
# Project Deliverable 1: Group Formation and Project Understanding

Group 4

Members - Dhruv Jani, Dhruvil Patel, Karan Issrani, Mohit Gupta, Yaw Frempong

Communication plan to include project artifact repository - git repository

We plan to meet every week and divide tasks which can be parallelized to ensure completing deadline specific tasks. We will meet at the Atkins library and zoom whenever possible.
Our links
Zoom link - https://uncc.zoom.us/j/6026928233
GitHub Repository - https://github.com/kissrani/ITCS6100

Our tentative plan looks like the below table

| Dates | Communication Plan |
|---|---|
| 04/13/2022 | Creating a Git Repository and studying the dataset. |
| 04/17/2022 | Studying the background and Setup |
| 04/18/2022 | Data Preparation |
| 04/24/2022 - 04/30/2022 | Creating a solution and Choosing the algorithm. |
| 05/1/2022 - 05/3/2022 | Final Changes |
| 05/5/2022 | Submission |

2) Selection of data to analyse from the Open Data Registry for Amazon Web Services

This dataset comprises movie ratings and has 1 lakh entries. This is prescribed by AWS
https://grouplens.org/datasets/movielens/

3) Business Problem or Opportunity, Domain Knowledge (link to information on domain relative to data, problem or opportunity)
Business Problem/Opportunity - Recommendation algorithms are at the core of the Netflix, Prime and Hulu products. They provide our members with personalised suggestions **to reduce the amount of time and frustration to find something great content to watch**.
Link - **https://research.netflix.com/research-area/recommendations**

Domain Knowledge - Recommender systems are Machine Learning techniques that serve the best advice for a potential buyer. They suggest the most relevant items to buy and, as a result, increase a company's revenue. These suggestions are based on users' behaviour and history that contain information on their past preferences.

4) Research Objectives and Question(s) (what you are trying to describe or predict with the data)

We plan to preprocess the data and design our own recommendation system, choosing our algorithm.

We are trying to describe the taste in movies for an individual based on the previous observations we obtain a trend and predict a movie based on the data we have. For this, we will make use of ML Algorithms in AWS Sagemaker. By doing this we aim to answer the following questions:

What did the user like in the past ?(describe the user)

What will the user like in the future? (Prediction)