Movie Recommendation

# **Sanjay Jaras DSC680 Applied Data Science GitHub Portfolio URL:** <https://sanjayjaras.github.io/>

# Domain and Background

Recommendation engines or recommender systems filter a large list of products or things to present only products or things that customers might be interested in. There are many applications of recommender systems like movies, series, various products on e-commerce sites, etc. When we search the similar movies on google, google uses a recommendation system in the background to find a similar movies list. There are four types of recommendation engines:

1. Content-based: This recommender system uses attributes of the products or movies to find the ranking and similar products and movies.
2. Collaborative filtering based: This recommender system uses different algorithms to find similar users by some activities they performed like the movie rating, product likes, or movie reviews.
3. Popularity-based: This type of recommender system will suggest products or movies depending on the popularity.
4. Hybrid: These recommender systems are created by using two or more types of the above-mentioned recommender systems.

References:

[1] Beginner Tutorial: Recommender Systems in Python – Aditya Sharma – May 2020 - <https://www.datacamp.com/community/tutorials/recommender-systems-python>

[2] Movies Recommendation - Şevval Yurtekin - <https://www.kaggle.com/sevvalyurtekin/movies-recommendation>

[3] Recommender system - <https://en.wikipedia.org/wiki/Recommender_system#:~:text=A%20recommender%20system%2C%20or%20a,would%20give%20to%20an%20item>.

[4] Introduction to recommender systems – Baptiste Rocca – Jun 2019 - <https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada>

[5] Recommendation systems: Principles, methods, and evaluation - F.O.IsinkayeaY.O.FolajimibB.A.Ojokoh – Nov 2015 - <https://www.sciencedirect.com/science/article/pii/S1110866515000341>

[6] Building a Movie Recommendation Engine in Python using Scikit-Learn – Suman Adhikari – Feb 2019 - <https://medium.com/@sumanadhikari/building-a-movie-recommendation-engine-using-scikit-learn-8dbb11c5aa4b>

[7] Build A Movie Recommendation Engine Using Python -   
Aug 2019 - <https://medium.com/analytics-vidhya/build-a-movie-recommendation-engine-using-python-scikit-learn-machine-learning-e68ba297e163>

[8] Movies Recommender System - Rounak Banik - <https://www.kaggle.com/rounakbanik/movie-recommender-systems>

[9] recommendation through movie\_lens\_rating - Karan Choudhary - <https://www.kaggle.com/karanchoudhary103/recommendation-through-movie-lens-rating?select=tag.csv>

[10] An In-Depth Guide to How Recommender Systems Work - Badreesh Shetty – Jul 2019 - <https://builtin.com/data-science/recommender-systems>

# Data Understanding

This dataset is made available by <https://movielens.org/>. The dataset contains a 5-star rating and free-text tags from MovieLense. It contains 25 million ratings and more than 1 million tag applications for 62 thousand movies. This data was created by 162 thousand users between 1996 and 2019. Every selected user has rated a minimum of 20 movies.

Rating file structure

1. userId
2. movieId
3. rating
4. timestamp

Tags file structure

1. userId
2. movieId
3. tag
4. timestamp

Movies file structure

1. movieId
2. title
3. genres: these are pipe separated

<https://grouplens.org/datasets/movielens/>

# Research Questions:

The number of movies available on platforms like youtube, Netflix, Hulu, Disney is increasing each day. With the increasing number of available movies, viewers spend a lot of time searching for movies of their likings. A good recommendation engine can save a lot of time of all the viewer that are looking for movies of their likings. A good recommender system will not only save the time of searching for good movies but also avoid viewers wasting time watching trailers or movies for some time and then stopping. Without a good recommendation system, a user might leave the platform if they cannot find a product or content of his liking. If we use the hybrid type of recommendation engine, it can drastically improve suggested recommendations.

# Method:

I will be starting with a Collaborative filtering-based recommendation engine. After a collaborative filtering-based recommendation engine, I will try to work on a hybrid recommendation engine. For the hybrid recommendation engine, I will use collaborative filter-based and content-based recommendation engines together. With a hybrid recommendation system, I am expecting better results than any one type of recommendation engine.

# Potential Issues:

I have not worked on a recommendation engine before, so there will be some learning curve before actually training a model. The data size is large, so, I might face some hardware limitations while training models. I am not sure, how much success, I can achieve in 3 weeks of project time. In this limited time, I might not able to complete a Hybrid type of recommendation engine.

# Concluding Remarks:

On the internet and media platforms like Netflix, youtube it is very to get lost while searching for products or content as per our likings. If the platform does not suggest any products or content, we might waste a lot of our time searching for products or content that we are interested in. Recommender systems are machine learning systems that help us discover new products, media content, and services depending on their earlier activities. A good recommendation system not only saves users/customers time but also keeps him/her engage with the platform. Without a good recommendation system, a user might leave the platform if they cannot find a product or content of his/her liking. With a good recommendation system, we can keep users/customers happy and engaged, so that they will keep using a platform for a longer duration.