

# A Survey on Sentiment Analysis based Social Unrest Detection

Prof.P.M. Tekade  
Information Technology  
Department of Engineering,  
Rajarshi Shahu College of  
Engineering, Tathawade, Pune

Naresh Shinde  
B.E (IT)  
Rajarshi Shahu College of  
Engineering, Tathawade, Pune

Yash Vyawahare  
B.E (IT)  
Rajarshi Shahu College of  
Engineering, Tathawade, Pune

Yash Kotalwar  
B.E (IT)  
Rajarshi Shahu College of  
Engineering, Tathawade, Pune

Ruturaj Nilawar  
B.E (IT)  
Rajarshi Shahu College of  
Engineering, Tathawade, Pune

## ABSTRACT

Social Unrest refers to public dissatisfaction towards topics, objects, or politics, and with the current population, they are varied and hard to detect. As such social unrest detection has become a new topic for research, and in recent years many methodologies have emerged; in this paper, we survey these various methodologies to help us in finding better ways for social unrest detection.

### Keywords

Sentiment Analysis, Machine Learning, Bag-of-words, Neural Networks, LSTM, Social Unrest

## 1. INTRODUCTION

Civil disorder, also known as a civil disturbance, civil unrest, or social unrest, is a situation arising from a group act of civil disobedience in which law enforcement has difficulty maintaining their authority[1]. Social Unrest mainly occurs due to public dissatisfaction with objects, events, or topics. Over the years, numerous studies have been organized on Social Unrest Detection.

It has been observed that most of these studies have one common technique used in them, i.e., Sentiment Analysis. Sentiment analysis is the analysis of people's views, attitudes, opinions, and emotions towards a subject or object [2]. Sentiment Analysis (SA) or Opinion Mining (OM) is the computational study of people's opinions, attitudes, and emotions toward an entity. The entity can represent individuals, events, or topics[3]. We have conducted a survey of various such study papers and different types of Sentiment Analysis

techniques to help outline the progress in this field and the current most efficient and accurate methods.

### 1.1 Sentiment Analysis

Sentiment analysis is the usage of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information[4]. Sentiment analysis is an approach to natural language processing (NLP) that identifies the emotional tone behind a body of text. This is an in-demand way for organizations to determine and classify opinions about a product, service, or idea. It implies the use of data mining, machine learning (ML), and artificial intelligence (AI) to mine text for sentiment and subjective information.

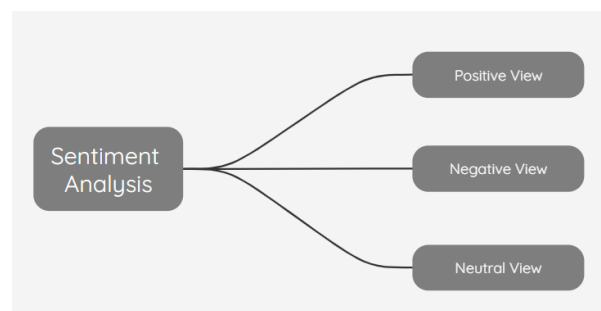


Figure 1. Types of Sentiment Analysis

Since nowadays humans express their thinking and feelings more openly, sentiment analysis is fast becoming an important tool to monitor and understand the sentiment in all types of data.

Analyzing Feedback, such as opinions in surveys and social media conversations, allows us to

understand what makes users happy or frustrated so that they can tailor products and services to meet their user's needs.

For example, using sentiment analysis, it will analyze its own 4,000+ responses in user satisfaction. The survey could help you discover why users are happy or unhappy.

## 1.2 Benefits of Sentiment Analysis-

### 1.2.1 Sorting Data at Scale

It's very difficult to sort manually by sorting thousands of tweets, user support conversations, and surveys. There's just much business data to process manually. Sentiment analysis helps businesses process vast amounts of unregulated data in an efficient and cost-effective way.

### 1.2.2 Real-Time Analysis

Sentiment analysis can identify critical problems in real-time. For example, Wearable Smart Watch provides real-time data like a person's heart rate. Sentiment analysis models can assist you in immediately identifying if any serious situations occur, so you can take action right away.

### 1.2.3 Consistent criteria

It's estimated that people only agree around 70-75% of the time when determining the sentiment of a particular text. Tagging text by sentiment is highly emotional and influenced by personal experiences, thinking, and trust. By using a centralized sentiment analysis system, companies can apply the same criteria to all of their data, helping them for better accuracy and gain better insights. Therefore the application of sentiment analysis is endless.

## 2. LITERATURE SURVEY

In [5], Elliot Mbunge (et al.) uses fuzzy search-based sentiment analysis on public sentiment towards the Zimbabwe election and finds out the general trend of public opinion toward the election period of 2017-2018.

In [6], the authors Saadati D (Et al.) used an SVM-based sentiment analysis model using techniques like bag-of-words, etc., on a Twitter-based data relation to protests. The paper concluded that there is high potential in predicting the onset of social unrest.

