## Thesis Title and Abstract

## Title

Computational Methods for Modelling and Explaining Perceived Emotions in Music

## Abstract

The rapid increase in musical content in various social media and other platforms necessitates data-driven computational modeling of perceived emotions to facilitate applications like music emotion recognition (MER). The description, measurement, collection, processing, and storage of perceived-emotion opinion data form an essential part of such studies, achieved using various emotion representations like discrete and dimensional models. Though significant studies exist based on these models on various music traditions, their effectiveness on Hindustani classical music (HCM) is relatively less studied. On the other hand, the Nava Rasa concept in Indian aesthetics provides emotion categories to denote aesthetic emotions, which have been explored sparsely in MER. In this thesis, first, we study an intensity-based, categorical emotion representation called the Emotion-word and Intensity-Value (EWIV) representation, where the emotion-words are taken from the Nava Rasa concept. We demonstrate the effectiveness of EWIV and validate the quality of self-reported emotions, on existing benchmark clip sets, and a newly introduced set of clips from HCM called the *EmoRaga* clip set for perceived emotion analysis in HCM. We also discuss representativeness of EWIV using goodness-of-fit measures for statistical models as our metric and demonstrate that it might provide a better fit for perceived emotion data under certain conditions.

Context is one of the key parameters that influence music emotion perception in listeners. We utilize excerpts from the EmoRaga clip-set to explore the influence of musical context on perceived emotions. We demonstrate that change in *immediate intrinsic musical context* changes the perception of musical excerpts, and term this phenomenon *intra-contextual influence*. Using EWIV emotion representation, we show how patterns of such influence emerge for dominantly happy and sad Sitar excerpts. This contributes to the modeling of the subtle nuanced variations in music-perceived emotions with different improvisations of the same music piece.

The dependence of perceived emotions from music on temporally distributed music segments makes most machine learning methods for MER unreliable, calling for human-understandable explanations of the model predictions. We present an attentive-LSTM-based, explainable dynamic emotion prediction model and show that it performs better than existing models on a benchmark dataset, using a benchmark feature set. We also demonstrate that a reduced feature set consisting of Spectral features gives comparable results. We apply the best models to the EmoRaga clip-set to successfully perform dynamic dominant and secondary emotion classifications, music emotion variation detection, and identification of the music segments with high probabilities of perceiving the dominant emotion. These studies also demonstrate the applicability of EWIV representations estimated from self-reported emotion data towards these MER applications. We compare the model-predicted important segments with those annotated with emotion motifs by experts, which yields significant overlap, demonstrating that these are indeed captured by the model, explaining the emotion predictions.

**Keywords**: Explainable music emotion recognition, emotion representation, MER in Hindustani classical music, emotion motif, intra-contextual influence