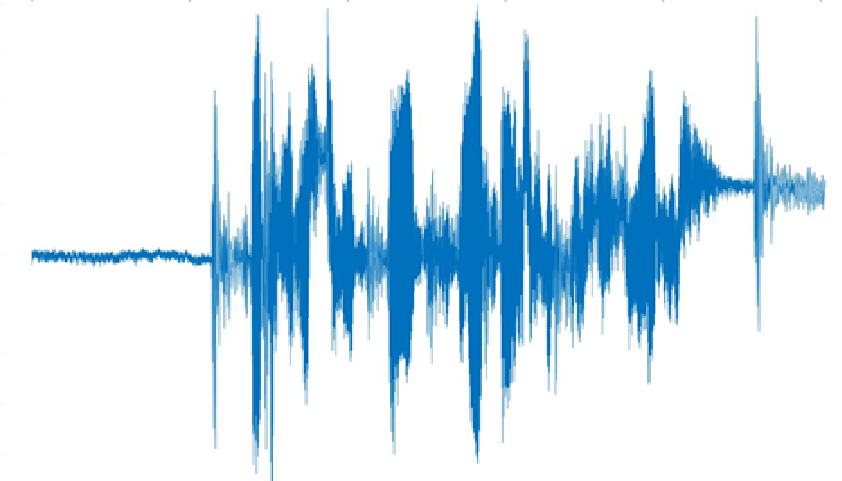
Chapter 4: Data Processing and Feature Extraction

This chapter explains the process adopted for data extraction and processing. The reference dataset used for modelling is DS\_1 (ref: TBD) which contains sounds classified into 4 classes viz Q1, Q2, Q3 and Q4. Each class corresponds to each quadrant of Russell’s Circumplex Model of emotion. The objective of the data extraction and processing approach is to determine the best dataset that can aid modelling based on data sampling, data scaling, anomaly detection and feature inclusion / exclusion.

Intuition

There are three dimensions to the characteristics of a sound – time dimension, frequency dimension and the strength or magnitude of the signal. A sound wave is typically perceived as a time variant signal and it is easy to visualize the time variability (on a coordinate plane x-axis represents the time and y-axis represents the magnitude). However, there’s a frequency axis that also varies with time (represented by z-axis in a coordinate plane) indicating that for a single unit of propagation of time along x-axis, there are several frequencies that are activated with varying degree of strength or magnitude. This characteristic make the sound “time-variant” and “frequency-variant” as illustrated in the figure below (Fig number: TBD). This characteristic also differentiates a sound signal from a normal signal where the frequency component is comparatively constant.



Original Sound signal

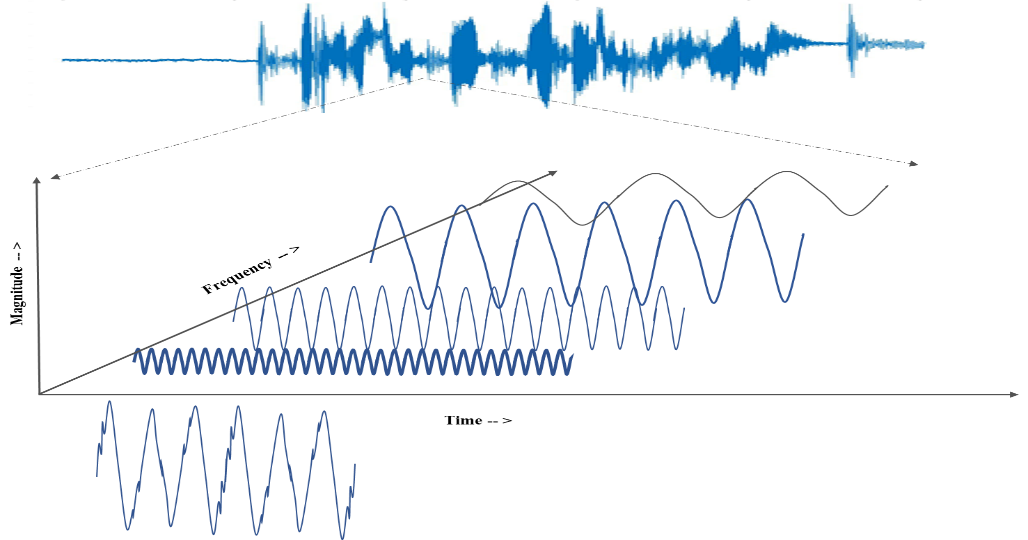


Illustration of a small window of sound signal into its frequency and time variant components (not to true scale)

In case of musical sounds, the presence of frequency components dramatically increases driven by the timbre of predominant sound (in case of monophonic sound) and each musical instrument in use (in case of polyphonic sound). This increase in frequency components imposes the need to extract appropriate frequencies from the sound along with its time variant characteristics. Based on this intuition, following features were extracted from the musical sounds in the dataset.

Feature Extraction

Data Sampling

Data Scaling

Anomaly Detection

Feature inclusion / exclusion