



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Special Issue 2, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

SENTEX- The Modern Sentiment Analysis

Mrs A Jayanthi ¹, Akshaya Baalaji S², Blessington J V ³, Dhinesh Kumar K⁴

Asst. Professor, Dept of Computer Science Engineering, Velammal Engineering College, Chennai,
Tamil Nadu, India¹

U.G Scholar, Dept of Computer Science and Engineering, Velammal Engineering College, Chennai,
Tamil Nadu, India^{2,3,4}

ABSTRACT : Sentiment analysis, also known as opinion mining, is a powerful tool that enables us to extract valuable insights from text data. By analyzing the opinions, emotions, and attitudes expressed in written language, we can better understand people's perspectives and make informed decisions. To perform sentiment analysis, we first need to preprocess the text by removing unwanted information and normalizing the words. This helps to ensure that our analysis is based on relevant features and attributes. Next, we need to identify the features or attributes that are most relevant to the sentiment analysis. This can be done using various techniques, such as bag of words, n-grams, or word embeddings. By extracting these features, we can build a model that can accurately classify the sentiment of the text as positive, negative, or neutral. To achieve this, we can use a variety of machine learning algorithms, such as Naive Bayes, Support Vector Machines (SVM), or neural networks. These algorithms are trained on a labeled dataset of text data with known sentiments, which enables them to learn how to accurately classify new text. Finally, we evaluate the performance of our model using various metrics, such as accuracy, precision, recall, and F1 score. By doing so, we can ensure that our model is accurate and reliable, and can be used to make informed decisions based on the sentiment of the text. Overall, sentiment analysis is a powerful tool that can help us gain valuable insights from text data and make more informed decisions based on people's opinions, emotions, and attitudes.

KEYWORDS: Bag of words, n-grams, or word embeddings, Naive Bayes, Support Vector Machines (SVM), or neural networks.

I. INTRODUCTION

Sentiment analysis is a cutting-edge application of natural language processing that is revolutionizing the way we understand and interact with textual data. With the explosion of social media and other online platforms, there is an enormous amount of data being generated every day that contains valuable insights into people's attitudes, opinions, and emotions. Sentiment analysis provides a way to extract this information and use it to gain a deeper understanding of the world around us. The process of sentiment analysis involves several stages, including data collection, text preprocessing, feature extraction, and machine learning modeling. Each of these stages requires careful consideration and attention to detail, as even small errors or oversights can lead to inaccurate results. One of the biggest challenges in sentiment analysis is dealing with the complex and nuanced nature of human language, which can often be ambiguous, contradictory, or difficult to interpret. Despite these challenges, sentiment analysis has a wide range of applications in many different fields. For example, it can be used to track public opinion on political issues, monitor brand reputation, or identify potential customer service issues before they escalate. It can also be used to identify trends and patterns in large datasets, providing valuable insights that can inform business strategy and decision-making. As technology continues to evolve, sentiment analysis is becoming increasingly sophisticated and accurate, with new tools and techniques being developed all the time. This is opening up new possibilities for its use, and there is no doubt that sentiment analysis will continue to play a crucial role in the future of natural language processing and data analysis.

II . RELATED WORK

In [4], "Opinion mining and sentiment analysis" is a seminal paper published in 2008 by Bing Liu and Lillian Lee, in which they provide a comprehensive survey of the field of sentiment analysis. The paper covers the history of sentiment analysis, its challenges and limitations, the techniques and algorithms used in sentiment analysis, and its applications in various fields, such as marketing, politics, and customer service. It is considered a foundational work in the field of sentiment analysis and has been widely cited in subsequent research. The accuracy of the existing system is about 70% .

In [7]. "A Deep Learning Approach to Sentiment Analysis in Twitter" by Tasnim Mohiuddin and Md. Ashraf Islam (2019) - This paper proposes a deep learning approach for sentiment analysis on Twitter data using Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) has an accuracy of 89.9% on the SemEval-2017 Task 4 dataset, which is a benchmark dataset for sentiment analysis on Twitter. They also report a comparative analysis with other state-of-the-art approaches, showing that their proposed approach outperforms them in terms of accuracy.. We have provided a solution with more accuracy than the existing system does not have a faster way of getting insight from customer data .The ability to act on customer suggestions.The accuracy and insightful of customer perceptions and feedback is low . The work in this paper is divided into two stages. 1) Tokenization 2) Sentence Splitter 3) SentiWord Interpretation. Based on the sentiment analysis technique used, the text is classified into a positive, negative, or neutral sentiment. Some techniques may also assign a sentiment score to each text, indicating the intensity of the sentiment. Visualization tools such as word clouds, sentiment histograms, and sentiment heat maps can be used to present the results in an easily interpretable format .

