JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# SENTIMENT ANALYSIS USING MACHINE LEARNING.

[1]Jaffarhussain A, [2]N S Sandhya, [3]Shashank M, [4]Sindhu A, [5]Dr. Jithendra PR Nayak

1,2,3,4B.Tech Student,[5]Associate Professor

[1,2,3,4,5]COMPUTER SCIENCE AND ENGINEERING

[1,2,3,4,5]Srinivas Institute of Technology Mangaluru, India

***Abstract:***

This project focuses on automated learning techniques and analyses of emotions based on deep learning to classify positive, negat ive, or neutral emotion textbooks. The analysis is performed using data records such as GosarchSearch and Sentmenta140. The co llected data is suffered by extensive preprocessing, including improving model accuracy using techniques such as text cleaning, to kenization, stop word removal, TF-IDF, Word Bags (ARC), and Word Instruments (Word2vec, Glove, Bert). Numerous auto-learning models including logistics crowds, Vector (SVM) support, Nyver Bayes, Random Forest, and more. According to the mo del, the model is evaluated with power metrics such as accuracy, accuracy, memory, and score F1, providing an essential assessm ent of effectiveness in classifying emotions. This project not only explores theoretical and practical aspects of emotional analysis, but also shows actual applications in social networking, brand reputation management, product check analysis, and interpretation of customer comments. Furthermore, this study highlights the importance of distinctive extraction techniques, optimizing model s election and performance, and optimizing performance in achieving solid classification of emotions..

## I. INTRODUCTION

In this project, we have developed a system of sentiment analysis designed to classify text based on its emotional tone. The aim was to build a model capable of distinguishing simple sentiments such as positive and negative, as well as more nuanced emotions such as joy, anger, sadness, fear and surprise. Sentiment analysis, key subfield processing of natural language (NLP) focuses on the extraction and interpretation of emotions from the text by analyzing words, phrases and structures. It starts from sources such as social media and customer reviews, followed by preliminary processing techniques such as text cleaning, tokenization, lemmatization and removal of tracks. Advanced methods such as verbal insertion (Word2vec, Glove) and sentiment Lexicons increase contextual interpretation, especially in cases such as sarcasm. For classification machine learning models (SVM, Naive Bayes, Random Forest) and deep architecture of learning (RNN, LSTM) have been implemented. The system offers a robust emotional detection, supports a better decision based on data and future progress in multilingual sentiment analysis.

## II. LITERATURE SURVEY

This part summarizes recent research works on the analysis of sentiment by means of ML and DL techniques. Sentiment analysis, the main task in NLP, identifies and categorizes emotions or opinions in the text. Selected papers explore different methods, challenges and applications and offer insight into progress caused by traditional algorithms and modern deep learning models.

1. "Sentiment analysis using machine learning algorithms" from Datrtaray G. Takale et al. (February 2024) compares subordinate models such as SVM, Naive Bayes and random forests with unlimited methods, emphasizing better accuracy with marked data and potential.
2. "Sentiment analysis: tasks, applications and techniques of deep learning" from ABM Shawkat Ali and Neera Sharma (July 2024) discusses the classification of sentiment, mining and emotion detection using CNN and RNNS, emphasizing their role in monitoring public opinion trends.
3. "Sentiment analysis: Access to machine learning" from Dipak Kawade and Kavita Oza (January 2024) compares traditional models of machine learning with deep techniques such as Bert and LSTM, using data sets such as IMDB, Yelp and Twitter to evaluate performance.

4. “Overview of sentiment analysis using machine learning techniques” from Arshad Kabir (April 2024) solves challenges such as sarcasm detection and multilingual sentiment, emphasizing Bert, GPT and Robert.
5. "Review of literature on the application of sentiment analysis using machine learning techniques" from Krishn Prasad to (APR 2024) Review the techniques of extraction (TF-IDF, Word2vec, Bert) and discusses the impacts of sentiment in politics, finance and health care.

## III. METHODOLOGY

The development of a Sentiment Analysis project using ML and DL techniques includes a systematic approach consisting of a several key steps:

Step 1: Data Collection First, relevant text data is taken from social media platforms, customer reviews, public surveys and other open data sets such as IMDB, Yelp, Twitter Sentiment140, or survey data. The collection of various and extensive data sets ensures that the model learns a wide range of sentiments and emotions.

