**Music Recommender System**

Technical Project Proposal

Automated Learning and Data Analysis

Ajay Ramani

Bhargava Pejakala Kakarnaya

Kristy Patel

Sandesh Veerapur

Shini Arora

# MUSIC RECOMMENDER SYSTEM

The Music Recommender System aims to predict which songs a user will listen to. The idea is to recommend the most popular songs to the user learning from his listening history.

This system makes use of the Million Song Dataset (MSD) which is a freely-available collection of audio features and metadata for a million users. The data available in MSD is in the form of <user ID, song ID, play count> triplets. The full listening history is available for about a million users which serves as the training set and the validation and test sets contain half of the listening history of about 110K users.

Last.fm dataset and musiXmatch dataset will be used in collaboration with MSD. Last.fm dataset is a collection of song-level tags and pre-computed song-level similarity for songs in MSD. musiXmatch dataset provides lyrics of MSD tracks in a bag-of-words format.

**References**:

1. <http://www.kaggle.com/c/msdchallenge>
2. <http://labrosa.ee.columbia.edu/millionsong/>
3. Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. In Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011