Table 1: Descriptions of the features

Feature	Description
	Mid-level Features
Melodiousness/Melody	To which excerpt do you feel like singing along?
Articulation	Which has more sounds with staccato articulation?
Rhythmic complexity	Is it difficult to repeat by tapping? Is it difficult to find the meter? Does the rhythm have many layers?
Rhythmic stability	Imagine marching along with music. Which is easier to march along with?
Dissonance	Which excerpt has noisier timbre? Has more dissonant intervals (tritones, seconds, etc.)?
Tonal Stability/Atonality	Where is it easier to determine the tonic and key? In which excerpt are there more modulations?
Modality/Mode	Imagine accompanying this song with chords. Which song would have more minor chords?
	Signal Processing Features
Danceability	Describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity
Loudness	The overall loudness of a track in decibels (dB). Loudness values are averaged across the entire track and are useful for comparing relative loudness of tracks.
Chord Change Rate	Bigger number means more chord changes
Dynamic Complexity	This algorithm computes the dynamic complexity defined as the average absolute deviation from the global loudness level estimate on the dB scale. It is related to the dynamic range and to the amount of fluctuation in loudness present in a recording.
Zero Crossing Rate	This algorithm computes the zero-crossing rate of an audio signal. It is the number of sign changes between consecutive signal values divided by the total number of values. Noisy signals tend to have higher zero-crossing rate. It can be used for voice detection or percussive sound detection (e.g a snare hit)
Chords Number Rate	The ratio of different chords from the total number of chords in the progression
Pitch Salience	How easy a human would clearly identify something as having a certain pitch
Spectral Centroid	The spectral centroid is the center of ‘gravity’ of the spectrum (e.g. it can be observed that higher values correspond to brighter sounds)
Spectral Complexity	The spectral complexity is based on the number of peaks in the input spectrum
Spectral Decrease	Spectral decrease represents the amount of decrease of the spectrum, while emphasizing the slopes of the lower frequencies
Spectral Entropy	The entropy can be used to capture the peakiness of the spectral representation. The resulting feature is low for a rate-map with many distinct spectral peaks and high for a flat rate-map spectrum
Spectral Flux	The spectral flux evaluates the temporal variation of the logarithmically-scaled rate-map across adjacent frames. It has been suggested to be useful for the distinction of music and speech signals, since music has a higher rate of change
Spectral Kurtosis	It can detect musical noise
Spectral Roll Off	Determines the frequency in Hz below which a pre-defined percentage of the total spectral energy is concentrated
Spectral Spread	Noise-like signals have usually a large spectral spread, while individual tonal sounds with isolated peaks will result in a low spectral spread
Onset Rate	How many distinguished notes exist per minute
Length	The length of the music track
Beats Loudness	How much loud is each beat
Spectral RMS	RMS for music shows the average loudness of the audio
Spectral Skewness	The feature will be zero for silent segments and high for voiced speech where substantial energy is present around the fundamental frequency
Spectral Energybands	This algorithm computes energy in a given frequency band of a spectrum including both start and stop cutoff frequencies
BPM	Beats per minute
Vocal - Instrumental	Instrumental if the track has no vocals and vocal if the track has vocals