

**Name: Sravanthi Kapu**

**Project Title: Movie recommendation System**

**Project Objective:** To build a recommendation system which recommends movies to a user, which he may be interested in watching based on his previous ratings and other users rating. Implementation to find the movies which user might be interested will be done by using collaborative filtering

**Project description:** In a movie recommendation system, many users rating's for many movies is analysed. By comparing the users who are having similar preferences, movies can be recommended to a user based on the ratings and genres liked by the user. The approach called Collaborative filtering is used in order to achieve this.

In this approach, if the user rating for a particular movie is missing, that is predicted by analysing the previous data. The matrix of users and rating for movies is filled by analysing the user pairs which are similar. These pairs are called latent factors in Alternate least squares algorithm, By using this algorithm and latent factors we filling in the missing values in a matrix in order to achieve our suggestion list for the user. Here we train our system and then recommend the movies. By using spark, we read the files in the dataset and create RDD's to analyse and work on the data. RDDA's provide faster way for iterative algorithms to execute. Resilient distributed dataset (RDD) is a fault-tolerant collection of elements that can be operated on in parallel on our cluster which makes our recommendation system fast.

Dataset: A dataset with all the data related to movies, rankings, genre, user id and other fields is taken.

<https://grouplens.org/datasets/movielens/> or <https://www.kaggle.com/orgesleka/imdbmovies>