International Journal of Advanced Research
in Arts, Science, Engineering & Management

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 6.551

Sentiment Classification on Women's Safety across Indian Cities using Natural Language Processing and Machine Learning

Mrs. Preethi M[1], Varun[2], Dishashree[3], J Treshika[4], Kajal Kumari[5]

Asst. Professor, Department of Computer Science & Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Karnataka, India[1]

Student, Department of Computer Science & Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Karnataka, India[2,3,4,5]

ABSTRACT: Nowadays women are experiencing lots of violence such as harassment in places in several cities. This starts from stalking which then leads to abusive harassment or also called abuse assault. In this paper, we mainly focus on the role of social media which can be used to promote women's safety in India, given more preference to the participation of many social media websites or applications such as Twitter, Facebook, and Instagram platforms. The proposed work focuses on developing the responsibilities among the common people in various parts of Indian cities so that the safety of women around them is ensured. Tweet on the Twitter application contains text messages, audio data, video data, images, smiley expressions, and hashtags. This tweet content can be used to read among the people and thus can educate them to take strict actions if tweets are abusive to women and hence can punish such people if the harassment is made. Applications that include hashtags such as MeToo has been considered for the study. Machine learning algorithms such as Naive Bayes, Random Forest, Support Vector Machine, and Logistic regression are analyzed.

I.INTRODUCTION

There are some very aggressive patterns of harassment and violence, including staring and commenting, and these unacceptable behaviors are often considered part of city life. There have been a number of studies held in Indian cities and women reporting similar patterns of sexual related harassment and comments made by others they do not know. Research conducted in the cities which are more famous in India which shows that more than half of the percentage of women do not feel safe when commuting to work or when traveling by any means

of transport. Any cities can be visited by women which means they are free to go whenever they want, whether it's at the Institute of Education or wherever the women want to go. But they are not safe in places such as shopping malls which they pass by on their way to work due to many strange eyes. for the harassment of girls. There are times when the harassment of girls was done by their neighbors while they were going to the school. Or the safety of girls where less which created a mindset of fear in the thinking of small girls. who suffered throughout their life due to that one incident that happened in their past life where they were forced to do something unacceptable. Or was harassed by their neighbors or anyone whom they don't even know. Safest cities idea approaches to women's safety by considering the rights of women which may affect the town from getting afraid of violence or harassment. Instead of controlling the actions of women, the mentality of the society can be changed in order to protect women. Analysis of Twitter data also includes the people names and the name of women who are against unusual or bad behavior of people in cities of India which make them uncomfortable to walk freely. Machine learning algorithms were applied on the twitter dataset to smooth the data by removing zero values and using appropriate theory to develop a method of data analysis and delete retweet and repeated data from the obtained dataset. Which gives us the clear view of women's safety status in Indian society. And we can also find the safest cities of the country where there is more safety to the women.

A. Abbreviations

SVM - Support Vector Machine.

RF - Random Forest.

LR - Logistic regression.

B. sentiment analysis

Sentiment analysis meant for extracting data from social media, analyzing using Natural language processing (NLP), statistics or machine learning techniques to identify the sentiment of extracted text unit. The sentiment can be positive, negative or neutral. Sentiment analysis is used in brand monitoring, analyzing customer satisfaction to improve the branding, market research analysis, etc. Sentiment analysis can be many types such as text summarization, recommendations, social sentiment analysis etc. However, analysis of social media streams is always restricted and basic analysis can only be done using the extracted texts. The following figure shows the general architecture of sentiment analysis from social media, considered here is twitter data. The following figure shows the models/ layers used for analysis.

II. LITERATURE REVIEW

Parikh et al. [1] performed twitter sentiment analysis using unigram Naïve bayes model. The author extracted twitter content using Java twitter API and they implemented using Java technologies. First they underwent pre-processing module, then applied unigram naïve bayes for sentiment analysis. They implemented multinomial unigram and Bernoulli unigram model and they evaluated performance of the model using F1 score measures. However their model achieved good results on analysis, they considered only small dataset. Their implementation lags in training large amount of data.

The authors from [2] analyzed sentiment classification using distant supervision model. They applied automatically classification of sentiment of Twitter messages, they are classified as positive or negative based on the query term. The authors proposed machine learning algorithms for classifying the sentiment of Twitter messages using distant supervision. They considered the training data with emoticons, as noisy labels. They showed machine learning algorithms namely Naive Bayes, Maximum Entropy, and SVM achieved accuracy more than 80% which is trained with emoticon. The authors also proved that pre-processing improves accuracy of sentiment classification.

