Please upload your progress report to the Github repo shared on CMT. The progress report should give us an idea of how you're implementing your proposal. It should answer 3 main questions:

1) Which tasks have been completed?

* Scrape data from IMDB API.
* Decided what data we needed to subtract from the dataset, which included **title**, **director**, **genre**, **region**, **average rating**, **main casts**, and **adult/non-adult title**.
* Set up a form of movies using the data scraped from IMDB API (Movie Title : Basic Information)
* Set up our Mysql database and upload our data on AWS.

2) Which tasks are pending?

* Research on recommendation algorithms based on users’ similarity
* Implement recommendation algorithms
* Boolean form of users (User : Movie Title) indicates whether (1 for yes; 0 for no) a specific user watched the movie.

3) Are you facing any challenges?

* We are stuck at how to quantify similarity
  + How to define the dimension of our vector space
  + How to position users’ vectors in vector space
  + How to calculate similarity between vectors
* How to find real user data to improve
  + First way to achieve: deep scrapy to obtain information about user having watched movies or not
  + Second way to achieve:let user use our website and label movies to “watched” or “not watched”