ISSN No:-2456-2165

An Application : Emo-Prediction Using Sentiment Analysis

[1]Dr. M.S. Chaudhari, [2]Ms. Nupur Banode, [3]Kaushal Lande, [3]Ankita Nademwar, [3]Siddhi Ghosekar, [3]Dnyanesh Lambat, [3]Sagar Lende
[1]Head of Department, Department of Information Technology, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India
[2]Assistant Professor, Department of Information Technology, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India
[3]UG Student, Department of Information Technology, Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

Abstract:- The research paper investigates the effectiveness of machine learning methods in predicting human emotions from text data through sentiment analysis. It reviews existing literature, explores diverse datasets, and applies various machine learning models to predict emotions accurately. Results demonstrate the success of these models, highlighting their real-world applications in marketing, mental health analysis, and user experience enhancement. The study concludes by outlining future research directions for improved emotion prediction using sentiment analysis, contributing to the advancement of understanding and utilizing emotional states derived from text.

Keywords:- *Sentiment Analysis, Polarity, Bag of words, SentiWordNet.*

I. INTRODUCTION

Understanding human emotions through text data is an evolving and promising field with wide-ranging applications. The extraction and prediction of emotions from textual content using sentiment analysis and machine learning techniques hold immense potential across various domains. These applications span from social media analysis and market research to mental health assessment and user experience enhancement. Written text can express lots of different feelings, making it a critical avenue for understanding user behavior, preferences, and mental well-being.

Aim of this research paper is to investigate the effectiveness of machine learning models in predicting emotions from textual data, thus contributing to the ongoing discourse in sentiment analysis methodologies. By exploring and analyzing diverse datasets and employing various machine learning algorithms, this study endeavors to provide valuable insights into the potential of sentiment analysis in decoding human emotions. Emotion prediction using sentiment analysis becoming an important thing to study, given its implications in understanding the emotional content embedded within text.

In Sentiment analysis methodologies has gradually extended to incorporate advanced machine learning techniques that capture the complexities of human emotions expressed through text. Various machine learning models aims to address these limitations by investigating and evaluating the efficacy of in predicting emotions, thereby contributing to the ongoing enhancement and refinement of sentiment analysis approaches.

By exploring diverse datasets and assessing the accuracy of emotion prediction, this research endeavors to offer valuable insights into the potential applications of sentiment analysis in deciphering human emotions. Through this exploration, the paper seeks to lay the foundation for further advancements in understanding and utilizing emotional states derived from text.

II. LITERATURE SERVEY

It can be seen that the role of online product reviews is mainly manifested. An objective of product market is Reviews of online product can affect product sales, revenue, and market competition and the beliefs of consumers. Online product review influences consumer judgment on products through information dissemination. It may actively promote the usefulness of reviews, and also cause decision-making difficulties for consumers.[1]In an empirical study of text-based emotion prediction. In order to make text-to-speech synthesis sound as natural and engaging as possible, it's important to convey the emotional stance in the text. This implies first to identify the appropriate emotional meaning of the corresponding text passage. [2] SENTIWORDNET 3.0 the results of this semi-supervised learning algorithm are only an intermediate step of the annotation process, since they are fed to an iterative random-walk process that is run to convergence. It gives output of the random-walk process after convergence has been reached.[3] CES-D (Center for Epidemiologic Studies Depression Scale) is used for Depression Screening Test. It is a questionnaire as the primary tool to determine the depression levels of the crowdworkers. It measures symptoms defined by the American Psychiatric Association Diagnostic and Statistical Manual (DSM-IV), and quantifies depressive feelings and behaviors during the past week.[4] By extracting the features or

different characteristics from the speech and a training is needed for a large number of speech database to make the system accurate identification of emotion can be done. For an emotional speech they used corpora which is selected or implemented then emotion specific features are extracted from those speeches and finally a classification model is used to recognize the emotions.[5] In this SER(Speech Emotion Recognition) is used to automatically determine the emotional state of the speaker via a speech signal. Changes may be observed when the waveform's frequency and intensity changed and compared with different emotionally coloured speech signals. It captures these variations using different discriminative acoustic features.[6] A study using the features Pitch and energy and classifier HMM was performed in impact of Attention Mechanism on Speech Emotion Recognition where the accuracy rate achieved was 86%.

Another study used Support Vector Machine as a classifier with Berlin Database where the overall recognition rate was 82.5%.