The author in [3] demonstrated that training data labeled with emoticons has much potential to be independent of domain, topic and time. The author constructed model for topic dependency, domain dependency and temporal dependency using naïve Bayes and SVM classification algorithm. The author constructed framework for emoticon corpus and emoticon trained classification using Naïve Bayes and SVM. However, their model shows the performance of dependency model, they have not considered much pre-processing. Some noisy data reduced the performance of their proposed model.

Barbosa et al [4] considered the study of sentiment analysis of Biased and Noisy twitter Data. The author classified twitter data to positive, negative and neutral. They extracted the feature using Part of speech (POS) tags. They analyzed three different source of twitter data namely Twendz, twitterSentiment and TweetFeel. However, their results showed good performance of classification on Bias and noisy data, there are certain limitations available such as sentences that contain antagonistic sentiments.

The authors in [5] studied entity based sentiment analysis on twitter dataset. They used Named Entity Recognition (NER) model and used unigram model for sentiment analysis. The author represented the classified data as sentiment cloud for different entities such as name Harry potter, organization Microsoft etc. Though their work represented some

interesting facts, this model cannot be applied on all the contexts. Thus in our proposed model, we applied the feature extraction techniques.

Agarwa et al. [6] proposed sentiment analysis on twitter dataset using POS specific polarity features. Instead of going for the known feature extraction model, they applied tree kernel and POS for feature extraction. The author investigated two kinds of models namely tree kernel and feature based models and experiment both these models and concluded that they outperform the unigram baseline. However the accuracy are very less compared to our model, they achieved around 60% accuracy.
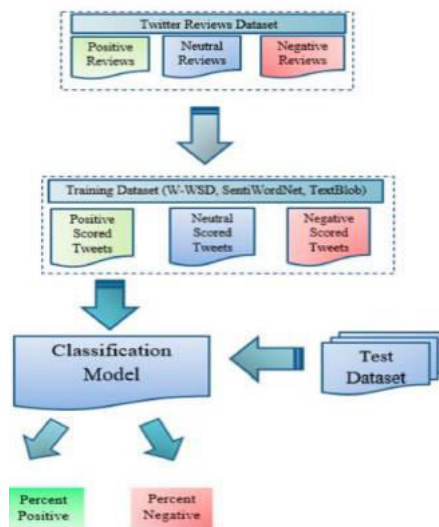

*Figure 1. **Sentiment Analysis***

The above figure is the overall representation and flow of Twitter sentiment Analysis. Sentiment analysis can be termed as the process of mining of attitudes, opinions, emotions and emotions from text, tweets and database using Natural Language Processing (NLP). Sentiment analysis classify opinions into categories namely "positive" or "negative" or "neutral". It's also termed as subjectivity analysis, opinion mining, and appraisal extraction. The terminologies namely opinion, sentiment, view and belief are used frequently in this study there are differences between them. Opinion refers to conclusion open to dispute, views refer to subjective opinion, whereas belief refers to acceptance and intellectual assent and finally sentiment refers to opinion representing individual's feelings.

III. IMPLEMENTATION DETAILS

This work is done using python with suitable libraries. We downloaded the dataset from twitter.com which consist of separate ram as well as test suite. train.csv is considered a training set and test.csv is considered a test set. Algorithms such as SVM, logistic regression and random forest are applied. Implementation steps The proposed application should be able to identify the sentiment of the tweet text. Feature extraction models used are Bag of words and TF/IDF. We used random forest, Logistic regression, and SVM classification models to predict the sentiment type.

1. Extract the feature using count vectorization.
2. Apply machine learning models.
3. Split train and test set.
4. On the test set, apply the machine learning algorithm Random Forest, Logistic regression, and SVM.
5. Predict the sentiment types.
6. Compare the machine learning algorithm's accuracy on each feature extraction model.

IV. APPROACHES

A. Support Vector machine

Among the supervised machine learning techniques, the most widely used technique is Support vector machine which is used for both regression and classification problems. The algorithm works in such a way that each datum is plotted as a point in n-dimensional space with characteristic values representing the values of each coordinate.
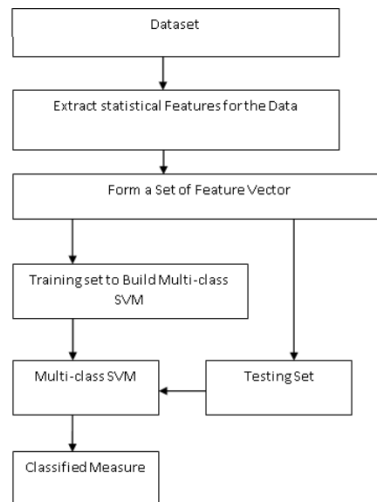

*Figure 2. **Flow chart of SVM algorithm***

B. Logistic Regression

In simple linear regression, the score of one variable is predicted from the other variable's score. The criterion is the variable which is predicted and is also called as Y. Predictor is the base variable which is used for prediction and is also known as X. Prediction method using only one predictor is known as simple regression. In simple linear regression, Straight line is formed when prediction of Y is expressed in terms of X.

