

December 2021

## A Review on Sentiment Analysis techniques and applications

Eshaa Gogia

*Bhilai Institute of Technology, Durg, eshaagogia@gmail.com*

Follow this and additional works at: <https://www.interscience.in/gret>



Part of the [Other Computer Engineering Commons](#)

---

### Recommended Citation

Gogia, Eshaa (2021) "A Review on Sentiment Analysis techniques and applications," *Graduate Research in Engineering and Technology (GRET)*: Vol. 1 : Iss. 3 , Article 4.

DOI: 10.47893/GRET.2021.1041

Available at: <https://www.interscience.in/gret/vol1/iss3/4>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in Graduate Research in Engineering and Technology (GRET) by an authorized editor of Interscience Research Network. For more information, please contact [sritampatnaik@gmail.com](mailto:sritampatnaik@gmail.com).

---

## A Review on Sentiment Analysis techniques and applications

### Cover Page Footnote

ACKNOWLEDGEMENT Right inspiration and motivation play an integral role in making a success story and the journey of the work worthwhile. I would like to acknowledge and give my sincere gratitude to my co-author Prof. Shankha De for playing a key role in assisting me throughout my research and writing work. His tireless support helped me throughout all stages of my work. I would like to pay my warmest token of thanks to Prof. Saurabh Singh as without his immense guidance and motivation it would have been impossible to carry out my research project. Additionally, I would like to express my gratitude to my entire college Bhilai Institute of Technology for believing in me and providing me the opportunity to prepare this article. Last but not the least, I am highly obliged to my friends and family for their valuable supervision and suggestions.

December 2021

## A Review on Sentiment Analysis techniques and applications

Eshaa Gogia

*Bhilai Institute of Technology, Durg, eshaagogia@gmail.com*

Follow this and additional works at: <https://www.interscience.in/gret>



Part of the [Other Computer Engineering Commons](#)

---

### Recommended Citation

Gogia, Eshaa (2021) "A Review on Sentiment Analysis techniques and applications," *Graduate Research in Engineering and Technology (GRET)*: Vol. 1 : Iss. 3 , Article 4.

Available at: <https://www.interscience.in/gret/vol1/iss3/4>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in Graduate Research in Engineering and Technology (GRET) by an authorized editor of Interscience Research Network. For more information, please contact [sritampatnaik@gmail.com](mailto:sritampatnaik@gmail.com).

---

## A Review on Sentiment Analysis techniques and applications

### Cover Page Footnote

ACKNOWLEDGEMENT Right inspiration and motivation play an integral role in making a success story and the journey of the work worthwhile. I would like to acknowledge and give my sincere gratitude to my co-author Prof. Shankha De for playing a key role in assisting me throughout my research and writing work. His tireless support helped me throughout all stages of my work. I would like to pay my warmest token of thanks to Prof. Saurabh Singh as without his immense guidance and motivation it would have been impossible to carry out my research project. Additionally, I would like to express my gratitude to my entire college Bhilai Institute of Technology for believing in me and providing me the opportunity to prepare this article. Last but not the least, I am highly obliged to my friends and family for their valuable supervision and suggestions.

---

This article is available in Graduate Research in Engineering and Technology (GRET): <https://www.interscience.in/gret/vol1/iss3/4>

# A Review on Sentiment Analysis techniques and applications

Eshaa Gogia\* Shankha De\*

[eshaagogia@outlook.com](mailto:eshaagogia@outlook.com), [shankhada2009@gmail.com](mailto:shankhada2009@gmail.com)

Department of Computer Science Engineering, Bhilai Institute of Technology Durg, Chhattisgarh, India

## I. INTRODUCTION

The concept of sentiment analysis has been for quite some time. The main idea is to leverage a wider audience and analyze their expressions and emotions. From extracting insights from customer behavior for better performance to analyzing the mindsets and consequent emotions of masses sums up the entire motive of the process of sentiment analysis. Although quantification of such large data at disposal and its subsequent refinement can be difficult but with implementation of various algorithms and regression techniques, the obtained metrics and insights can prove to be beneficial not only to a specific organization but also to individuals to study patterns in human behavior. The more the number of people, more are the data points. Sentiment analysis is also termed as opinion mining as these data points serve as a vast dataset of opinions<sup>1</sup>. In this paper, we will have a brief review about the sentiment analyzer model and discuss a modified approach for the same.

