INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 1785-1786
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

EMOTIONAL DETECTOR USING NLP

Mrs Kiruba Devi M[1], Srinivasan D[2], Tharaneesh M[3], Sanjiv Kumar S[4]

[1,2,3,4]Proffessor, Information Technology, Sri shakthi institute of engineering and technology, Coimbatore, Tamil Nadu, India.

ABSTRACT

Sentiment Analysis, a key area within Natural Language Processing (NLP), focuses on identifying and extracting subjective information from text, such as emotions, opinions, and sentiments. This project, "Sentiment Analysis Using NLP," aims to analyze and categorize the sentiment embedded in textual data, which can be used to understand public opinion, gauge brand perception, monitor feedback, or detect emotional trends in social media. With the rapid growth of textual data on platforms like Twitter, Facebook, and product review sites, sentiment analysis has become crucial for businesses and organizations that wish to understand and respond to their audience's needs and concerns..

1. INTRODUCTION

Sentiment analysis, also referred to as opinion mining, is a critical subfield of Natural Language Processing (NLP) that focuses on determining the sentiment expressed in a piece of text. With the increasing amount of textual data available from social media platforms, customer reviews, sentiment analysis has become an essential tool for businesses, researchers, and decision-makers.We explore the methodologies used in sentiment analysis and evaluates the efficacy of different NLP techniques, such as tokenization, part-of-speech tagging, and sentiment scoring, in classifying text data into positive, negative, or neutral sentiments.

2. METHODOLOGY

1. Problem Definition and Objectives :

Goal Identification: Clearly define the primary goals and objectives of the sentiment analysis project, such as identifying in specific types of text data (e.g., product reviews, social media posts).

Scope of Analysis: Determine the scope, including the type of sentiment (e.g., emotions like anger, joy) and specific domains (e.g., product reviews, political opinions).

2. Data Collection :

Data Sources: Identify and collect data from relevant sources, such as social media platforms (e.g., Twitter, Instagram,Facebook),

Data Preprocessing: Clean the collected text data to remove noise, such as irrelevant characters, stopwords,etc . Text normalization (e.g., stemming, lemmatization) will be applied as needed.

Labeling and Annotation: If using supervised learning, label data with sentiment categories (e.g., positive, negative, neutral).

3. Text Pre-processing :

Tokenization: Divide text data into individual words or tokens for easier analysis.

Vectorization: Convert text data into numerical representations using techniques like Bag of Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF)

4. Model Training and Tuning :

Data Splitting: Divide the dataset into training, validation, and test sets to ensure that the model can generalize well to new data.

Cross-Validation: Apply cross-validation to assess the model's robustness and avoid overfitting.

5. Error Analysis :

Review Mis-classified Instances: Identify patterns in mis-classified samples to gain insights into where the model struggles (e.g., handling sarcasm, irony, or complex language).

Iterative Model Improvement: Based on error analysis, adjust pre-processing, feature extraction, or model architecture as necessary.

3. MODELING AND ANALYSIS

1. Model Performance and Comparison :

Key Observations:

Logistic Regression: Logistic Regression achieved slightly higher performance than Naive Bayes, benefiting from the simplicity and robustness of the model. It can capture linear relationships.

2. Limitations and Challenges :

Difficulty Handling Long-Term Dependencies: These models do not consider the order or context of words in a sentence, which limits their ability to accurately interpret sentiment in longer sentences or documents.

Baseline Comparisons: These models provide an effective baseline for sentiment analysis. They are easy to interpret, providing insight into which features (words) are most indicative of positive or negative sentiment.

Decent Performance with Proper Preprocessing: With well-designed pre-processing steps, such as removing stopwords, stemming / lemmatization, and converting text to numeric representations (like TF-IDF), these models can capture fundamental sentiment patterns in the data.

3. Error Analysis :

Misclassified Samples:

Ambiguous Language: Texts with mixed sentiment (e.g., "The product is great but the service was terrible") posed challenges, as the models had difficulty determining the overall sentiment.

Out-of-Vocabulary Terms: Certain domain-specific terms or slang were not recognized well, which may be improved by further fine-tuning with more specific datasets.

4. Limitations :

Data Dependency: The models' performance heavily depends on the quality and representativeness of the training data. Limited diversity in training data can lead to biases or inadequate performance on new types of text.

5. Future Improvements :

Improved Handling of Sarcasm and Irony: Incorporating more complex training data or specialized models for sarcasm detection could improve performance in ambiguous cases

Domain-Specific Fine-Tuning: Fine-tuning on domain-specific data (e.g., industry-specific language or slang) would likely improve the model's ability to understand and classify nuanced sentiment accurately.

4. CONCLUSION

This approach demonstrates the effectiveness of deep learning in sentiment analysis, with potential for further enhancement.The findings underscore the value of NLP and deep learning in automating sentiment extraction, helping organizations make data-driven decisions based on customer sentiment

ACKNOWLEDGEMENT

Everyone who helped to make this emotional detector using NLP project a success has my sincere gratitude. Firstly, I would like to sincerely thank [Mrs.M.Kirubadevi], my project adviser, for all of her help and assistance during the creation of this project. This guidance has been priceless. Overcoming the obstacles encountered during the project was made possible in large part by their knowledge and -support.

My sincere gratitude goes out to my co-workers, [D.Srinivasan], [S.Sanjiv Kumar], and [M.Tharaneesh], for their diligent labor, teamwork, and commitment. Achieving our objectives required our team's synergy and mutual support.We extend our sincere gratitude to [Sri Shakthi Institute of Engineering and Technology] for supplying the tools and setting that support our R&D activities. I am also grateful for the comments and guidance from the professors and other students who looked over our work and offered insightful recommendations. The development of our project was tremendously aided by the contributions and resources provided by the open-source community, for which I am also thankful.

Ultimately, I want to express my gratitude to my friends and family for their constant support and inspiration, which inspired me to keep going and finish this project. We appreciate your support and donations, everyone.

5. REFERENCES

[1] Ali, A. 2001.Macroeconomic variables as common pervasive risk factors and the empirical content of the Arbitrage Pricing Theory. Journal of Empirical finance, 5(3): 221–240.

[2] Basu, S. 1997. The Investment Performance of Common Stocks in Relation to their Price to Earnings Ratio: A Test of the Efficient Markets Hypothesis. Journal of Finance, 33(3): 663-682.

[3] Bhatti, U. and Hanif. M. 2010. Validity of Capital Assets Pricing Model.Evidence from KSE-Pakistan.European Journal of Economics, Finance and Administrative Science, 3 (20).