ABSTRACT

Sentiment analysis is the classification problem where an algorithm recognizes the primary emotions that a media input evokes. Emotions are subjective, so classifying media into distinct groups is a challenging problem. Most human subjects agree in broad strokes on emotional classifications. However it is not uncommon to find media inputs where there is no consensus, leading to inconsistencies in class groupings. As a result, sentiment analysis is more challenging to incorporate into real-world applications because an algorithms recognition accuracy might differ between classes.

This paper introduces a multimodal approach to the sentiment analysis problem. This method leverages the variety of information available to minimize class inconsistencies. By analyzing songs through two independent media spaces — audio and lyrics — the algorithm combines the inherent emotional information to reach a single overall classification.

Keywords: Signal Processing, Sentiment Analysis, Machine Learning, Feature Fusion, Multimodal Classification