## II. APPLICATIONS

Sentiment analysis makes it easier to understand and analyze unstructured data and provides deep insights about human behavior and psychoanalytic perspective. Companies have mountains of customer feedback collected and when it comes to summarizing the feedback to obtain certain actionable insights, sentiment analysis comes into play. It is commonly used as a part of business analysis where the customer behavior and satisfactory levels are continuously studied and examined<sup>1</sup>. In some cases it may

happen that larger amount of useful data comes from social media handles like Twitter, Facebook and Instagram where people are actively involved in sharing their reviews and experiences about certain products and services. Ratings and reaction comments are also a potential source of opinion mining and expression analysis. Analyzing sentiments behind certain forum posts or policy proposition polls give a better idea for strategizing and planning. In addition to this, social media has a wide growing audience expressing their thoughts and opinions on various societal issues, trends and also sharing their personal experiences in the feed as well<sup>2</sup>. Here too, sentiment analysis can play a vital role in monitoring and having a broader overview of attitudes and mindsets of not only specific targeted audience but also for large masses as well. Different social media posts can be analyzed with various motives such as enhanced neurological study, psychological researches, calculation of happiness and other emotion rates of a particular region or country and a wider opinion mining<sup>3</sup>.

## TECHNIQUES FOR SENTIMENT ANALYSIS –

Machine learning is a subset of AI and sentiment analysis makes use of both of these applications without being explicitly programmed<sup>3</sup>. In addition to these, making use of natural language processing and computational linguistics also help in the overall mechanism and ease of machine training. With a thorough study and analysis of certain set of expressions, machine learning techniques can be used to assign weighted sentiment scores to the given set of

entities. There are various approaches for sentiment analysis, some of them are listed below:

- **Rule based sentiment analysis**  
A group of words are combined to form a dictionary, which contains various words depicting different sentiments. A set of rules are followed to mitigate and make use of these words in a well-defined manner. A large number of clauses, phrases and expressions are made out of these words which impart different tone for sentiments like sarcasm, negations etc. These are further rated on a scale of 1-10 with some pre-defined remarks.
- **Machine learning based sentiment analyzer**  
A well-defined training set can be used as labels to train a ML model to analyze various opinions and sentiments<sup>3</sup>. This approach depends largely on the type of algorithm and the quality of the training data used. The training data has proved to be a great help to gauge the polarity of sentiments in the sentences<sup>4</sup>. Advancements in deep learning architecture have even come up with sentiment neurons which extensively make use of Natural Language Processing (NLP) to train an improved dataset for efficient performance.

### III. ALGORITHMS USED

Along with the training dataset of words, various supervised learning algorithms are used to classify and analyze the data. Some of these algorithms are discussed below:

- **NAIVE BAYES CLASSIFIER**  
It is one of the simplest and most commonly used classifier which uses the concept of **posterior**

**probability**. It also uses the Bayes theorem to predict the probability that a given feature set belongs to a particular label. Although all features are independent but Naive's classifier suggests the usage of prior establishment of certain set of words.

$$P(\text{label} | \text{features}) = \frac{P(\text{label}) * P(\text{features} | \text{label})}{P(\text{features})}$$

The accuracy of the Naïve Bayes Classifier can be improved by using the feature selection method.

#### **Feature Selection Method –**

This method primarily uses two approaches including Lexicon Analysis which makes use of human annotations ranging from root words to larger strings and Statistical Analysis which is usually automated making the process of indexing easier and faster.

- **LINEAR REGRESSION**

Linear regression calculates the scalar relationship between variables on both the axes. Taking words and phrases on the X coordinate and sentimental output on the Y coordinate axis helps derive the scalar relationship between the two variables and calculates the polarity of the obtained output with great accuracy.

