|  |
| --- |
| AMPBA-Class of 2020 Winter  Term IV |
|  |
| ***Team: Group-1***  Unsupervised Machine Learning-II  Project Progress Report  ***Karthik Shankar S (PGID: 11920008)***  ***Pundareek Chandrashekhar (PGID: 11920049)***  ***Rajesh Rangaswamy (PGID: 11920072)***  ***Praveenkumar Agoorkaisetty (PGID: 11920096)*** |

**

***Project Goals and Objectives***

*Listening to music is obviously one’s personal choice. But with the arrival of smart gadgets and availability of popular online entertainment channels, listening to music no longer is completely one’s personal choice rather it can be rephrased like, a listener can now be coerced or recommended to listen a song based on their listening history in an entertainment ecosystem.*

*By listening history means, a song listened by an user can be tagged to ‘n’ number of attributes for suggesting any other user with same or similar taste profile.*

*The Project titled* ***“Song Recommendation System”*** *is an attempt to our Assignment Fulfilment for Unsupervised Learning-II as part of the Academic Curriculum.*

***Goal of the Project***

*To devise an Unsupervised Machine Learning model which is expected to recommend a song to an user, based on collaborative filtering techniques.*

***Objective of the Project***

*Objective of the attempt is to develop a Proof of Concept (PoC) for the Song Recommendation System by training and testing the model with the subset data first hand.*

*We aim to use machine learning to suggest or recommend a song based on both user and song characteristics with the metadata about a song (such as user listened and # of times listened) From the models, we hope to also get a deeper insight into the features that are most suggestive of a song popularity, and understand what makes certain songs more popular than others.*

*Our problem statement is thus*

1. *How to suggest or recommend a song based on user listening history?*
2. *What would be the song recommendation for a new user based on the model being developed?*
3. *What type of modelling technique would yield a better song recommendation or suggestion?*

*We propose to define success of this project by analysing the metric which would be obtained from the ROC Curves later.*

*The Million Song Data Set (MSDS) is a freely available collection of audio feature and one of the extensive database of popular contemporary tracks spanning decades of western music. With the advent of high end communication technologies coupled with speed and coverage, every one is connected. The music industry (Sony Music Entertainment, Universal Music Group, Warner Music Group, etc.) is highly invested in identifying trending features, and would be especially interested in an algorithmic approach to evaluating the potential popularity of a new song.*

*Objective is Song Recommendation.*

*Goals:*

*SMART*

*Data Collection Steps*

*Why this data suits to our Project Goals*

*Basic Analysis of Data (Descriptive Statistics, Visualization)*

*Preliminary Analysis*

*Description of the Way Forward.*