Emotion Intensity Prediction in Text: A Comparative Study

Abstract:

This study compares the effectiveness of statistical and deep learning models for predicting emotion intensity in text. The dataset consists of texts labeled with four emotions: anger, fear, joy, and sadness, each assigned a real-valued intensity score. Statistical models include a lexicon-based approach and the VADER method, while deep learning models use Long Short-Term Memory (LSTM) networks. Results indicate that LSTM outperforms statistical models, highlighting the effectiveness of deep learning for emotion intensity prediction in text.

Introduction:

Understanding emotion intensity in text is crucial for various applications, including sentiment analysis and mental health research. This study compares statistical and deep learning models for predicting emotion intensity in text, aiming to provide insights into their effectiveness and applicability.

Methodology:

The dataset includes texts labeled with emotion intensity scores, divided into training, validation, and test sets. Statistical models, including a lexicon-based approach and VADER, are compared with deep learning models like LSTM. Models are trained on the training and validation sets and evaluated using RMSE on the test set.

Results:

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

Discussion:

The LSTM model outperforms other models, including VADER, the lexicon-based approach, and linear regression, achieving a low test loss. This suggests that LSTM is more effective in capturing nuances in emotion intensity in text.

~Farman ali

Farmanali27000@gmail.com