**Documentation**

In this notebook, We attempted at implementing a few recommendation algorithms (content based and popularity based) and try to build an ensemble of these models to come up with our final recommendation system. With us, we have MovieLens datasets.

We used the IMDB Ratings to come up with our **Top Movies Chart.** We used IMDB's *weighted rating* formula to construct my chart. We sorted them in decreasing order of this weighted rating to get top charts. By taking input as Genre we filtered top movies of that Genre.

To personalize our recommendations more, We built an engine that computes similarity between movies based on certain metrics and suggests movies that are most similar to a particular movie that a user liked. Since we used movie metadata (or content) to build this engine, this also known as **Content Based Filtering.**

We first tried to built a recommender using movie descriptions and taglines. We do not have a quantitative metric to judge our machine's performance so this will have to be done qualitatively using cosine similarity.

To build our standard metadata-based content recommender, we merged our current dataset with the crew and the keyword datasets.

What we planned on doing is creating a dataset for every movie which consists of **genres, director, main actors and keywords.** We then used a **Count Vectorizer** to create our count matrix as we did in the Description Recommender. The remaining steps are similar to what we did earlier: we calculate the cosine similarities and return movies that are most similar.

We then did a small amount of pre-processing of our keywords before putting them to any use. As a first step, we calculated the frequency counts of every keyword that appears in the dataset.

For example, when we search for Dark Knight, The recommendations seem to have recognized other Christopher Nolan movies (due to the high weightage given to director) and put them as top recommendations. We enjoyed watching **The Dark Knight** as well as some of the other ones in the list including **Batman Begins**, **The Prestige** and **The Dark Knight Rises**.

One thing that we notice about our recommendation system is that it recommends movies regardless of ratings and popularity. It is true that **Batman and Robin** has a lot of similar characters as compared to **The Dark Knight** but it was a terrible movie that shouldn't be recommended to anyone.

Therefore, we added a mechanism to remove bad movies and return movies which are popular and have had a good critical response.