Step 2: Pre -processing data Pre -processing is necessary for the preparation of the raw text for analysis. This includes several sub-steps: Text Cleaning: Removal of undesirable elements such as special characters, numbers, punctuation, URL and emoi addresses. Tokenization: Distribution of text into individual units such as words or phrases. Removing the words stop Words: elimination of common words that do not contribute significant meaning (eg ",", "The", "and"). Lemmatization/stemming: a reduction in words to their root form for uniformity (eg "run" to "run"). Lower occurrence: transfer all text to lowercase letters to avoid duplication due to sensitivity of cases. These steps help create clean and meaningful data inputs for the model.

Step 3: Extraction of functions After pre -work, text data is transformed into a numerical format that can interpret the machine learning models. Techniques include: Bag of Words (Bow): It represents the text as a matrix of words. Inserting words: Word2vec, Glove or Bert.

Step 4: Model selection Machine learning models such as vector machine support (SVM), Naive Bayes is trained for extracted functions. In parallel, DL models such as CNN, RNNS, LSTM and transformer -based models (eg Bert, Robert) are developed for better context understanding. Hyperparameters are tuned using techniques such as grid search helps in improving model performance.

Step 5: Model Rating Models are evaluated with the help of matrices such as accuracy, memories and score F1. Cross validation is performed to ensure the robustness of the model and prevent excessive continuity.

Step 6: Comparative analysis A comparison is made between traditional machine learning models and deep learning models to determine the most effective method, considering factors such as dataset complexity, training time, and classification accuracy.

Step 7: Application and deployment Finally, the best powerful model is deployed for real -world applications such as social media monitoring, feedback analysis from customers and brand management.

Step 8: Challenges and Ethical Reflections The methodology also deals with challenges such as sarcasm detection, multilingual analysis of sentiment and ethical concerns such as personal data protection and algorithmic distortion.

## IV. SYSTEM DESIGN

Sentiment analysis systems are designed to process textual data and reach mood categories such as positive, negative, or neutral The system architecture consists of five main components:

Data Collection: The system collects raw data from sources such as social media platforms (Twitter, Facebook), user input, and public data records (IMDB, YELP)**.** This variety of inputs ensures widespread use in a variety of areas.

Element Preparation and Extraction: Unprocessed text is subject to steps such as text cleaning (noise removal), tokenization (splitt ing words text), stopwatch removal, and lemmatization (reducing words to basic formats). To extract elements, convert texInput i nto numerical representations suitable for model input using techniques such as TF-IDF, Word2vec, and glove.

Sentiment classification: The hybrid approach is used in the classification of sentiment, comparing ML models (Naive Bayes, SVM, Random Forest) and DL (LSTM, CNN, Bert). These models analyse features and classify text into the sentiment categories, with deep learning models capture context and nuances for higher accuracy.

Output: The system generates forecasts of sentiment that are categorized into positive, negative or neutral. The output is shown by the user via text results and visual representations such as charts and charts.

Visualization and applications: The results are presented interactively and allow applications in areas such as social media monitoring, feedback analysis from customers, and brand reputation.
The system ensures scalability and performance optimization for efficient processing of large data sets.

## V. FUTURE WORK OF THE SYSTEM

1. Expand the Sentiment category:

Expand the system to detect a wider range of emotions beyond the basic categories of sentiment (positive, negative, neutral) such as joy, sadness, anger and surprise.

2. To incorporate multilingual support:

Improve the ability of the system to analyze text in multiple languages, making it more versatile and worldwide.

3. Sentiment analysis in real time:

To develop real -time applications that can immediately process and analyze sentiment from social media, customer reviews and interactions with a living cottage.

4. Integration of contextual sentiment models:

Integrate advanced context models such as transformers (eg Bert, GPT) to improve understanding of complex and ambiguous terms of sentiment in the text.

5. Improving efficiency of model and scalability:

Optimize models for faster and more accurate sentiment analysis with reduced computing overhead costs, allowing deployment in environments limited to resources or mobile devices.

6. Detection and alleviating distortion:

Implement methods for detecting and reducing distortion in sentiment analysis and ensuring that the system produces fair and balanced sentiment classification across different demographic groups.

7. Cross -domain adaptability:

Strengthen the ability of the system to adapt to various domains (eg health care, financing, policy) by training on data sets specific to the domain of the sentiment classification in specialized areas.

## VI. IMPACT ON SOCIETY

The sentiment analysis system uses advances in machine learning and NLP to provide valuable knowledge of public opinion and emotions. It has significant consequences for industries such as marketing, customer service and social media monitoring, helping businesses understand feedback, detect trends and improve customer satisfaction.