III. METHODOLOGY

There are **5 steps** to analyze sentiment data and here's the graphical representation of the methodology to do the same.



Fig 1. Methodology for sentiment analysis

Data Collection :Data collection is the process of gathering relevant information or data from various sources in order to use it for analysis and decision-making. Data collection is a crucial step in the data science process as the quality and quantity of data collected can greatly impact the accuracy and reliability of the analysis results. Consumers express opinions on various platforms using different languages and slang, making the data overwhelming for businesses. Advanced technologies like natural language processing and sentiment analysis can help businesses understand and improve their offerings.

Text Preparation :Text preparation for sentiment analysis involves cleaning, tokenization, POS tagging, stemming/lemmatization, sentiment labeling, and feature extraction. The goal is to transform unstructured textual data into a structured format for analysis using machine learning techniques. Quality text preparation is crucial for accurate sentiment analysis results. It is nothing but filtering the extracted data before analysis. It includes identifying and eliminating non-textual content and content that is irrelevant to the area of study from the data.

Sentiment Detection :Sentiment detection is a vital component of sentiment analysis that involves identifying the overall sentiment expressed in text. This is accomplished using machine learning algorithms, such as Naive Bayes or Deep Learning models, that are trained on labeled datasets to predict the sentiment of new, unlabeled text. To perform sentiment detection, text must be preprocessed, including text cleaning, tokenization, and sentiment labeling. Overall, sentiment detection is essential for understanding the sentiment of textual data and can provide valuable insights for businesses and organizations. At this stage, each sentence of the review and opinion is examined for subjectivity. Sentiment analysis distinguishes subjective expressions from objective ones and applies computational techniques like Unigrams, lemmas, and negation to analyze sentiment at different levels.

Sentiment Classification :Sentiment classification is a crucial aspect of sentiment analysis that involves categorizing text as positive, negative, or neutral. This is achieved through machine learning algorithms like Naive Bayes or Deep Learning models, using preprocessed text with features extracted using techniques like bag-of-words or TF-IDF vectorization. Sentiment classification provides valuable insights into customer opinions and brand perception. Sentiments can be broadly classified into two groups, positive and negative. At this stage of sentiment analysis methodology, each subjective sentence detected is classified into groups-positive, negative, good, bad, like, dislike.



Top emojis for Positive :

[😂, 😄, 😁, 😊, 😋, 😌, 😍, 😎, 😏, 😐, 😑, 😒, 😓, 😔, 😕, 😖, 😗, 😘, 😙, 😚, 😛, 😜, 😝, 😞, 😟, 😠, 😡, 😢, 😣, 😤, 😥, 😦, 😧, 😨, 😩, 😪, 😫, 😬, 😭, 😮, 😯, 😰, 😱, 😲, 😳, 😴, 😵, 😶, 😷, 😸, 😹, 😺, 😻, 😼, 😽, 😾, 😿, 🙀, 🙈, 🙉, 🙊, 🙋, 🙌, 🙍, 🙎, 🙏]

Top emojis for Neutral :

[😏, 😄, 😁, 😊, 😋, 😌, 😍, 😎, 😏, 😐, 😑, 😒, 😓, 😔, 😕, 😖, 😗, 😘, 😙, 😚, 😛, 😜, 😝, 😞, 😟, 😠, 😡, 😢, 😣, 😤, 😥, 😦, 😧, 😨, 😩, 😪, 😫, 😬, 😭, 😮, 😯, 😰, 😱, 😲, 😳, 😴, 😵, 😶, 😷, 😸, 😹, 😺, 😻, 😼, 😽, 😾, 😿, 🙀, 🙈, 🙉, 🙊, 🙋, 🙌, 🙍, 🙎, 🙏]

Top emojis for Irrelevant :

[😏, 😄, 😁, 😊, 😋, 😌, 😍, 😎, 😏, 😐, 😑, 😒, 😓, 😔, 😕, 😖, 😗, 😘, 😙, 😚, 😛, 😜, 😝, 😞, 😟, 😠, 😡, 😢, 😣, 😤, 😥, 😦, 😧, 😨, 😩, 😪, 😫, 😬, 😭, 😮, 😯, 😰, 😱, 😲, 😳, 😴, 😵, 😶, 😷, 😸, 😹, 😺, 😻, 😼, 😽, 😾, 😿, 🙀, 🙈, 🙉, 🙊, 🙋, 🙌, 🙍, 🙎, 🙏]

Top emojis for Negative :

