Emotion Intensity Prediction in Tweets

Abstract:

Emotion intensity analysis in text is crucial for understanding the underlying sentiment and mood of individuals. This study compares the effectiveness of statistical and deep learning models for predicting emotion intensity in tweets. The dataset includes tweets labeled with four emotions: anger, fear, joy, and sadness, each assigned a real-valued intensity score. For statistical model, a lexicon-based approach and the VADER method are used. For deep learning model, Long Short-Term Memory (LSTM) networks are employed. Results show that the LSTM model outperforms the statistical models, indicating that deep learning models are more effective for emotion intensity prediction in tweets.

1. Introduction

Background

Emotion intensity analysis aims to quantify the degree of emotion expressed in text, providing deeper insights into the emotional content of communication. Understanding the intensity of emotions, such as anger, fear, joy, and sadness, in tweets can be valuable for various applications, including sentiment analysis, social media monitoring, and mental health research.

Significance

Accurately predicting emotion intensity in tweets can help in understanding the emotional dynamics of individuals and communities, leading to improved communication strategies, targeted interventions, and better mental health support. However, the task of emotion intensity prediction poses several challenges, including the subjective nature of emotions and the nuances in their expression in text.

Objectives

The primary objective of this study is to compare the effectiveness of statistical and deep learning models for predicting emotion intensity in tweets. Specifically, the study aims to:

- Evaluate the performance of statistical models, including a lexicon-based approach and the VADER method, for emotion intensity prediction.
- Compare the performance of statistical models with deep learning models and machine learning models, such as LSTM and Linear regression, for emotion intensity prediction.
- Analyse the strengths and limitations of each approach and provide insights into their effectiveness for emotion intensity prediction in tweets.

2. Methodology

The study uses a dataset of tweets labelled with emotion intensity scores for anger, fear, joy, and sadness. The dataset is divided into training, validation, and test sets for model training and evaluation.

Statistical models, including a lexicon-based approach and the VADER method, are employed for sentiment analysis. Deep learning models, such as LSTM, and Linear regression model are also used for comparison.

The models are trained on the training and validation sets and evaluated using RMSE on the test set. The performance of each approach is analysed and compared to determine the most effective method for emotion intensity prediction in tweets.

2.1. Statistical models:

2.1.1. **VADER Method:**

The VADER method is a lexicon and rule-based sentiment analysis tool specifically designed for social media text. It analyses the sentiment of text based on a predefined list of lexical features, including:

- **Polarity Scores:** VADER calculates the sentiment polarity of text based on the presence of positive, negative, and neutral words in the text. It assigns a sentiment score to the text, ranging from -1 (most negative) to +1 (most positive).
- **Valence Shifters:** VADER considers valence shifters (e.g., "not," "very") to adjust the sentiment intensity of words in context. For example, "not good" would be considered negative despite "good" being positive, due to the presence of "not."
- **Emoticons and Punctuation:** VADER also considers emoticons and punctuation (e.g., exclamation marks) as features that indicate sentiment intensity in text.

- **Advantages:**

  VADER is particularly effective for social media text due to its ability to handle informal language, emoticons, and slang. It provides a quick and easy way to analyse sentiment in large volumes of text.

- **Limitations:**

  VADER's performance may vary depending on the domain or context of the text. It may struggle with sarcasm, irony, and nuanced language, as it relies heavily on lexical features.

2.1.2. **Lexicon-Based Approach:**

The lexicon-based approach using the NRC Emotion Lexicon involves assigning emotion labels to words in text and aggregating these labels to determine the overall emotional content.

The NRC lexicon contains words annotated with emotions like joy, sadness, anger, fear, trust, disgust, and anticipation, with binary values indicating their association with each emotion. This approach is used to assign emotion labels to words in tweets and determine the overall sentiment (positive, negative, neutral) of the tweets.

- **Advantages:**

  Simplicity and Interpretability: Lexicon-based approaches are simple to implement and interpret, making them accessible for quick sentiment analysis tasks.

- **Limitations:**

  Lack of Nuance: Lexicon-based approaches may oversimplify sentiment analysis, lacking the nuance to understand context, sarcasm, or subtle emotions in text.

Overall, while lexicon-based approaches are straightforward and easy to implement, they may not always provide the most accurate or nuanced sentiment analysis results, especially in cases where context and subtlety are important.

2.1.3. **Linear Regression**

- **Description:** Linear regression is a statistical model that assumes a linear relationship between the independent variables (features) and the dependent variable (target). It is used to predict the target variable based on the values of the features.
- **Training:** The linear regression model is trained using the training data, where the features are the pre-processed tweet text and the target variable is the emotion intensity score. The model learns the coefficients for each feature to minimize the prediction error.
- **Evaluation:** The model is evaluated using Root Mean Squared Error (RMSE) on the validation set. RMSE measures the difference between the predicted intensity scores and the actual scores, providing a measure of the model's performance.