In addition to commercial use, the sentiment analysis can influence the public discourse by monitoring public sentiment on political, social and political issues and contributing to informed decision -making. It also supports healthier online conversations by helping individuals to interpret emotions behind the text.

In short, the system increases data -based decision -making and benefits both organizations and companies by providing a deeper understanding of emotions in textual content.

## VII. RESULT OF THE SYSTEM

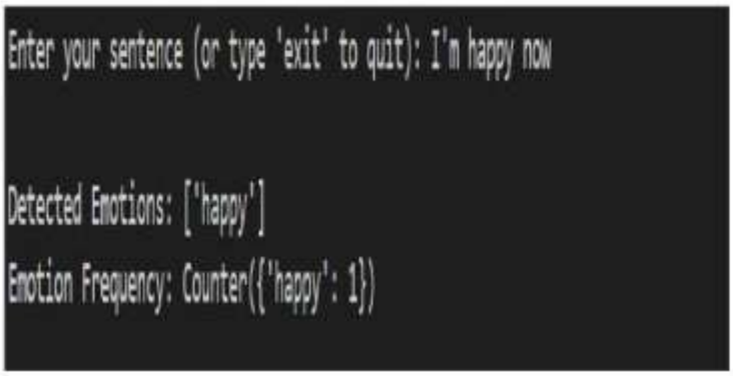


Fig 1: Input

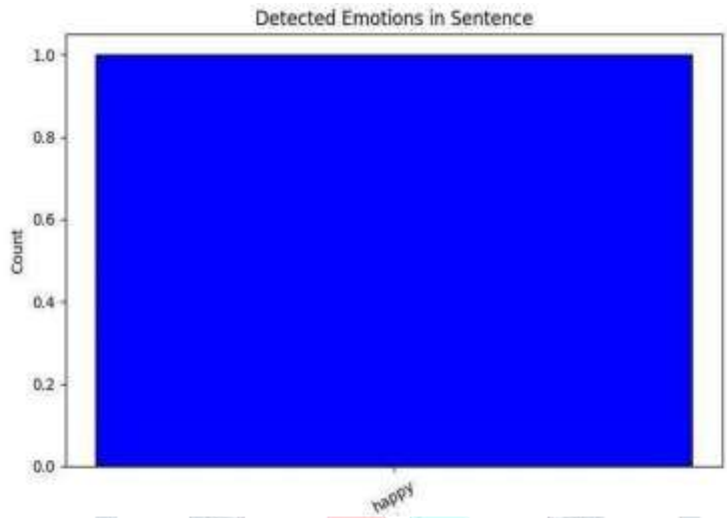


Fig 2: Emotion

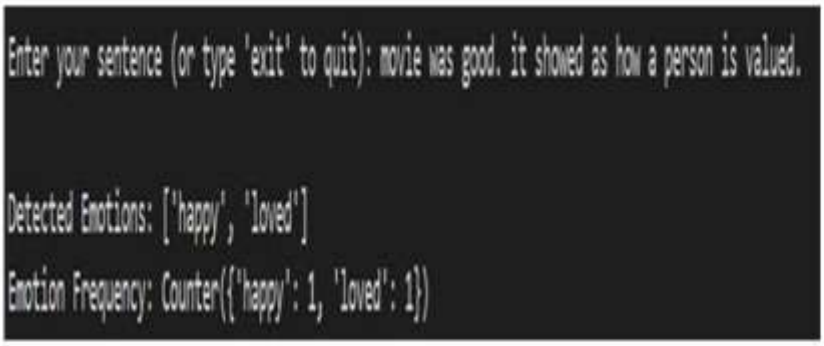


Fig 3: Input

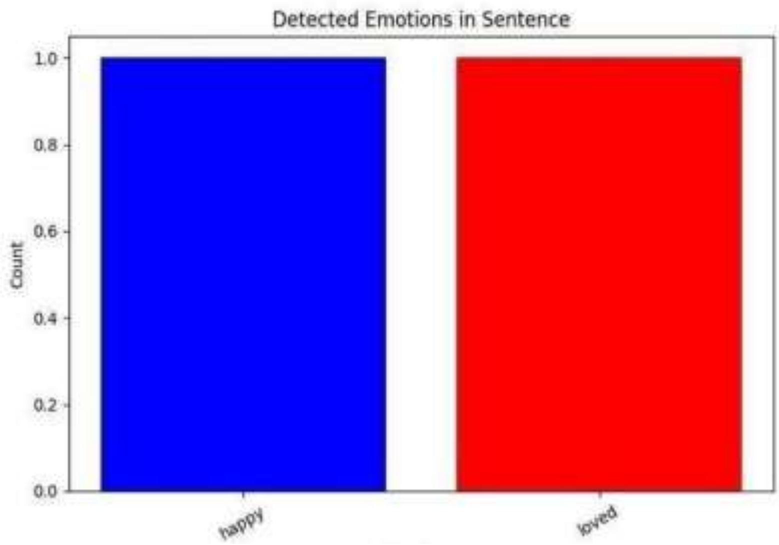


Fig 4: Emotion

## VIII. CONCLUSION

The Sentiment Analysis System is an efficient technology/efficient tool for understanding and classifying the mood of textual data using NLP, ML, and DL techniques. It is classified accurately as positive, negative, and neutral. This is valuable for applications such as social media monitoring, customer feedback analysis, and brand management.

Looking forward, it has potential for future development, including multilingual support for global companies, real -time sentiment analyzes for faster responses and increased detection of complex emotions such as frustration and sarcasm. Tato vylepšení by byla adaptabilnější a bystré.

As the sentiment analysis continues to proceed, it will remain a necessary tool for businesses to understand customers' emotions, improve communication strategies and control informed decision -making in the data -based world.

## IX. ACKNOWLEDGMENT

The authors would like to express their heartfelt gratitude to Dr. Jithendra PR Nayak, Faculty of CSE, Srinivas Institute of Technology, for his valuable guidance, consistent support, and encouragement throughout the development of this project. We are also thankful to our institution for providing the necessary resources and a conducive environment for successful completion of this research. Lastly, we acknowledge the support of our peers and team members whose collaboration and feedback were instrumental in the realization of the Sentiment Analysis.

## REFERENCES

[1] Takale, D. G., Patil, A., Jadhav, S., & Masram, S. (2024). Analysis of sentiment using applications of machine learning algorithms. It compares subordinate models (SVM, Naive Bayes, Random Forest) to unlimited methods, emphasizing higher accuracy with marked data and potential without supervision of learning in social media analysis.

