Feature Weighting in Multimodal Affect Prediction and Emotional Inference

Keziah, Virginia (School: Fairview High School)

Emotion inference is the ability to infer how another individual is feeling and is crucial to social interaction and well-being. The predominant form of investigation involves unimodal, simplified cues that participants use to evaluate others' feelings. However observers in real-world situations must rely on multiple factors in determining the emotion expression of others (e.g. facial expression, voice, prosody). Although daily emotion inference requires multimodal cue integration, little is known about the relative importance of expressive features, specifically which are most predictive of emotional valence. Emotion inference also becomes increasingly important in creating emotionally intelligent AI as technology is further integrated into daily life. Here I investigate the process of emotion inference, and the relative importance of expressive features. I created three models to infer emotion based on vocal, facial, and multimodal cue inputs, with an output of instantaneous predictive ratings of emotional valence. Such features were extracted from continuously rated naturalistic videos. Each model was correlated to the storytellers' actual ratings and the average observers' ratings to evaluate their relative performances. I then used a system of lesioning each feature within instances in which the multimodal model was successful to investigate which features had the biggest effect on the model's error, and were therefore most predictive of emotional valence. This project both builds groundwork for more emotionally intelligent AI and introduces a system of feature classification to aid human observers with poor emotion inference capabilities.

Awards Won: