Group Project Proposal

Huiying Chen*
Anju Kunnumpurathu James*
George Torres*
Fabio Tardivo*
hchen@nmsu.edu
anju@nmsu.edu
gtorresz@nmsu.edu
ftardivo@nmsu.edu
New Mexico State University
Las Cruces, NM, USA

CCS CONCEPTS

• **Computer systems organization** → **Machine Learning**; *information filtering systems*; • **Recommendation System**;

KEYWORDS

machine learning, neural networks, movie recommendation

ACM Reference Format:
Huiying Chen, Anju Kunnumpurathu James, George Torres, and Fabio Tardivo. 2021. Group Project Proposal. In *Proceedings of (CS487/519 Applied Machine Learning)*. ACM, New York, NY, USA, 1 page.

1 MOTIVATION

The rapid growth of data collection has led to a new era of information. Data is being used to create more efficient systems and this is where Recommendation Systems come into play. Recommendation Systems are a type of information filtering systems as they improve the quality of search results and provides items that are more relevant to the search item or are related to the search history of the user.

They are used to predict the rating or preference that a user would give to an item. Almost every major tech company has applied them in some form or the other: Amazon uses it to suggest products to customers, YouTube uses it to decide which video to play next on autoplay, and Facebook uses it to recommend pages to like and people to follow. Moreover, companies like Netflix and Spotify depend highly on the effectiveness of their recommendation engines for their business and success.

*All authors contributed equally to this research.

CS487/519 Applied Machine Learning, Group Project Proposal,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

2 PROBLEM DEFINITION

We'll propose a baseline Movie Recommendation System using TMDB 5000 Movie Dataset. We are going to define a function that takes in a movie title as an input and outputs a list of the 10 most similar movies.

3 RELATED WORK

- **Demographic Filtering**: They offer generalized recommendations to every user, based on movie popularity and/or genre. The System recommends the same movies to users with similar demographic features. Since each user is different , this approach is considered to be too simple. The basic idea behind this system is that movies that are more popular and critically acclaimed will have a higher probability of being liked by the average audience.
- **Based Filtering**: They suggest similar items based on a particular item. This system uses item metadata, such as genre, director, description, actors, etc. for movies, to make these recommendations. The general idea behind these recommender systems is that if a person liked a particular item, he or she will also like an item that is similar to it.
- **Collaborative Filtering**: This system matches persons with similar interests and provides recommendations based on this matching. Collaborative filters do not require item metadata like its content-based counterparts.

4 DATASET

TMDB 5000 Movie Dataset contains metadata for all 45,000 movies listed in the Full MovieLens Dataset. The dataset consists of movies released on or before July 2017. Data points include cast, crew, plot keywords, budget, revenue, posters, release dates, languages, production companies, countries, TMDB vote counts and vote averages. This dataset also has files containing 26 million ratings from 270,000 users for all 45,000 movies. Ratings are on a scale of 1-5 and have been obtained from the official GroupLens website.

5 EXPERIMENTS

6 REFERENCE