III. PROPOSED METHODOLOGY

- Data Collection: Gather a diverse and representative dataset of textual content for training and testing the sentiment analysis model. This dataset should include examples of text with known sentiment labels (positive, negative, neutral).
- Preprocessing: Perform text preprocessing, including lowercasing, tokenization, and removal of stop words, punctuation, and special characters. Additionally, apply stemming or lemmatization to standardize the text.
- Feature Extraction: Extract relevant linguistic features from the preprocessed text, such as word frequencies, n-grams, and word embeddings (e.g., Word2Vec or GloVe). These features will serve as input for the sentiment analysis model.
- Sentiment Analysis Model Selection: Choose an appropriate NLP-based sentiment analysis model, such as a supervised machine learning model (e.g., Support Vector Machines, Random Forest, or deep learning models like LSTM or BERT) or lexicon-based approaches.
- Model Training: Train the selected sentiment analysis model on the labeled dataset, using a portion for training and another for validation. Fine-tune the model parameters for optimal performance.
- Model Evaluation: Assess the model's performance using various evaluation metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, to determine its effectiveness in classifying text into positive, negative, or neutral sentiments.
- User Interface Design: Develop a user-friendly interface for the Emo-Prediction App. This interface should allow users to input text for sentiment analysis and receive results in a clear and comprehensible format.
- Real-Time Analysis: Implement the capability for real-time sentiment analysis, ensuring that users receive immediate feedback on the emotional tone of the provided text.
- Multilingual Support: Incorporate language detection and support for multiple languages to make the app accessible to a global audience.
- Confidence Score Calculation: Calculate a confidence score for each sentiment classification to indicate the reliability of the prediction. This score can help users gauge the certainty of the analysis.
- Integration and Customization: Allow users to integrate the Emo-Prediction App into their systems or workflows and customize it to suit their specific requirements, such as industry-specific sentiment lexicons.
- Testing and Validation**:** Conduct extensive testing and validation to ensure the app's accuracy and reliability. Test it on a diverse set of text samples to verify its performance across different domains and contexts.
- Documentation: Prepare comprehensive documentation that explains how to use the Emo-Prediction App, including the API if applicable, and provides insights into the sentiment analysis model used.
- Deployment and Monitoring: Deploy the app on the chosen platform (web, mobile, desktop, etc.) and continuously monitor its performance. Regularly update the sentiment analysis model to adapt to changing linguistic patterns and user needs.
- User Feedback and Improvement: Collect user feedback and suggestions for app improvements and incorporate these into future updates. This iterative process ensures the app remains relevant and effective.

The interpretation of the findings involved a detailed analysis and discussion of the implications within the broader context of sentiment analysis. The focus was on the role of accurately predicting emotions from textual data and the potential practical applications derived from these insights.

IV. TECHNIQUES USED

- Bag-of-Words(BoW): This technique involves representing text as a collection of words, disregarding grammar and word order. It quantifies the occurrence of words within the text and is a foundational approach for sentiment analysis.

- Machine Learning Models: Various machine learning models are employed for sentiment analysis and emotion prediction, including:

- Naive Bayes: A probabilistic model that calculates the probability of a sentiment or emotion based on the occurrence of words.
- Support Vector Machines (SVM): Used for classification tasks, SVM separates data points into different classes based on the features extracted from the text.
- Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM): Deep learning models suitable for sequence data analysis, capable of capturing dependencies within text.

- Sentiment Lexicons: Lexicons or dictionaries containing lists of words associated with sentiment (positive, negative, neutral) or specific emotions. These lexicons assign scores or labels to words, aiding sentiment analysis.
- Hybrid Approaches: Combining multiple techniques, such as using both lexicon-based sentiment analysis and machine learning models, to enhance the accuracy of emotion prediction.
- Preprocessing Techniques: Text preprocessing involves steps like tokenization, removing stop words, stemming or lemmatization, and handling special characters to clean and prepare the text data for analysis.

V. CONCLUSIONS

In conclusion, our research delved into the exploration of various machine learning models for emotion prediction through sentiment analysis. The comparative analysis revealed that while traditional models like Naive Bayes and Support Vector Machines (SVM) offer respectable accuracy in predicting emotions from text data, the application of deep learning, particularly Long Short-Term Memory (LSTM) networks, showcased superior performance. LSTM's ability to capture nuanced emotional context within textual data, reflected in its higher accuracy and balanced precision and recall metrics, underscores its potential for more intricate emotion detection. These findings highlight the significance of model selection and the evolution toward leveraging advanced deep learning architectures for more accurate and nuanced emotion prediction.

REFERENCES

[1]. Huifen Wang, Yang Wang "A Review of Online Product Reviews" Journal of Service Science and Management, 2020, 13, 88-96
[2]. Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat, 2005, "Emotions from text: Machine learning for text-based emotion prediction." In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 579--586.
[3]. Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani, 2010, "SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining." In LREC, Vol. 10. 2200--2204.
[4]. Munmun De Choudhury, Michael Gamon, Scott Counts, Eric Horvitz "Predicting Depression via Social Media" Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media.
[5]. Saikat Basu, Jaybrata Chakraborty, Arnab Bag, Md. Aftabuddin "A Review on Emotion Recognition using Speech", International Conference on Inventive Communication and Computational Technologies (ICICCT 2017)
[6]. Eva Lieskovská, Maroš Jakubec, Roman Jarina, Michal Chmulík "A Review on Speech Emotion Recognition Using Deep Learning and Attention Mechanism", Faculty of Electrical Engineering and Information Technology, University of Žilina, Electronics 2021
[7]. Shaikh Nilofer R. A., Rani P. Gadhe, R. R. Deshmukh, V. B. Waghmare, P. P. Shrishrimal, "Automatic emotion recognition from speech signals: A Review", International Journal of Scientific & Engineering Research, Volume 6, Issue 4, April-2015 636 ISSN 2229-5518