- **SUPPORT VECTOR MACHINES**

Support Vector Machines (SVM) can be termed as an advanced version of Linear Regression. These are supervised learning models working on the principle of linear separator or best hyperplanes for better understanding and modelling of the data. It works more towards accurate analysis than prediction taking into consideration opinion

subjectivity as well as expresser's credibility <sup>5</sup>.

language pair with small amount of training set <sup>6</sup>.

- **MAXIMUM ENTROPY CLASSIFIER**

Also known as Conditional Exponential Classifier, it uses the process of encoding to convert labels to vectors which are then analyzed based on their assigned weight scores. Apart from these features it also detects parallel sentences between different

$$P(fs | label) = \frac{\text{dotprod}(weights, \text{encode}(fs, label))}{\sum(\text{dotprod}(weights, \text{encode}(fs, l)) \text{ for } l \text{ in labels})}$$

#### IV. PROPOSED METHODOLOGY

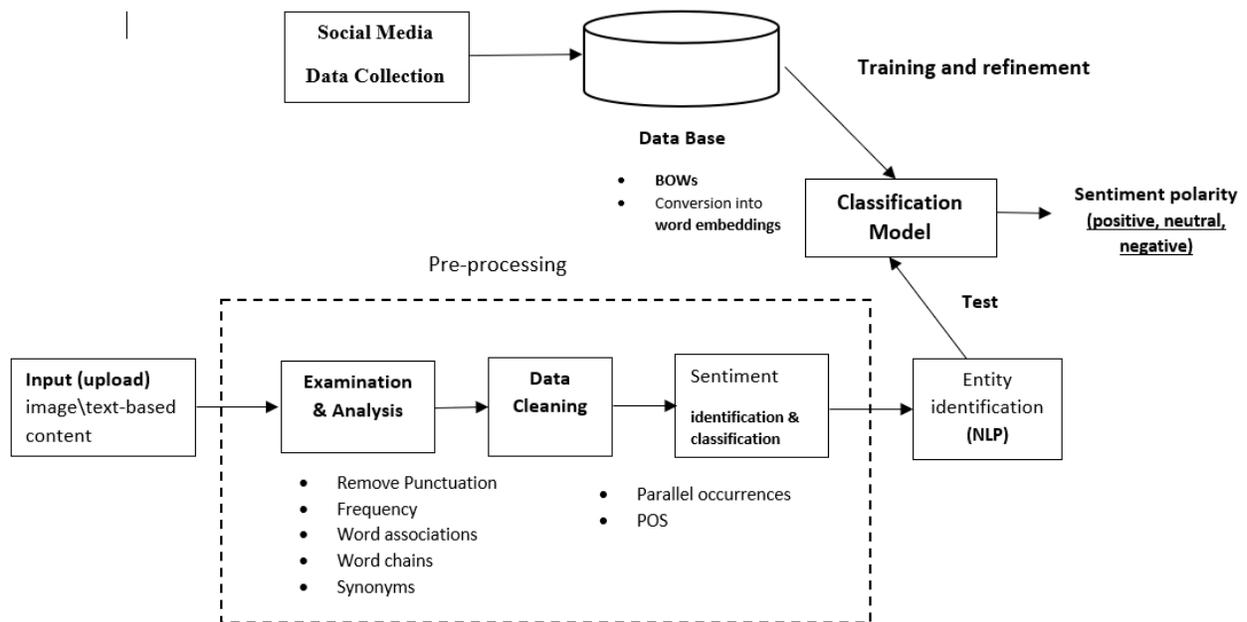


FIGURE .1 Sentiment Analyzer

The existing model of sentiment analyzer can be slightly modified for better user experience. Following is a sequence diagram depicting a revised model with incorporation of few changes:

