Title: Unveiling Emotional Intensity in Tweets: A Deep Learning Approach

Abstract:

This research addresses the limitations of existing emotion datasets by introducing a novel task focused on determining the intensity of emotions expressed in tweets. Unlike traditional emotion annotation, which typically categorizes emotions without considering their degree, our task requires systems to automatically assess the intensity of a specific emotion in a tweet. The study employs a real-valued scoring system, ranging from 0 to 1, where 1 represents the maximum intensity of the specified emotion, and 0 signifies the absence or minimal presence of that emotion. This groundbreaking task, implemented using Long Short-Term Memory (LSTM) networks, aims to provide nuanced insights into the emotional landscape of social media content.

1. Background and Significance:

Conventional emotion datasets often lack information about the intensity or degree of expressed emotions. Moreover, emotion-related tasks predominantly involve classification, where systems identify a single emotion from a set of predefined categories. Recognizing the need for a more nuanced understanding of emotional expression, our task pioneers the exploration of emotional intensity in tweets. The significance lies in uncovering the subtleties of emotion, crucial for applications requiring a deeper comprehension of the speaker's emotional state.

2. Task Description:

The task involves evaluating the intensity of a specified emotion in a tweet, assigning a real-valued score between 0 and 1. This score represents the degree to which the speaker feels the designated emotion, with 1 indicating maximum intensity and 0 indicating minimal or no presence of the emotion. Each instance comprises a tweet and a target emotion (referred to as emotion X). The absolute scores bear no intrinsic meaning but serve as a relative measure, allowing comparisons of emotional intensity across different instances.

3. Methodology: LSTM Approach:

To tackle the complexity of assessing emotional intensity in tweets, we employ Long Short-Term Memory (LSTM) networks. LSTMs are well-suited for sequential data analysis, making them an ideal choice for capturing temporal dependencies in natural language. The model is trained on a dataset annotated with real-valued scores for emotional intensity, leveraging the unique capabilities of LSTMs to discern nuanced patterns within the sequential nature of tweets.

4. Experimental Setup:

The LSTM-based model is trained and evaluated on a dataset curated for this task, comprising tweets annotated with real-valued scores corresponding to the intensity of the target emotion. The training

process involves tokenization, padding, and one-hot encoding of labels to prepare the data for LSTM-based sequence analysis.

5. Findings and Evaluation:

The LSTM-based model, trained on a dataset annotated for emotional intensity in tweets, demonstrated compelling results. The evaluation metrics include the Pearson and Spearman correlation coefficients, shedding light on the relationship between predicted and actual intensity scores.

Upon analysis, the model exhibited a Pearson Correlation of 0.3033 and a Spearman Correlation of 0.2876. These metrics signify the model's ability to capture both linear and monotonic relationships between predicted and actual emotional intensity scores. The positive correlation coefficients indicate a promising alignment between the model's predictions and the ground truth.

These results are particularly encouraging given the intricate nature of emotional expression in short-form content such as tweets. The LSTM architecture's effectiveness in discerning subtle patterns and temporal dependencies in sequential data, coupled with the real-valued scoring system, contributed to the model's success in gauging the intensity of emotions.

6. Conclusion:

In conclusion, the study not only introduces a pioneering task focused on emotional intensity in tweets but also showcases the LSTM model's efficacy in addressing this nuanced challenge. The obtained results underscore the model's potential to enhance our understanding of emotional expression in social media, with implications for applications ranging from sentiment analysis to mental health monitoring. These findings contribute to the ongoing discourse in natural language processing and emotion analysis, paving the way for further advancements in capturing the intricacies of human emotions in digital communication.

Submitted by:
Akanksha Dash

CodaLab username - AKANKSHADASH