[😏, 😄, 😁, 😊, 😋, 😌, 😍, 😎, 😏, 😐, 😑, 😒, 😓, 😔, 😕, 😖, 😗, 😘, 😙, 😚, 😛, 😜, 😝, 😞, 😟, 😠, 😡, 😢, 😣, 😤, 😥, 😦, 😧, 😨, 😩, 😪, 😫, 😬, 😭, 😮, 😯, 😰, 😱, 😲, 😳, 😴, 😵, 😶, 😷, 😸, 😹, 😺, 😻, 😼, 😽, 😾, 😿, 🙀, 🙈, 🙉, 🙊, 🙋, 🙌, 🙍, 🙎, 🙏]

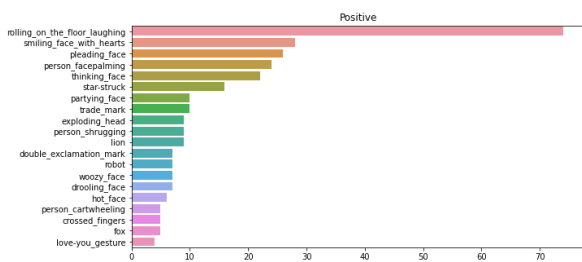


Fig 5.1 .Positive word cloud

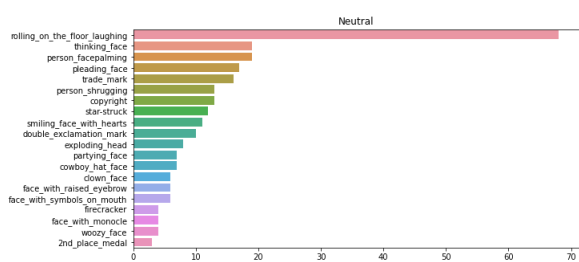


Fig 5.2 .Neutral word cloud

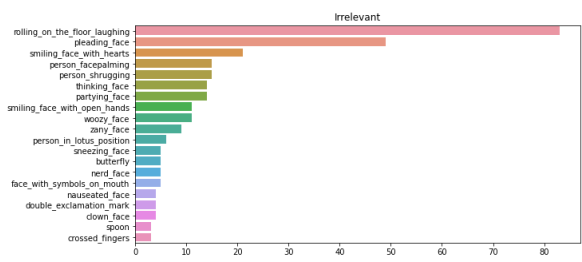


Fig 5.3 .Irrelevant word cloud

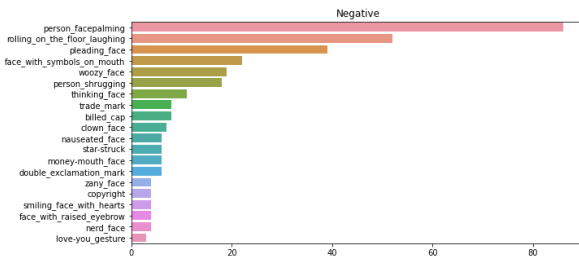


Fig 5.4 .Negative word cloud

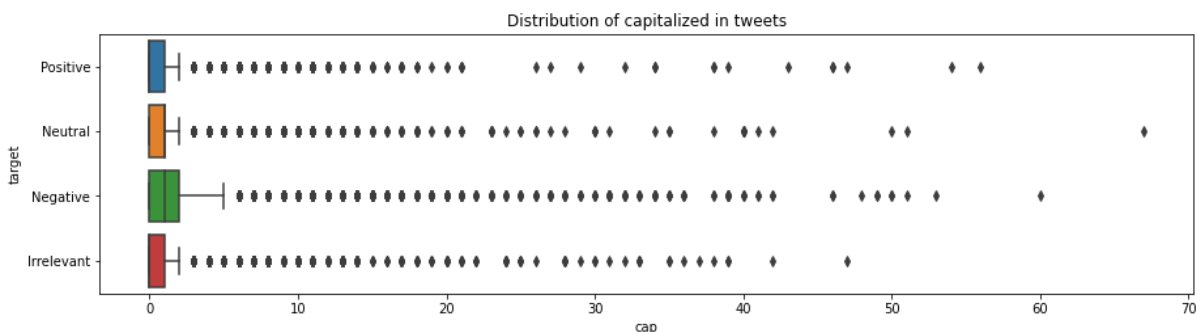


Fig 6 . Comparison of Sentiment

V. CONCLUSION

On concluding , sentiment analysis is a valuable technique for understanding the opinions and emotions expressed in text data. It can be used in a variety of applications, including social media monitoring, market research, customer feedback analysis, and more. However, it is important to note that sentiment analysis is not a perfect tool and has its limitations. It relies heavily on the quality of the data and the accuracy of the algorithms used. It also struggles with sarcasm, irony, and other forms of figurative language that can be difficult to interpret . Overall, sentiment analysis can provide valuable insights when used appropriately and with careful consideration of its limitations. Sentiment analysis has Strong commercial interest because Companies want to know how their products are being reviewed by customers

