Movie Recommendation Systems

CSE575 - Project Proposal

Team Members

1) Dharmit Prajapati (dprajap3@asu.edu)
2) Aakash Mathai (amathai1@asu.edu)
3) Abhinav Venepally (avenepal@asu.edu)
4) Sunayana Gupta (sgupt279@asu.edu)
5) Jayashree Natarajan (jnatara@asu.edu)
6) Vishnuraj Kunnathveetil (vkunnath@asu.edu)

Description of the proposal

Due to an increase in e-commerce and online media consumption over the last two decades, recommendation systems have grown in popularity. We've witnessed an explosion of different media firms launching their own streaming services in recent years, necessitating the development of systems that can recommend suitable content to viewers in real-time from millions of selections in a company's movie repertoire. Because the issue of streaming services has grown in popularity at such a quick rate, we chose to limit our endeavor to movies alone. People have attempted a variety of techniques to address this specific issue because it is such a well-researched topic.

We will be using the below algorithms:

1) Collaborative Filtering
In this, the users are asked for their preferences and this forms the parameters for each user. We can use these parameters as well as use the ratings given by the user to provide movie recommendations.

2) Unsupervised K-Clusters

We start by extracting features from the movie dataset using techniques such as PCA feature extraction to reduce the number of genres and thereby reduce the dimensionality for more efficient processing. Once the features are extracted, these are then fed into a K-clustering algorithm which groups the movies into a specific set of clusters. After each movie is assigned to its cluster classification, for each user, the highest rated movies of each user are compiled if the movie exceeds a certain threshold and then determine which K-clusters the user prefers the most.

3) Matrix Factorization

Matrix factorization works by characterizing items and users by vectors of factors that are inferred from the item rating patterns, with high correspondence between item and user resulting in a recommendation

4) Content-based recommender systems

Based on the genre of the movie and how the user has rated other movies we can predict how the user would rate other movies (movies not yet watched by the user) and suggest the top 20-30 movies based on this.

We will be using the K-Fold cross-validation and the model with the best result should be our ideal model. We will be comparing the models based on the accuracy(how successfully the models identify fresh recommendations for users) and efficiency(how long each model takes to complete)

Timeline

For this project, we have reached a decision to experiment with several machine learning practices, from which we will be comparing results to choose the more suitable and modified algorithm. For the project proposal, we have researched a few of these approaches, and plan on immediately implementing them by the milestone(Mar 14th- Mar 18th) and post that we will focus on improving the accuracy and performance of these 4 models and compare them on various parameters. We also plan on dividing the work in order to reach faster results while maintaining regular team meetings in order to share knowledge and maintain a similar pace among all divisions.

REFERENCES

Schafer, J. B., Frankowski, D., Herlocker, J., & Sen, S. (n.d.). Collaborative filtering recommender systems. *The Adaptive Web*, 291–324. https://doi.org/10.1007/978-3-540-72079-9_9

Yehuda Koren, Robert M. Bell, Chris Volinsky: Matrix Factorization Techniques for Recommender Systems. IEEE Computer 42(8): 30-37 (2009)

Yassine, A., Mohamed, L., & Al Achhab, M. (2021). Intelligent recommender system based on unsupervised machine learning and demographic attributes. *Simulation Modelling Practice and Theory*, *107*, 102198. https://doi.org/10.1016/j.simpat.2020.102198