paper aims to build a model that performs a sentimental analysis of people's opinions related to the social issues of women that is a very critical issue these days in many countries of the world. A dataset of tweets has been collected from Twitter by using a twitter scraper in python programming and then cleaned the dataset using the nltk library to remove noise from the dataset. To analyze the sentiments of people's Machine Learning tools and techniques are used. To classify each tweet as positive, negative, or neutral using Text blob in Python based on the polarity of sentiments. Datasets were collected on two women’s hashtags like #Women and #Metoo. The proposed study shows which hashtag is more popular or used by people to share experiences, opinions, and issues related to women. Different machine learning algorithms were used to train and test the model.

Keywords—Social Issues, Sentiment Analysis, Twitter, Python, Machine Learning Techniques

1. 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, honor 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 60 % of the 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.

Women are continuously harassed in our society every day. Each and every city has some parts or localities where women harassment is a major issue. The survey of metropolitan cities has uncovered that 60% of women are harassed and do not feel safe while going out of their houses. These harassments range from passing comments to body shamming which is a matter of concern for all of our society. Women while travelling via public transport feel unsafe according to the recent analysis. There are many cases in a society where women are continuously harassed in their neighborhoods, shopping malls, and on their way to their work. These issues of harassment lead to the discouragement of women class to work in a safe environment. Building a safe and harassment-free work environment for women can encourage them to work and prosper. One incident of harassment for a woman or girl can carve a lifetime bad memory and leave a scar for that woman or girl. Our society needs to approach woman safety with a perspective which will empower them to live a carefree life without having to concern with their safety and harassment.

Objectives

The main objective of this project is to explain about the importance of safety of women which leads to the growth in society and country. For this here we are using sentiment analysis by machine learning where estimation of learning may be positive or negative but it can be solved by different techniques. Web-based media contains tremendous measure of slant rich information as tweets, announcements, blog entries and so on Contrasted with general notion investigation, Twitter estimation examination is minimal troublesome because of the presence of slang words and incorrect spellings.

2. RELATED WORK

Existing System

In the previous techniques analysis of sentiments of twitter and youtube data are done and then analyze the people's opinions using machine learning tools and techniques and also using the python programming language for coding and also method of collected the tweets using the Twitter API from social media platforms such as twitter and then analyzed the sentiments of users using machine learning techniques by using different classifiers, such as Naïve Bayes, maximum Entropy and also used the binary and TF-IDF model for accuracy of sentiment analysis

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

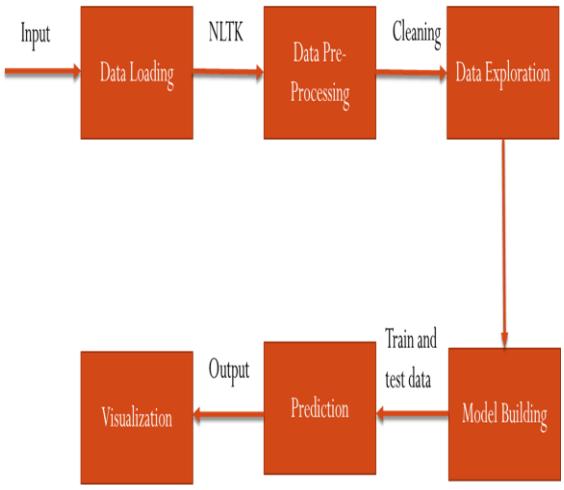
and it focus on analyzing the sentiments and subjectivity of tweets.

Proposed System

The proposed study shows which hashtag is more popular or used by people to share experiences, opinions, and issues related to women. Different machine learning algorithms were used to train and test the model in these Collection of data, pre-processing the data, extraction of features, choosing base features, detection of sentiments and classification of sentiments using machine learning approaches or simple computations are the basic steps to perform sentiment analysis.

3. IMPLEMENTATION

System architecture



Data Loading

Data loading is the process of copying and loading data or data sets from a source file, folder or application to a database or similar application. It is usually implemented by copying digital data from a source and pasting or loading the data to a data storage or processing utility.

