In this paper, we focus on the automatic classification of jazz records. We propose a novel approach where we break down the ambiguous task, which is commonly referred to as genre classification, into three more specific semantic levels.

First, the rhythm feel (swing, latin, funk, two-beat) characterizes the basic groove organization and most often relates to the rhythm section. Second, the tonality type (functional, blues, bebop harmony) describes the general harmonic organization of the underlying composition, which serves the soloist as a guideline for the improvisation. Third, the tempo class (slow, medium, up) provides a rather broad categorization of the overall tempo of a song. We consider three individual classification tasks with respect to the different descriptors. A set of 229 jazz recordings was selected and labeled by musicology and jazz students as part of the Jazzomat Research Project.

For the three tasks, we conduct several classification experiments in order to investigate the usefulness of certain features and feature groups in general as well as pre-processing methods such as feature selection and feature space transformation. For a baseline classification experiment, we use low-level audio features which are widely used in Music Information Retrieval (MIR) research. As second step, we test several recently proposed mid-level features for modeling rhythmic or harmonic content of audio data. We systematically evaluate different modifications and combinations of these features.

By testing decision trees as classifiers, the obtained decision rules provide an opportunity for a musicological interpretation of the classification process. Beyond that, we also investigate which features were selected for the different tasks by an automatic feature selection algorithm. The results of this interdisciplinary study have potential implications for jazz research as well as for content-based audio analysis tasks such as music similarity search and music recommendation in general.