Emotion Intensity Prediction in texts.

Abstract:

Understanding the intensity of emotions expressed in texts, including anger, fear, joy, and sadness, is crucial for analyzing sentiment and user behavior. This study examines the effectiveness of two methods, the lexicon-based approach and the VADER method, for predicting emotion intensity in texts. The lexicon-based approach assigns emotion labels to words in the text and aggregates them to determine overall emotional content. In contrast, the VADER method is a rule-based sentiment analysis tool that calculates sentiment polarity based on lexical features.

The study finds that the lexicon-based approach is effective in predicting emotion intensity in texts. This approach provides valuable insights into sentiment analysis, enabling a deeper understanding of emotional content in texts. Accurate prediction of emotion intensity allows for better understanding of user sentiments and behaviors in various contexts. Overall, this study emphasizes the importance of considering emotion intensity in sentiment analysis and highlights the effectiveness of the lexicon-based approach for this purpose.

Introduction:

Emotion intensity analysis in texts provides deeper insights into the emotional content of communication. This study focuses on predicting emotion intensity in texts for four emotions: anger, fear, joy, and sadness. By accurately predicting emotion intensity, we can better understand the emotional dynamics of individuals and communities on social media platforms. The study uses a dataset of labeled texts to train and test models for predicting emotion intensity. Two methods, the lexicon-based approach and the VADER method, are compared for their effectiveness. The results show that the lexicon-based approach is effective in predicting emotion intensity, highlighting its potential for sentiment analysis. This study contributes to the field of sentiment analysis by emphasizing the importance of considering emotion intensity in understanding user emotions and behaviors in online communication.

Objectives:

- Evaluate the effectiveness of the lexicon-based approach and the VADER method for predicting emotion intensity in texts.
- Compare the performance of these methods and provide insights into their effectiveness for sentiment analysis.
- Explore the potential applications of emotion intensity prediction in texts.

Methodology:

Data:
We use a dataset of texts labeled with emotion intensity scores for anger, fear, joy, and sadness. This dataset allows us to train and test our models for predicting emotion intensity.

Text Preprocessing:
Before training our models, we preprocess the text by removing Twitter handles, URLs, and stopwords. We also tokenize the text to prepare it for analysis.

Model Training:
We use the TF-IDF vectorization technique to convert the preprocessed text into numerical features. Then, we train a LinearSVR model for each emotion to predict the intensity scores.

Results:

The performance of our models is evaluated using Root Mean Squared Error (RMSE) on the test set. Here are the RMSE values we achieved:

- Anger: 0.0629
- Fear: 0.0704
- Joy: 0.0723
- Sadness: 0.0654

Conclusion:

The lexicon-based approach, combined with the LinearSVR algorithm, proves to be effective for predicting emotion intensity in texts. The simplicity and interpretability of the lexicon-based approach make it a valuable tool for sentiment analysis. The use of the LinearSVR algorithm allows for efficient prediction of emotion intensity, providing insights into the emotional content of texts. Further research could explore other simple methods and compare them with more complex ones to improve emotion intensity prediction in texts.

~Farman ali
Farmanali27000@gmail.com
+91 9449987882