NLTK

NLTK (Natural Language Toolkit) is a suite that contains libraries and programs for statistical language processing. It is one of the most powerful NLP libraries, which contains packages to make machines understand human language and reply to it with an appropriate response.

Data Preprocessing

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the crucial step while creating a machine learning model. In this the data is prepared and made ready for model to continue its process.

Data Exploration

Data exploration refers to the initial step in data analysis in which data analysts use data visualization and statistical techniques to describe dataset characterizations, such as size, quantity, and accuracy, in order to better understand the nature of the data.

Model Building

A machine learning model is built by learning and generalizing from training data, then applying that acquired knowledge to new data it has never seen before to make predictions and fulfill its

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

purpose. Lack of data prevents from building the model, and accessing of data is not enough.

Visualization

Data visualization is a technique that uses an array of static and interactive visuals within a specific context to help people understand and make sense of large amounts of data. The data is often displayed in a story format that visualizes patterns, trends and correlations that may otherwise go unnoticed.

4. EXPERIMENTAL RESULTS

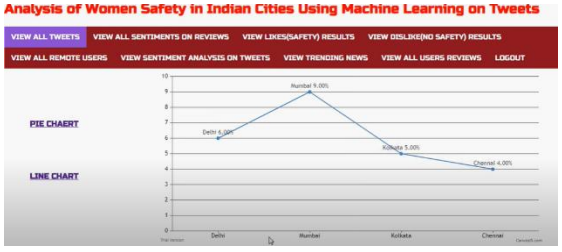


Fig:2 Result on line graph

Fig:3 Result on pie graph

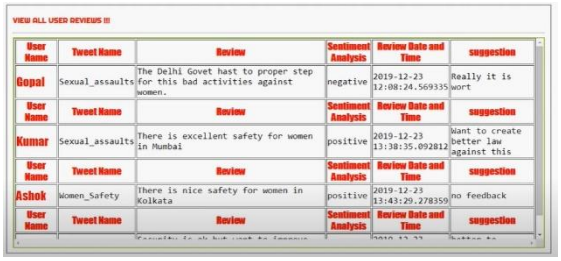


Fig:-4 Tweet

5. CONCLUSION

The primary target is to zero in on how an awareness of others' expectations on piece of Indian culture can be created with the commoners so we should zero in on the security of ladies encompassing them. For the twitter information that incorporates a large number of tweet and messages each day, AI calculation assists with getting sorted out and perform investigation. Straight mathematical factor model and SPC calculation are a portion of the calculations which are compelling in dissecting the huge information that give classification and convert them into important datasets, Future EnhancementThis concept of integrating collected information from multiple social networking sites and deriving strategic graphical form on necessary concepts like child labour, child abuse, eve teasing which are showing their adverse effects on society thereby provide guidelines so as to lessen such worse scenarios prevailing in the society.

6. REFERENCES

[1] Agarwal, Apoorv, FadiBiadsy, 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.

[6] Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau RJ. Sentiment analysis of twitter data. InProceedings of the workshop on language in social media (LSM 2011) 2011 Jun (pp. 30-38).

[7] Liu B. Sentiment analysis and opinion mining. Synthesis lectures on human language technologies. 2012 May 22;5(1):1-67.

[8] Reyes-Menendez A, Saura JR, Alvarez-Alonso C. Understanding# WorldEnvironmentDay user opinions in Twitter: A topic-based sentiment analysis approach. International journal of environmental research and public health. 2018 Nov;15(11):2537.

[9] Bakshi RK, Kaur N, Kaur R, Kaur G. Opinion mining and sentiment analysis. In2016 3rd international conference on computing for sustainable global development (INDIACom) 2016 Mar 16 (pp. 452-455). IEEE.

[10] Bakshi RK, Kaur N, Kaur R, Kaur G. Opinion mining and sentiment analysis. In2016 3rd international conference on computing for sustainable global development (INDIACom) 2016 Mar 16 (pp. 452-455). IEEE.