In [7], Benkhelifa E (Et al.) created a cloud-based system that uses data from multiple social media platforms to predict social and civil unrest. They designed a framework that relies on social networks to achieve the goals.

In [8], machine learning algorithms like clustering methods were used. They have concluded that the

classification approach to collecting scholarly articles targets the topics and subtopics of our research scope.

In [9], the authors use K-means. And they compared their methods with three different OSLOM, CPM, and hub percolation overlapping community detection methods.

In [10], authors use machine learning algorithms like LSTM and CNN. In this, they have concluded all the latest neural network-based approaches in terms of accuracy and model complexity.

In [11], they have used Adaptive Deep Recurrent Neural Network to overcome issues like misspelling, special characters, etc., and developed a real-time sentiment and visualization analysis for the social media platform Twitter.

In [12], the authors Buscald (Et al) used a SA based approach on tweets about the genoa floodings to extract information that can be helpful for disaster management.

## 3. DETECTION TECHNIQUES

### 3.1 Logistic Regression

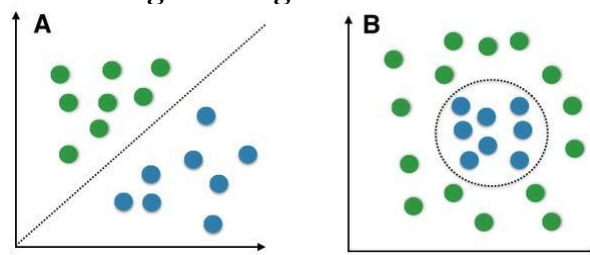


Figure 2. Logistic Regression

Logistic regression is a supervised learning technique. It will give the deterministic value between 0 and 1. Generally, it is used for classification problems. The assumption for the logistic regression is that the independent variable should not have multicollinearity. The tweets can then be classified into positive, negative, or neutral. Logistic regression act as or is used as a classifier and it has unigram as a feature vector.

### 3.2 Decision Tree

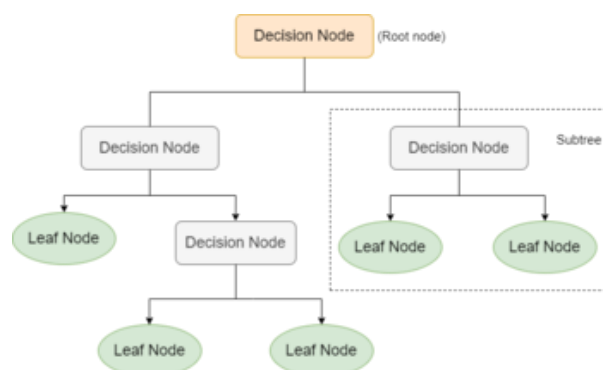
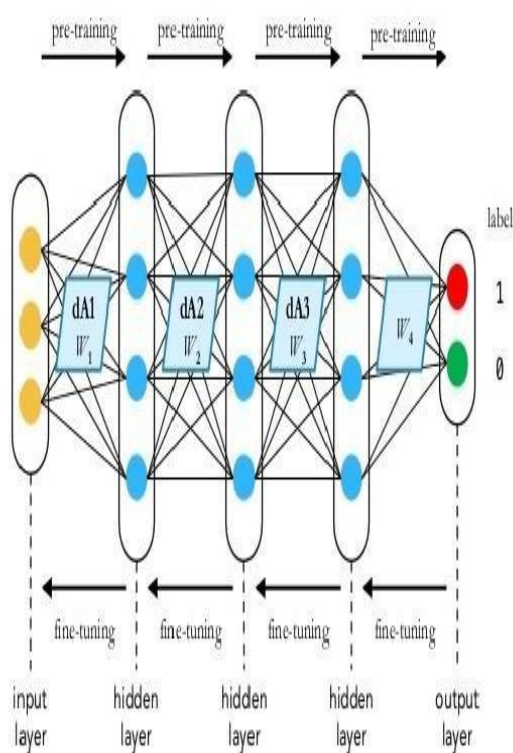


Figure 4. Decision Tree

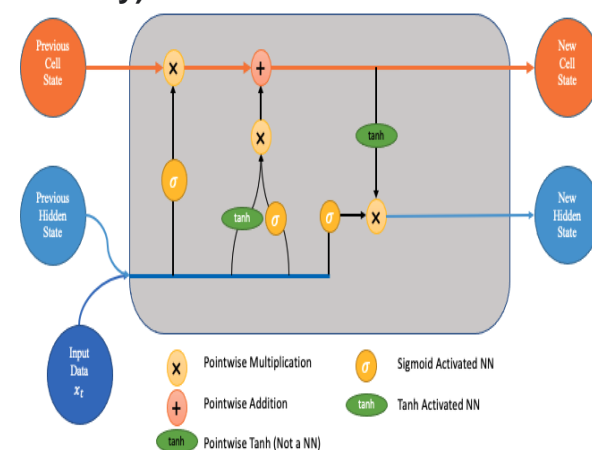
It comes under supervised machine learning, and this tool is mainly used for classification and counting the prediction. Using a decision tree, we can make predictions and take the decisions on previous data. Text mining techniques, stem, classification, and others are used to build classification and analysis of sentiment analysis.