[2] Ali, A. S., & Sharma, N. A. (2024). Sentiment analysis: tasks, applications and techniques of deep learning. It examines the classification of sentiment, mining of opinions and detecting emotions using CNN and RNN focusing on monitoring public opinion trends.

[3] Kawade, D., & Oza, K. (2024). Sentiment analysis: Machine learning approach. It compares the traditional machine learning models with deep learning techniques such as BERT and LSTM, evaluates the performance on data sets such as IMDB, Yelp and Twitter.

[4] KABIR, A. (2024). Overview of sentiment analysis using machine learning techniques. It discusses challenges such as sarcasm detection and multilingual sentiment, emphasize improvements with BERT, GPT and RoBERTa.

[5] Prasad, K. (2024). Overview of literature on the application of sentiment analysis using machine learning techniques. Reviews contain extraction techniques (TF-IDF, Word2Vec, BERT) and discuss the impact of the sentiment analysis in politics, finances and health care focused on ethical concerns.

[6] Sentiment analysis at the function of the function: Survey Non -S. Joshi and Suhasini A. It examines the approaches of machine learning for the analysis of sentiment at the function level.

[7] Lexical and machine learning methods: Blinov et al. (2013) examine the combination of lexical and machine learning methods for sentiment analysis.

[8] Machine learning and deep learning for sentiment analysis: The study deals with the application of machine learning and the technique of deep learning for the analysis of sentiment in government organizations.

[9] Sentiment analysis techniques: Research illustrates various techniques of sentiment analysis, including machine learning methods such as TF-IDF and Naive Bayes, and deep learning methods such as word embedding and LSTM.

[10] Analysis of Twitter Sentiment: The study focuses on the analysis of sentiment using machine learning algorithms for Twitter data.

[11] Comparative study Techniques of Sentiment Analysis: Walaa Medhat et al. (2014) compare different techniques of sentiment analysis, including machine learning and deep learning approaches.

[12] Machine learning techniques for sentiment analysis: Kumar Ravi and Vadlors Ravi (2015) explore techniques of machine learning for sentiment analysis, including Naive Bayes, SVM and Max Entropy.

[13] Deep learning for sentiment analysis: Lei Zhang et al. (2018) survey deep learning approaches for sentiment analysis, including CNN, RNN and LSTM.

[14] Convolutional neural networks for sentiment analysis: Yoon Kim (2014) examines the use of convolutional neural networks for sentiment analysis.

[15] Recurrent neural networks for sentiment analysis: Pengfei Liu et al. (2015) examine the application of recurrent neural networks for sentiment analysis.