Recommendation Systems Milestone 1

Karan Uppal

Problem Definition

➢ Why is this problem important to solve?

Creating a recommendation system from a music dataset

➢ What is the intended goal?

Extract the recommended data points from the large dataset.

➢ What are key questions that need to be answered?

Can a item-item recommendation be best suited for this type of project to project which song will be popular or can other systems be used.

➢ What is that we are solving using data science?

Getting recommendation systems based on

desired traits from the dataset given

Data Exploration

➢ Data Description - What is the background of this data? What does it contain?

The data consists of user profiles that shows the songs they listened to and how many times. It includes contents such as titles of the songs, artists, release of song, User ID and Song ID(used to merge data sets).

➢ Observations & Insights - What are some key patterns in the data? What does it mean for the problem formulation? Are there any data treatments or pre-processing required?

Most popular songs turn out in the 21st century which may be due to technological influence surrounding the time of which those songs were released. Pre-processing the data included making sure Song ID information was matched with both data sets and excluded irrelevant information like creator and artist information.

Proposed approach

➢ Potential techniques - What different techniques should be explored?

Using classification of what genre of music using clustering methods maybe an advisable method to organizing and making better recommendations using this dataset.

➢ Overall solution design - What is the potential solution design?

Using new potential techniques as classification would need hard Data analysis and full visualization of principal component analysis to see any relationships within the data that may be revealing.

➢ Measures of success - What are the key measures of success to compare potential techniques?

Statistical relationships that are significant between item-item recommendation systems and a forecast of popular music.