

# **Summary of project**

## ***Datasets:***

1. Primary dataset - Amazon - Movies and TV Ratings.csv

## ***Tasks:***

### **Exploratory Data Analysis:**

1. Which movies have maximum views/ratings?
2. What is the average rating for each movie? Define the top 5 movies with the maximum ratings.
3. Define the top 5 movies with the least audience.

### **Recommendation Model:**

Some of the movies hadn't been watched and therefore, are not rated by the users. Netflix would like to take this as an opportunity and build a machine learning recommendation algorithm which provides the ratings for each of the users.

1. Divide the data into training and test data
2. Build a recommendation model on training data
3. Make predictions on the test data

## ***Steps involved:***

1. Pandas library has been used to import data from datafile and doing exploratory data analysis.
2. Multiple algorithms like Graph Neural Network, Apriori algorithm(association rule) and SVD algorithm from Surprise library can be used to build a recommender system. Here, SVD from Surprise library has been used to split the data in train and test set and later fitting the model to predict the recommendation for test data. TSNE class from sklearn.manifold has been used to plot the data with the recommendations or movie associations.