### 3.3 CNN (Convolutional Neural Networks)



CNN is a Convolutional Neural Network and these are basically used in the analysis of images, it detects the objects and recognizes them. In CNN we have three layers, I. Convolutional Layer II. Pooling Layer III. Fully-Connected Layer. Analysis of each object is done using a support vector machine classifier. In image processing it is conventionally designed to process pixel data.

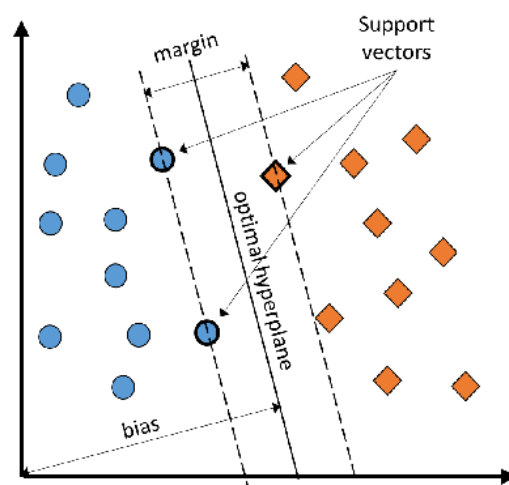
### 3.4 LSTM (Long Short-Term Memory)



**Figure 6. LSTM**

It is a type of recurrent neural network capable of sequence prediction. It is a type of complex area in deep learning. As a functional simulation of the behavior of human brains and one of the most successful deep neural models for sequential data, long short-term memory (LSTM) networks are excellent in learning implicit knowledge from data. LSTM network also generates an output at each time step and this output is used to train the network using gradient descent.

### 3.5 Support Vector Machine (SVM)



SVM is a supervised machine learning algorithm and using this technique we can do classification, regression and outlier detections. The distance between two support vectors is called the margin. The quantitative

#### 4. CONCLUSION

By using Machine Learning social unrest and global issues can be identified in a faster and more accurate manner. We believe that by using this framework much meaningful insights can be gathered for social analysis and threat detection not purchase.

#### 5. REFERENCES

- [1] "Field Manual No. 19-15: Civil Disturbances" (PDF). United States Army. Retrieved 3 February 2018.
- [2] Medhat Walaa, Hassan Ahmed, and Korashy Hoda, "Sentiment analysis algorithms and applications:A survey," Ain Shams Engineering Journal, pp. 1093– 1113, 2014.
- [3] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey", 2022. .
- [4] Sentiment analysis - Wikipedia. (2022). Retrieved 10 April 2022, from [https://en.wikipedia.org/wiki/Sentiment\\_analysis](https://en.wikipedia.org/wiki/Sentiment_analysis)
- [5] Mbunge, Elliot, Fanwell Vheremu, and Kinsley Kajiva. "A tool to predict the possibility of social unrest using sentiment analysis-case of Zimbabwe politics 2017-2018." *Int J Sci Res* 6.10 (2020).
- [6] Saadati, Dan, Farah Uraizee, and Tariq Patanam. "Social Unrest: Classification and Modeling, 229." (2016).
- [7] Benkhelifa, Elhadj, et al. "Exploiting social networks for the prediction of social and civil unrest: A cloud based framework." *2014 international conference on future internet of things and cloud*. IEEE, 2014.
- [8] Bayhaqy, Achmad, et al. "Sentiment analysis about E-commerce from tweets using decision tree, K-nearest neighbor, and naïve bayes." *2018 international conference on orange technologies (ICOT)*. IEEE, 2018.
- [9] Varghese, Raisa, and M. Jayasree. "Aspect based sentiment analysis using support vector machine classifier." *2013 international conference on advances in computing, communications and informatics (ICACCI)*. IEEE, 2013.
- [10] A. Dey, "Attention Based LSTM CNN Framework for Sentiment Extraction from Bengali Texts," 2020 11th International Conference on Electrical and Computer Engineering (ICECE), 2020, pp. 226-229, doi: 10.1109/ICECE51571.2020.9393107.
- [11] Kavitha, P. "Twitter Sentiment Analysis Based On Adaptive Deep Recurrent Neural Network." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 12.9 (2021): 2449-2457.
- [12] Buscaldi, Davide, and Irazú Hernandez-Farias. "Sentiment analysis on microblogs for natural disasters management: a study on the 2014 genoa floodings." *Proceedings of the 24th international conference on world wide web*. 2015.