

## **Netflix Recommendation System**

Nisarg Vinchhi - nvinchh

Rohit Mandge - rnmandge

Vaibhav Rajadhyaksha - vrajadh

### **Abstract:**

Our application will analyze the existing Netflix data to identify the relationship between movies and the user based on the ratings. We will utilize this relationship to suggest new movies to the user based on his preferences. We plan to use collaborative filtering techniques or similarity matrix to identify the relationships between the users and ratings. Our application will also be predicting the rating a user might give to the movie based on his previous ratings and ratings given to the movie by similar users. The Netflix Prize dataset consists of around 100 million records. Each movie name is mapped to a movie ID. For each movie ID a separate record is available which is of the form “userid, rating, time”. Userid is an integer and rating defined as an integral scale from 1 to 5. Date is represented as a timestamp. We plan to replicate this dataset in order to generate a larger dataset that we can utilize to work with and test large scale data computing.