and also Prospective of consumers want to know what existing users think. We have seen that Sentiment Analysis can be used for analyzing opinions in blogs, articles, Product reviews, Social Media websites, Movie-review websites and Chatbot chats where a third person narrates his views. Our analysis revealed that the majority of tweets were positive in sentiment, with users expressing satisfaction and enthusiasm towards the entity or topic. However, a notable proportion of tweets were negative, with users expressing frustration, disappointment, or criticism. By examining the most frequently used words and phrases in the tweets, we were able to identify key topics and themes that were associated with positive and negative sentiment. This information can be used to improve the entity or topic in question and address areas of concern. Overall, our sentiment analysis provides valuable insights into the attitudes and opinions of Twitter users towards a particular entity or topic. The information gathered can be used to inform decision-making and strategy development, and to improve the entity or topic in question. We have also studied NLP and Machine Learning approaches for Sentiment Analysis and we have analysed that it is easy to implement Sentiment Analysis via Senti WordNet approach than via Classifier approach.

REFERENCES

- [1] V.K. Singh, R. Piryani, A. Uddin, P. Waila, "Sentiment Analysis of Movie Reviews and Blog Posts", 3rd IEEE International Advance Computing Conference (IACC), 2013
- [2] Mostafa Karamibekr, Ali A. Ghorbani, "Sentiment Analysis of Social Issues", International Conference on Social Informatics, 2012
- [3] Alaa Hamouda, Mohamed Rohaim, "Reviews Classification Using Senti WordNet Lexicon", The Online Journal on Computer Science and Information Technology (OJCSIT), Volume 2, August-2011
- [4] Liu, B. Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1-167) . (2012).
- [5] Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. Eighth International Conference on Weblogs and Social Media
- [6] Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis In: Proceedings of the conference on human language technology and empirical methods in natural language processing, 347–354.. Association for Computational Linguistics, Stroudsburg, PA, USA..
- [7] Tasnim Mohiuddin and Md. Ashraful Islam "A Deep Learning Approach to Sentiment Analysis in Twitter" (2019)
- [8] Jiang L, Yu M, Zhou M, Liu X, Zhao T (2011) Target-dependent twitter sentiment classification In: Proceedings of the 49th, Annual Meeting of the Association for Computational Linguistics: Human Language Technologies- Volume 1, 151–160.. Association for Computational Linguistics, Stroudsburg, PA, USA
- [9] Tan LK-W, Na J-C, Theng Y-L, Chang K (2011) Sentence-level sentiment polarity classification using a linguistic approach In: Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation, 77–87.. Springer, Heidelberg, Germany
- [10] Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis In: Proceedings of the conference on human language technology and empirical methods in natural language processing, 347–354.. Association for Computational Linguistics, Stroudsburg, PA, USA
- [11] Eftymios Kouloumpis, Theresa Wilson, and Johanna Moore. "Twitter Sentiment Analysis: The Good the Bad and the OMG!" Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM). 2011.
- [12] Alec Go, Richa Bhayani, and Lei Huang. "Sentiment Analysis of Twitter Data." Proceedings of the Workshop on Languages in Social Media (LSM). Association for Computational Linguistics, 2009.
- [13] Aditi Gupta and Riddhiman Roy. "Sentiment Analysis on Twitter Data: A Comprehensive Study." Journal of Information Processing Systems, vol. 15, no. 2, 2019.
- [14] Rosenthal, S., Farra, N., & Nakov, P. (2017). SemEval-2017 Task 4: Sentiment analysis in Twitter. In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (pp. 502-518).
- [15] Zhang, Y., & Chen, L. (2019). A review on sentiment analysis research in Twitter. Journal of Information Science, 45(6), 759-790.
- [16] Ghosh, D., & Guha, R. (2013). An analysis of the Twitter sentiment analysis literature: Mining the digital gold rush. ICWSM, 5(2013), 17-21.
- [17] Ruan, L., Yang, C., & Wu, X. (2019). Multi-level attention based neural networks for sentiment analysis of tweets. Knowledge-Based Systems, 165, 1-11.
- [18] Surya, S., & Selvi, S. T. (2021). Sentiment Analysis of Twitter Data using Machine Learning Techniques. International Journal of Engineering and Advanced Technology, 10(1), 272-277.



- [19] Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. Data classification: algorithms and applications, 37-64.
- [20] Das, D., & Bandyopadhyay, S. (2019). Sentiment Analysis of Twitter Data: A Comparative Study of Machine Learning Techniques. In Handbook of Research on Machine Learning Innovations and Trends (pp. 330-353). IGI Global.



SJIF: 8.423



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details