Benefits:
1. Faster.
2. Installation is not needed.
3. Very easy to understand.
4. Understood.

Inconveniences:
1. Does not have the ability to give accurate prediction.
2. The relationship between response and features are assumed to be linear.

If the non-linear relationship is as high as in many situations, the linear relationship will not model the relationship efficiently and its predictions will be inaccurate.
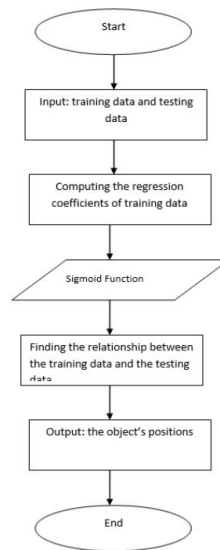
Figure 3. **Flow chart of Logistic Regression algorithm**

Relevance to current Research

C. Random Forest Model
1. Suppose the training dataset contains 'n' number of cases. and subsamples are randomly selected with substitutions. Individual trees are constructed using random subsamples taken from the training dataset.

2. By Assuming k input variables, some d is chosen such that d should be less than k. d variables are randomly chosen from k variables at each node. From d variables, the best split is selected for division of node. The value of m is maintained throughout the growth of the forest.
3. Each tree is planted as tall as possible without pruning.
4. The prediction of class of the new object is on the basis of majority of votes obtained from the combination of all decision trees.

It works on bayes theorem and is a simple family of "probabilistic classifiers" with strong (naive) independence assumptions between features.
It assumes that presence of one feature is independent of other. Naive Bayes models are easy to build and are especially useful for very large data sets. Naive Bayes is simple algorithm that performs better than most complex classifications methods.
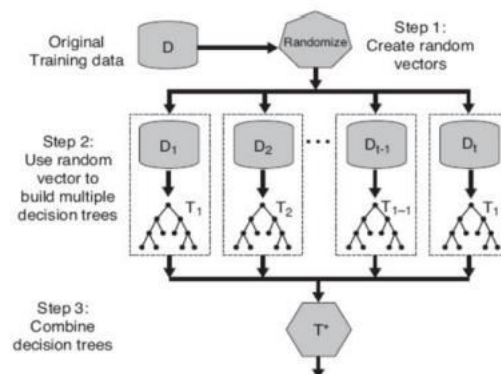

Figure: Flow chart of Random Forest

Figure 4. Flow chart of Random Forest

D. Naive Bayes

$$P(c \mid x) = \frac{P(x \mid c) P(c)}{P(x)}$$

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

P(c|x) is the posterior probability of the class (c, target) with the predictor (x, attribute).

P(c) is the prior probability of the class. P(x|c) is the probability of the predictor for a given class. P(x) is the prior probability of the predictor

V. CONCLUSION AND FUTURE WORK

Machine learning algorithm has been discussed throughout the project. For the twitter data that includes millions of tweet and messages every day, machine learning algorithm helps to organize and perform analysis. Logistic regression, Naive Bayes algorithm, SVM and Random forest are some of the algorithms which are effective in analyzing the large data that provide categorization and convert into meaningful datasets. We have done front end for the project where anyone can use it. Hence we can perform machine learning algorithms to achieve sentimental analysis and bring more safety to women by spreading the awareness.

For the future enhancement, we can extend to apply these machine learning algorithms on different social media platforms like facebook and instagram also since in our project only twitter is considered. Present ideology which is proposed can be integrated with the twitter application interface to reach larger extent and apply sentimental analysis on millions of tweet to provide more safety. Mainly we can add an extra feature in the website. When tourists search for a state it should show how much safe the state is. So that they can take some measures.

REFERENCES

[1] Apoorva Agarwal, Fadi Biadsy, 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] Luciano Barbosa 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] Adam Bermingham 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] Michael Gamon. "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] Soo-Min Kim and Eduard Hovy. "Determining the sentiment of opinions." Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics, 2004.

[6] Dan Klein and Christopher D. Manning. "Accurate unlexicalized parsing." Proceedings of the 41st Annual Meeting on Association for Computational LinguisticsVolume 1. Association for Computational Linguistics, 2003.

[7] Eugene Charniak and Mark Johnson. "Coarse-to-fine nbest parsing and MaxEnt discriminative reranking."

Proceedings of the 43rd annual meeting of the association for computational linguistics. Association for Computational Linguistics, 2005.

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

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)