2.2. Deep Learning Models

Long Short-Term Memory (LSTM)

- **Description:** LSTM is a type of recurrent neural network (RNN) designed to model sequential data. It is particularly effective for capturing long-term dependencies in sequences, making it well-suited for text analysis tasks.
- **Architecture:** The LSTM model consists of an embedding layer, an LSTM layer, and a dense output layer. The embedding layer converts the input text into dense vectors, which are then fed into the LSTM layer. The LSTM layer processes the sequences and learns the underlying patterns in the data. The dense output layer produces the final prediction.
- **Hyperparameters:** The hyperparameters of the LSTM model include the embedding dimension, the number of LSTM units, the dropout rate (for regularization), and the learning rate.
- **Training:** The LSTM model is trained using the training data, where the input sequences are the pre-processed tweet text and the target variable is the emotion intensity score. The model is trained to minimize the prediction error using an optimization algorithm such as Adam.
- **Evaluation:** The model is evaluated using RMSE on the validation set to assess its performance. The RMSE provides a measure of how well the model predicts the emotion intensity scores compared to the actual scores.

3. Results

- **VADER:** RMSE: 0.76
- **Lexicon-based Approach (NRC Emotion Lexicon):** Accuracy: 0.42
- **Linear Regression:** anger-RMSE: 0.24, fear-RMSE: 0.21, joy-RMSE: 0.23, sadness-RMSE: 0.27
- **LSTM:** Test Loss: 0.0042

The VADER model has the highest RMSE on the test set, indicating that it has the lowest performance among the models evaluated. The lexicon-based approach using the NRC Emotion Lexicon has a moderate accuracy of 0.42, indicating that it performs better than VADER but not as well as the other models.

Among the statistical models, linear regression performs reasonably well, with RMSE values ranging from 0.21 to 0.27 for different emotions. The LSTM model outperforms all other models, achieving a very low-test loss of 0.0041, indicating high performance in predicting emotion intensity from text.

Overall, the LSTM model shows superior performance compared to the other models, highlighting the effectiveness of deep learning approaches for emotion intensity analysis in text.

4. Discussion

Interpreting the Results:

The results demonstrate that the LSTM model outperforms the other models in predicting emotion intensity from text. It achieved a significantly lower test loss compared to the lexicon-based approach and linear regression. This indicates that the LSTM model is better able to capture the nuanced relationships between words and emotions in the dataset.

Implications of the Findings:

The findings suggest that deep learning models, specifically LSTMs, are effective for emotion intensity analysis in text. This has implications for various applications such as sentiment analysis, customer feedback analysis, and social media monitoring, where understanding the intensity of emotions expressed in text is important.

Comparison to Previous Studies:

Previous studies have also highlighted the effectiveness of deep learning models for emotion analysis tasks. The results of this study are consistent with these findings and further demonstrate the superiority of deep learning models over traditional lexicon-based and statistical approaches.

Strengths and Limitations of the Approach:

Strengths:

- The LSTM model can capture complex patterns in text data, allowing it to effectively predict emotion intensity.
- The use of the NRC Emotion Lexicon in the lexicon-based approach provides a structured way to analyse emotions in text.

Limitations:

- The lexicon-based approach relies heavily on the quality and coverage of the emotion lexicon, which may not capture all nuances of human emotions.
- The performance of the LSTM model may be sensitive to hyperparameters and model architecture, requiring careful tuning.

Overall, the findings suggest that deep learning models, particularly LSTMs, are a promising approach for emotion intensity analysis in text, offering improvements over traditional lexicon-based and statistical methods.

5. Conclusion

In conclusion, our study demonstrates the effectiveness of deep learning, particularly LSTM models, for predicting emotion intensity in text. The practical implications of this work are significant, as accurate emotion intensity analysis can improve the performance of various natural language processing applications, such as sentiment analysis, chatbots, and recommendation systems.

Future Research Directions:

Future research could focus on improving the interpretability of deep learning models for emotion intensity analysis, exploring techniques to reduce the reliance on labelled data, and extending the approach to other languages and domains. Additionally, investigating the impact of context and linguistic cues on emotion intensity prediction could further enhance the performance of these models.

References:

1. Hutto, C.J., & Gilbert, E. (2014, May). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth International Conference on Weblogs and Social Media (ICWSM-14).
2. Mohammad, S., & Turney, P. (2013). Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), 436-465.
3. Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013, October). Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 conference on empirical methods in natural language processing (pp. 1631-1642).
4. NRC Emotion Lexicon: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm
5. Task on Emotion Intensity:http://saifmohammad.com/WebPages/EmotionIntensity-SharedTask.html

Author

Muthukumaran A

muthukumaran382@gmail.com