Movie Recommendation Using Social Network Analysis

Riya Shetty
Master's in Data Science
Indiana University - Bloomington.

Ruchik Dama
Master's in Data Science
Indiana University - Bloomington.

Gautam Chauhan
Master's in Data Science
Indiana University - Bloomington.

Abstract - Movies are a good example of a soothing pastime. People spend their days scrolling through various OTT platform watching diverse movies, some of which relaxes and calms our thoughts. As a result, in this project, we'd like to propose a method in which the movie player application recommends the most popular movies that display the social network's trend. We use social network analysis to provide better suggestions by constructing graphs representing the relationships between users and the movies. Moreover, this study will help us understand why social network analysis will provide better suggestions since the project will attempt to grasp the value of every node in the network.

Keywords - Movie Recommendation System, Collaborative Filtering, Social Network analysis, Content-based filtering, Hybrid Systems.

1. Introduction

A social network is a collection of interconnected entities that are typically represented as a graph, with nodes describing the entities and links describing their relationships. A large part of how a person lives is determined by how they are linked to a larger social network. Recommender systems are software tools that recommend products or items that are appropriate for a customer's taste.

The problem arises when a new user enters the system, for such users that the system does not have enough information to make a decision. This is known as the Cold Start Alleviation problem. We have seen that in previous projects a movie recommender system is capable to give recommendations based on the user profile that is his/her liking but sometimes the generated data may turn out to be inefficient. We make use of social network analysis in this scenario where we gather information from various social networking platforms of the user. The user information includes age, gender, location, most liked genre, et al. This can be implemented using content-based filtering which is described in the upcoming sections.

2. Literature Survey

[1] Johann Stan et al. Study sheds some light on Recommender systems and social networks along with their challenges.

[2] Dolgikh, Dmitry & Jelinek, Ivan. proposed a novel method by taking into consideration the individual traits of a particular user.

[3] Xinchang Khamphaphone et al. identified communities based on the movies that the users have watched and used betweenness centrality and identified the Largest Cumulative modularity score. Group Modifications were further done by cosine similarity and collaborative filtering.

[4] Chang Li et al. proposed a graph-based method to predict users' attitudes towards unseen songs. A graph model was created from users and songs, further communities were formed based on the popularity of the songs. The final model was tested on the number of likes that song received, the number of paths between the song and the users, and the rate of path number and song liked.

[5] B. Fields et al. presented an extensive analysis of a sample of a social network of movies. Hybrid graphs and distance measures were constructed and used to examine the community structure of the artist network. Finally, the results of these investigations were shown to be mostly orthogonal between these distance spaces for Recommendation and Discovery.

3. Overview of the Proposed System

Aim of the proposed theory

To propose a theory to develop a movie recommender system using the techniques of collaborative filtering and content-based filtering. For the scope of this proposal, we have put forward a system that can recommend movies using hybrid systems and develop the visualizations for the same.
In this part, we'll go over

1. Recommender Systems
    a. Collaborative filtering
    b. Content-Based Filtering
    c. Hybrid Recommender System
2. Social Network Analysis

SNA can be utilized with a recommendation system that can not only recommend movies but also be more widely used in the business than recommendation systems that employ collaborative filtering. In this research, we intend to propose a theory for a movie

recommendation system that solves the problem by combining social network analysis with hybrid recommender systems. This eliminates any potential for confusion while using the hybrid recommender systems..

1. Recommender Systems:
   Recommender systems are in use across a number of industries, including playlist generators for video and movie services, product recommendations for online businesses, content recommendations for social network platforms, and open web content recommendations. Within and across platforms, it is possible to use these systems with a single input, e.g. movie, or multiple inputs, e.g. news, books, and search queries. Popular recommender systems are available for specializing themes such as restaurants and online dating. A recommender system has been used to explore research papers, experts, collaborators, and financial services. The same recommendation systems that employ one type of filtering to forecast ratings and customer satisfaction make it possible for consumers to purchase things based on their interests or requirements. Information about a person's life might provide insight into how they would behave in certain situations. A recommender system is a valuable replacement for search because it allows consumers to discover things they might not have otherwise found. Search engines that do not use standard indexing are commonly used in this way.

   a. Collaborative filtering
      The idea behind collaborative filtering is that individuals who have agreed in the past will again agree in the future, and will reap the benefits of similar products. This algorithm produces suggestions based solely on ratings for various individuals or things. Users and items are produced suggestions using this community by looking for peer users and items who have a rating history similar to the current user or item.

   b. Content-Based Filtering
      Content-based filtering algorithms are created using the item description and a user profile. When a user cannot be identified, these strategies work best with item information (name, location, description, etc.). Content-based recommenders approach recommendation as a user-specific classification issue, in which attributes of an item are used to create a candidate classifier for the user's likes and dislikes.

   c. Hybrid Recommender Systems
      There are many situations in which hybrid approaches can be applied. It can be used by itself or in conjunction with collaborative filtering to determine content-based and collaborative filtering predictions. Using a collaborative approach, it can be utilized to improve content-based functionality. All of these elements may be combined into a single format using it.