- Collecting vast data from various social media handles which may be in the form of text-based captions, comments, reactions, image-based posts, tweets, reviews and status.
- This collected data can serve as a potential dataset when analyzed and classified into various major and minor categories of sentiments, emotions and opinions. Tokenization can be the next step.
- Based on this categorical data, a large Bag of Words (BOWs) can be created with thousands of words depicting different emotions, sentiments, moods and opinions.
- Each word from the BOWs are converted into word embeddings which are nothing but weighted vectors or vectors with specified dimensions.
- The semantics of the word are embedded across the dimensions of the vector <sup>6</sup>.
- Words depicting nearly same meaning must be having semantically similar vectors <sup>7</sup>.
- Words having similar associations, word chains, parallel occurrences, synonyms and POS are identified.
- Based on the requirement, various classifier and algorithms will be used.  
A custom algorithm can be made with accuracy higher than Naïve Bayes Classifier with added features of Feature Selection Method and Support Vector Machines.
- On the platform of Sentiment Analyzer, user or analyst may upload any text or image-based post from social media handles.
- The selected upload will be thoroughly examined and analyzed.

- Term's presence, frequency, associations and semantics of each word of the upload will be detected and identified.
- Sentiment or emotion behind the respective words as well the entire uploaded file will be identified with BOWs and NLP <sup>8</sup>.
- Subsequent sentiments will be classified for various sub-categories (if any).
- Sentiment polarity will be checked and displayed after the entire mechanism <sup>9</sup>.
- The analyst or user may opt for a scalar analysis of the previous 3 uploads for better understanding of the mindset.

## V. CONCLUSION

A proper study and analysis of various emotions and sentiments people express is important to understand various mindsets and opinions. Introducing the combined features of various classifiers would help improve the overall performance. This online Sentiment Analyzer will act as a virtual door to people, rendering its higher precision features to analyze the closest possible emotions behind their social media feeds.

## VI. REFERENCES

- [1] Medhat. Walaa, Ahmed. Hassan, Hoda. Korashy (2014). "Sentiment analysis algorithms and applications: A survey". In the proceedings of the Ain Shams engineering journal, 5 (2014/12/1).
- [2] Dumbleton. Rob (2020), Article on "Complete guide to sentiment analysis", <https://getthematic.com/insights/sentiment-analysis/> [ Accessed: Oct.10, 2020].
- [3] Cambria. Erik (2016). "Affective Computing and Sentiment Analysis" in the proceedings of the IEEE Intelligent Systems (Volume: 31, Issue: 2, Mar-Apr.2016).

- [4] Badjatiya, Pinkesh, Gupta. Shashank, Gupta. Manish, Varma. Vasudeva (2017). “Deep learning for hate speech detection in tweets”. In proceedings of International World Wide Web Conference Committee (IW3C2) (June 1<sup>st</sup>,2017).
- [5] Mozetič. I, Grčar. M, Smailović. J (2016). “Multilingual Twitter Sentiment Classification: The Role of Human Annotators”. In the proceedings of the PLoS ONE 11(5): e0155036. <https://doi.org/10.1371/journal.pone.0155036>.
- [6] Maas. Andrew, E Daly. Raymond, T Pham. Peter, Huang. Dan, Y Ng. Andrew, Potts. Christopher (2011). “Learning word vectors for sentiment analysis”. In the proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies (2011/6).
- [7] Nasukawa. Tetsuya, Yi. Jeonghee (2003). “Sentiment analysis: capturing favorability using natural language processing”. In the proceedings of the 2<sup>nd</sup> international conference on Knowledge capture (2003/10).
- [8] Whitelaw. Casey, Garg. Navendu, Argamon. Shlomo (2005). “Using appraisal groups for sentiment analysis”. In the proceedings of the 14<sup>th</sup> ACM international conference on information and knowledge (2005/10).
- [9] Lin. Chenghua, He. Yulan (2009). “Joint sentiment/topic model for sentiment analysis”. In the proceedings of the 18<sup>th</sup> ACM conference on Information and knowledge management (2009/11).
- [10] Genc. Ozgur (2019), Article on, “The basics of NLP and real time sentiment analysis with open source tools”, <https://www.morganclaypool.com/doi/abs/10.2200/s00416ed1v01y201204hlt016> [Accessed: Oct.7,2020].