**Proposal**

***Subject of our Project : Movie Recommender (Based on Movie notations).***

Our dataset is composed of several files. The first file contains details on the movies (movieID, Movie titles and Genres). The second file contains the ratings data (userdID, MovieID and ratings ). Ratings are made on a 5-star scale, with half-star increments (0.5 stars - 5.0 stars).

This dataset contains 2,500,0095 ratings on 62,423 movies by 162,541 users. Users were selected at random . All selected users had rated at least 20 movies .These data were created between January 09, 1995 and November 21, 2019.

The features of our model are the following :

UserID, MovieID, Movie Title, Genres and Ratings.

Based on the ratings that a User gave on several movies (at least 20) , we want to predict the rating that this very same user would give on a movie (or several movies) he never saw (from the dataset).

Furthermore, we want to create a list of 20 potential movies such that , given a user from the dataset, the list will contain movie names that the user would likely give a good rate.

We will use a combination of both Linear and Logistic Regression models in order to classify the movie ratings.