2. Social Network Analysis
   (SNA) uses networks and graph theory to examine social structures. As a result, the concept of a networked system divides actors, persons, and objects into nodes (individuals), edges (relationships), and links (interactions or relations between them). SNA is capable of processing vast amounts of relational data and describing the overall topology of a relational network. SNA context and choose the parameters to use to identify the "center" based on the network's characteristics. The communication structure and location of persons may be precisely characterized by examining nodes, clusters, and relations.

Recommendation System Using SNA

Now, let's understand how hybrid systems can be used with the social network analysis to generate movie recommendations to the new user entering the network by eliminating the cold start problem.

1. Initially in the generic recommendation system, we face the cold start problem that is for any new user that enters the system, the system does not have enough information about the user to make a decision.
2. Via this project we intend to put forth a theory that can eliminate the cold start problem, that is whenever a new user enters the system, the system already has a list of the most popular movies in each community of people based on their age, gender, occupation,and zip code that can be recommended to the user.

Similarly, this can be done using content-based filtering wherein we collect the user information of all the users in the network

1. Communities are generated depending on the similar movie taste among the users. That is there can be multiple users who have liked the same movie, hence a community can be formed among these users.
2. We used the Louvain community detection method to generate the communities among the users who similar taste in the movies. We used the louvain community detecetion because it optimizes a modularity score for each community, where modularity is a measure of how well nodes are assigned to communities. This involves determining how much more tightly connected nodes within a community are than they would be in a random network.
3. Accordingly, in this case, the nodes in the network will be the users and the edges are the number of similar movies watched by the users, that is the count of the number of movies. Hence, the users are connected to one another with similar tastes in movies, that is who like the same movies.
4. Furthermore, whenever a new user is created, the user attributes are considered and based on these user data, we strive to discover the most comparable person already in the community. When a new user is matched with someone from one of the network's

communities, the new user is joined to that community, and the new user gets shown movies from that community.

5. Hence this new user is recommended movies depending on the movies liked by the users in the community of these users in the network.

Working of the Model

1. For the scope of this project, we are developing a model for movie recommendations that will use the hybris system and social network analysis for generating movie recommendations.
2. The data considered for this project has a set of 50 users, we have used the Movie Lens dataset for this project. To describe more about this dataset this data has been cleaned up users who had less than 20 ratings or did not have a complete demographic information was removed from this data set. It has 100,000 ratings (1-5) from 943 users on 1682 movies. Each user has rated at least 20 movies. Simple demographic info for the users (age, gender, occupation, zip)
3. We intend to build a network that will provide movie recommendations to individuals depending on their age, occupation, gender, and zip code. For this we use the Louvain community detection method. The Louvain Community Detection technique is a greedy optimization tool for increasing the "modularity" of a network split. The optimization procedure is divided into two stages. To begin, the method hunts for "small" groups by maximizing modularity on a local level. Second, it gathers nodes from the same community and builds a new network that includes the communities as nodes. These processes are continued until a community hierarchy is built and a maximum level of modularity is obtained.
4. We can couple the nodes together based on their commonalities after we have a count of the users who has viewed the same movies. This is accomplished by determining if the users are of the same gender, occupation, age difference, and zip code, and assigning values to them. Once we have the count of the users, we sum the attribute values and take into consideration the user which represents the highest sum of the attributes. The higher the value, the higher the similarity between the users.

Cold Start Problem

The Cold Start Problem occurs when a new user joins the system and the system lacks sufficient knowledge to make a decision for them. When things are added to the catalog, the item cold-start problem occurs when they have no or very few interactions. This is particularly problematic for collaborative filtering algorithms, which generate suggestions based on the item's interactions. A pure collaborative algorithm cannot propose an item if there are no interactions available. Although a collaborative algorithm will be able to recommend it if just a few interactions are available, the quality of those recommendations will be low. This raises a new dilemma, one that has nothing to do with

new things and everything to do with unpopular items. In some circumstances, a small number of things may receive a large number of interactions, whereas the majority of the items receive only a few fraction of them.

Attempt to alleviate the problem of Cold Start Problem

We've designated 20 nodes as testing data in our model. As a result, whenever a new user is created, the user attributes that is age, gender, occupation and zip code are considered and based on these user data, we strive to discover the most comparable person already in the community. Hence, when a new user is matched with someone from one of the network's communities, the new user is joined to that community, and the new user gets shown movies from that community. This is how the problem of a cold start may be solved.

