

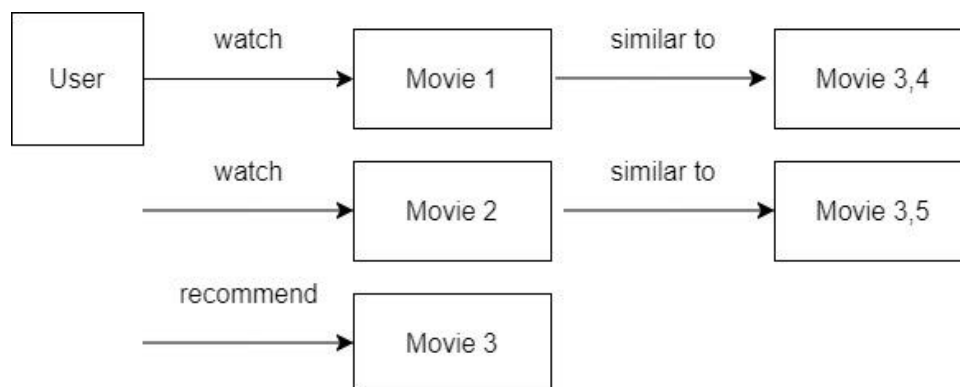
Movie Recommendation systems

In this assignment we will be creating a different recommender model which will result in a different kind of recommendation.

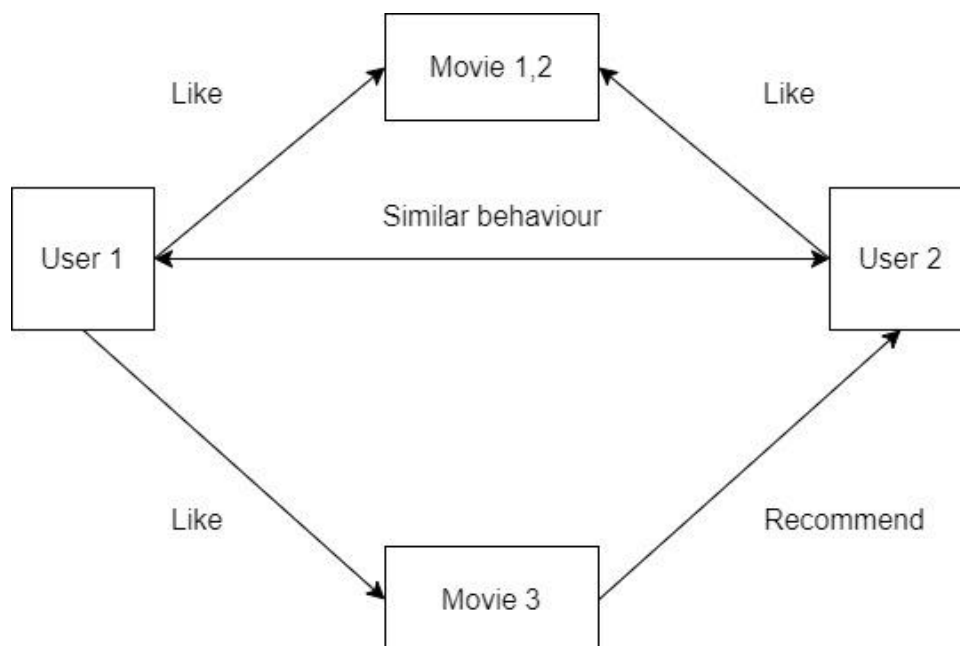
Let first address the type of recommender systems.

A recommendation system typically fits in either one of the two categories which is

- Content-based system: This method will base the recommendations on the similarity of the movies attributes such as genres, main actor, producer and other factors that are relevant. For example, if a user watches a sci-fi movie like Iron man, the system will likely recommend some sci-fi movie that are related to Iron man such as The Avengers.



- Collaborative Filtering system: This method, which became popular during the Netflix Prize challenge in 2006, bases the recommendation on user interaction with the movies. It will find the similarities in the user and how they rated the movies that they had watched. Then try to find which user is more likely to behave the same way as the user we are considering and base the recommendation on that data.



Implementation

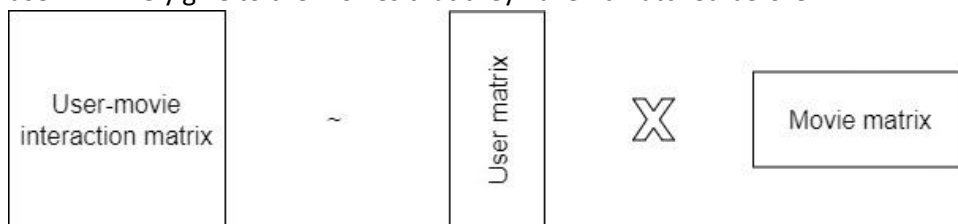
The goal is to recommend 5 movies for the first 5 users.

