**Problem Statement:**

The aim of our project is to identify the genre of a song based on the audio input.Extracting vocals and the instrumentals of a song separately. It can be extended to speech separation.

Separation is usually carried out in the time-frequency domain, recovering the short-time Fourier transformation (STFT) of the source signals for each time frame t and frequency given the STFT of the mixed speech. We also want to build a recommendation system for the user based on his current song choices history.

**Milestones:**

1) Speech to text conversion

2) genre detection based on lyrics

3) End to end genre classification based on music audio only

4) recommendation system for related songs (If time permits)

**Data Set:**

CCMixter (50 mixes) - vocal & background track

http://www.loria.fr/~aliutkus/kam/

MIR-1k (1000 excerpts) - vocal & background track

<https://sites.google.com/site/unvoicedsoundseparation/mir-1k>

\*We are planning to create are our own dataset if we the existing dataset don’t work