4. Result and Analysis

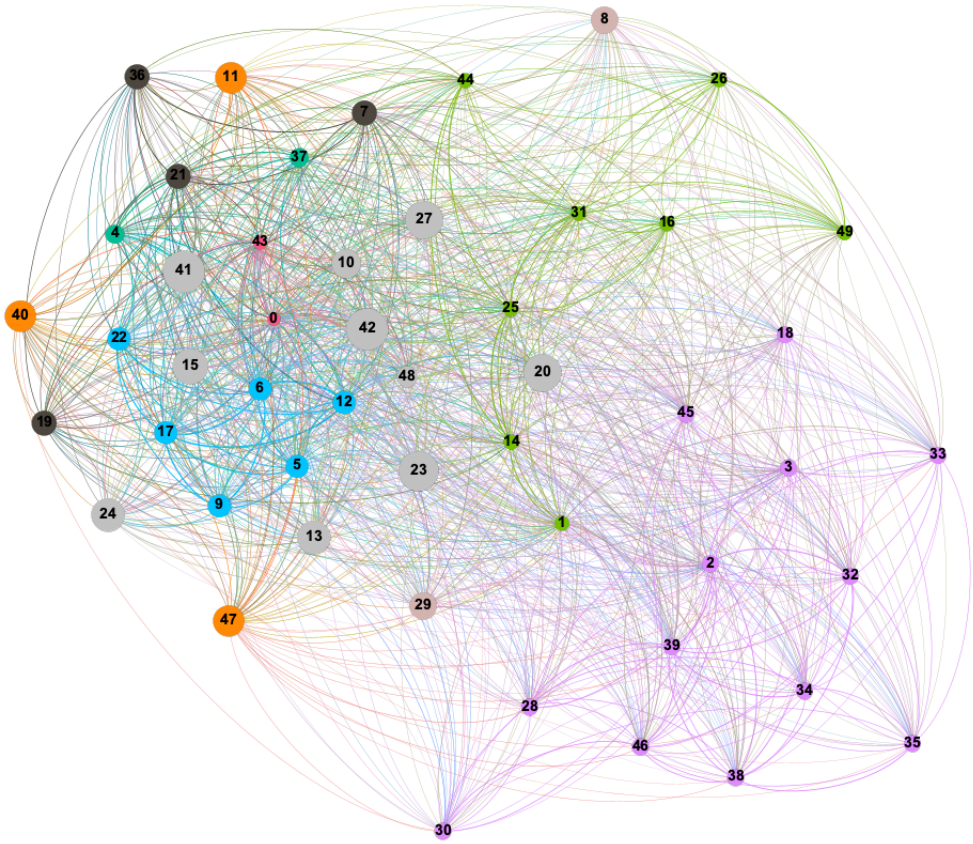

Fig - Visualisation of the User Graph

1. For this project, we employed the hybrid approach, which combines collaborative and content-based filtering. Content-based because the technique takes into account the user data, such as age, gender, occupation, and zip code, and Collaborative Filtering because it does not take each user's preferences into account.
2. Understanding trends through social media is a more easier way to put recommendation systems in place, and information dissemination is more efficient in these networks.
3. By using social network analysis we can understand what each age group, or occupation group in the society prefer to watch and hence due to this we can categorize the new user that enters the network.

5. Future Scope

The creation of this model can be used on real-time data, as well as its deployment on OTT platforms like Netflix, Amazon Prime Video, and others. We may also include the genre of each film and recommend films depending on the user's preferences for each category.

6. Acknowledgements

We take this opportunity to express our profound gratitude and deep regards to our guide Prof. Yeong Yeol Ahn for her exemplary guidance, monitoring. No project is ever complete without the guidelines of the experts who have already established a mark on this path before and have become masters of it. So, we would like to take this opportunity to thank all those who have helped us in the whole journey.

7. References

[1] Johann Stan, Fabrice Muhlenbach, Christine Largeron. Recommender Systems using Social Network Analysis: Challenges and Future Trends. Alhajj, Reda, and Rokne, Jon. Encyclopedia of Social Network Analysis and Mining, Springer, pp.1-22, 2014, 978-1-4614-6171-5.
Link for reference: (PDF) Recommender Systems Using Social Network Analysis: Challenges and Future Trends.

[2] Dolgikh, Dmitry & Jelinek, Ivan. (2015). Graph-based music recommendation approach using social network analysis and community detection method. 221-227. 10.1145/2812428.2812453.
Link for reference: Graph-based music recommendation approach using social network analysis and community detection method | Request PDF.

[3] Xinchang, Khamphaphone, Vilakone, Phonexay, Park, Doo-Soon “Movie Recommendation Algorithm Using Social Network Analysis to Alleviate Cold-Start Problem”
Link for reference: Movie Recommendation Algorithm Using Social Network Analysis to Alleviate Cold-Start Problem -Journal of Information Processing Systems.

[4] Chang Li, Jiaxiang Chen, Yangyue Wan, Building Music Recommender Systems with Graph-based Method
Link for reference: https://cseweb.ucsd.edu/classes/wi15/cse255-a/reports/fa15/008.pdf

[5] B. Fields, K. Jacobson, C. Rhodes, M. d'Inverno, M. Sandler and M. Casey, "Analysis and Exploitation of Musician Social Networks for Recommendation and Discovery,"
Link for reference: Analysis and Exploitation of Musician Social Networks for Recommendation and Discovery | IEEE Journals & Magazine | IEEE Xplore.