With a condition that the movies we recommended must be the one they haven't seen before.

We have tried using two methods, which resulted in a different recommendation.

- Content-based recommendations using nearest neighbours algorithm
 - We based our recommendations on how similar the movies are to each other in terms of genres.
 - We measured the similarity using cosine similarity.
 - The strength of this method is how easy it can be implemented, and the fact that it can recommend the movie that is close, or almost the same kind of movie with the one it was based on.
 - The weakness of this method is that although it can capture a really good relationship with the same kind of movies, it doesn't take into account the fact that even the movies in the same genres may have a different story telling strategy and the content of the movie might be different. So, it can recommend a movie with a bad rating but in the same genres, which is not ideal.
 - Since this recommendation is based on the movies that the users had watched, we create a recommendation by ranking the movies that the user have rated first. And we randomly select one of the movies that in the list of highest rating for that user as a base of our recommendation system.
 - The result for the first 5 users are
 - Recommendations for Vincent based on similar genres with Stone Cold:
 - 1: The Amazing Catfish
 - 2: House of Flying Daggers
 - 3: Ghost Dog: The Way of the Samurai
 - 4: Cleopatra
 - 5: Get Hard
 - Recommendations for Edgar based on similar genres with Fahrenheit 9/11:
 - 1: Scrooged
 - 2: Blade
 - 3: Wolves
 - 4: Contagion
 - 5: Half Past Dead
 - Recommendations for Addilyn based on similar genres with That Thing You Do!:
 - 1: Collateral
 - 2: Mozart's Sister
 - 3: Survivor
 - 4: Adam Resurrected
 - 5: La Famille Bélier

- Recommendations for Marlee based on similar genres with Now You See Me 2:
 - 1: I Love You Phillip Morris
 - 2: Original Sin
 - 3: Pirates of the Caribbean: Dead Man's Chest
 - 4: Life During Wartime
 - 5: Mutual Appreciation
- Recommendations for Javier based on similar genres with The Emperor's New Groove:
 - 1: Palo Alto
 - 2: Monsters, Inc.
 - 3: Quarantine
 - 4: The SpongeBob SquarePants Movie
 - 5: Valentine's Day
 -
- Collaborative filtering system using matrix factorization
 - We based our recommendations on user-movies interaction, where we use the user ratings for each movie as a starting point. By decomposing the user-movie interaction matrix into the product of two lower dimensional matrices. The process is implemented in a way that when we combine those two lower dimensional matrices we can predict the ratings that the user will likely give to the movies that they haven't watched before



- The strength of this method is the fact that it takes a human preference into account and trying to recommend a movies for a similar kind of users based on their behaviour. And for the matrix factorization method, it does work fairly well with a sparse data which are unknown to us, and it could work even though we don't have any details on the movies.
- The weakness of this method is that if formulated badly, it might take a long time to compute and as the number of users increases it can be hard to scale. Furthermore, while it can capture the trend of the data, we a mere human can have a hard time trying to interpret the pattern and might not be able to identify why does the system behave this way.
- Since this recommendation is based on the predicted ratings that the user might give to the movie, we recommend the movies by choosing the one that have the highest predicted ratings that the users haven't seen before.
- The result for the first 5 users are
 - Recommendation for Vincent
 - 1.The Dangerous Lives of Altar Boys
 - 2.The Importance of Being Earnest
 - 3.Risen
 - 4.The Curious Case of Benjamin Button
 - 5.Across the Universe
 - Recommendation for Edgar
 - 1.The Other End of the Line
 - 2.A Scanner Darkly
 - 3.Fahrenheit 9/11
 - 4.New York, New York
 - 5.Good Kill

- Recommendation for Addilyn
 - 1.Dysfunctional Friends
 - 2.ATL
 - 3.The Magic Sword: Quest for Camelot
 - 4.Lilyhammer
 - 5.That Thing You Do!
- Recommendation for Marlee
 - 1.Now You See Me 2
 - 2.The Cabin in the Woods
 - 3.Old Joy
 - 4.Lilyhammer
 - 5.When Harry Met Sally...
- Recommendation for Javier
 - 1.The Best Exotic Marigold Hotel
 - 2.George and the Dragon
 - 3.Perrier's Bounty
 - 4.Adventureland
 - 5.Dear John

Even though we have created a usable recommendation system, it still far from the deployment quality. To further asses the quality of this system we must find a way to evaluate user satisfaction based on our recommendations and using that data to improve our system further.

And to evaluate user satisfaction is another challenging task, it can be hard to determine the satisfaction of the user given the fact that most users might not even rate the movies that they had watched. And some matrices such as watch time can't really determine how much the user enjoys the movie either. So, this is another challenging task that will need to be consider in the further development of the system.