CSCE 5222: Feature Engineering

Project Proposal

Project Title : Movie Recommendation System

Team Members :

1. Reshmi Chowdary Divi - 11520230
2. Phani Sagar Kothapalli - 11555373
3. Harshith Reddy Nagireddy - 11517457

Idea description :

Recommender frameworks being a piece of data sifting framework are utilized to gauge the predisposition or appraisals the client will in general give for a thing. Among various types of suggestion draws near, shared sifting procedure has a high ubiquity due to their adequacy. These conventional collective sifting frameworks can even work viably and can deliver standard suggestions, in any event, for wide running issues.

For thing dependent on their neighbor's inclinations Collaborative separating strategies makes preferred proposals over others. Though different strategies like substance-based experiences helpless exactness, versatility, information sparsity and enormous blunder forecast. To discover these prospects, we have utilized thing based synergistic sifting approach. In this Item based collective separating method we initially inspect the User thing rating framework and we distinguish the connections among different things, and afterward we utilize these connections to figure the suggestions for the client.

In this project we are going to use Movie Recommendation System using collaborative filtering. The dataset used in this is the Movie-Lens dataset.

Goals and Objectives:

Goals :

KNN's exhibition will experience the ill effects of revile of dimensionality in the event that it utilizes "euclidean distance" in its goal work. Euclidean distance is pointless in high measurements since all vectors are practically equidistant to the inquiry question vector (focus on film's highlights). All things considered, we will utilize cosine closeness for closest neighbor search

Objective:

Recommendation System is a filtration program whose superb objective is to foresee the “rating “ or “inclination” of a client towards an area explicit thing or thing. For our situation, this space explicit thing is a film, thusly the principle focal point of our proposal framework is to channel and anticipate

just those motion pictures which a client would favor given some information about the client

oneself.

Motivation:

Collaborative filtering is a type of recommendation engine that uses both user and item data. More specifically, ratings from individual users on individual items. This way, items are recommended based on the ratings from other users, thus, collaborative. This data can be represented in a utility matrix, with one axis being users and one axis being items. The goal of collaborative filtering recommendation engines is to fill in the gaps in a utility matrix since not every user has rated every item, and then output the top-rated, previously-unrated items as recommendations.

Significance:

Another significant advance in information preprocessing is to normalize the dataset. This cycle makes the mean of the multitude of info highlights equivalent to zero and furthermore changes their difference over to 1. This guarantees that there is no predisposition while preparing the model because of the various sizes of all info highlights. In the event that this isn't done the neural organization may get befuddled and give a higher load to those highlights which have a higher normal incentive than others.

We actualize this progression by bringing in the StandardScaler strategy from the sklearn.preprocessing library. We launch the variable sc with the StandardScaler() work. After which we utilize the fit_transform work for actualizing these progressions on the X_train and X_test datasets. The y_train and y_test sets contain double qualities, subsequently they need not be normalized. Presently that the datasets are prepared, we may continue with building the Artificial Neural Network utilizing the Keras library.

Literature Survey:

In [1] Author used recommendation system using collaborative filtering. It is implemented using Apache Mahout and takes the ratings given to movies to provide movie suggestions. Our system considers the user ratings to recommend movies.

In [2] Author the recommendation system recommends different movies to users. Since this system is based on a collaborative approach, it will give progressively explicit outcomes contrasted with different systems that are based on the content-based approach.

In [3] Author includes a summary review of literature studies related to the movie recommendation system based on collaborative filtering.

Objectives:

- Subclass of Information filtering system that seek to predict the 'rating' or 'preference' that a user would give to them.
- Helps deciding in what to wear, what to buy, what stocks to purchase etc.
- Applied in variety of applications like movies, books, research articles.
- Recommendation systems has mainly two elements Item and User.

Features:

Latent features just refer to some abstraction of all the items' or users' features. As long as you have the **same number of latent features for items and users**, you can multiply the matrices out to produce a matrix with the same dimensions as your utility matrix. The number of latent features is a hyperparameter you can tune in your model. Based on matrix multiplication, we can also see that the rating value of U1 for I3 is affected by the I3 row of the item matrix and the U1 column of the user matrix.

Because we can't decompose matrices with missing values, we have to take another approach. This is where machine learning comes in. As previously mentioned, we're now going to try to **recreate the utility matrix with our item matrix and our user matrix**. This is done using a method of gradient descent: **Alternating Least Squares**

Expected outcome:

userId	1	2	3	4	5	6	7	8	9	10	...	601	602	603	604	605	606
movieId																	
1	4.0	0.0	0.0	0.0	4.0	0.0	4.5	0.0	0.0	0.0	...	4.0	0.0	4.0	3.0	4.0	2.5
2	0.0	0.0	0.0	0.0	0.0	4.0	0.0	4.0	0.0	0.0	...	0.0	4.0	0.0	5.0	3.5	0.0
3	4.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	3.0	0.0	0.0

References :

1. Ananya Agarwal, S. Srinivasan "Movie Recommendation System" (2020)
2. F. Furtado, A. singh, "Movie Recommendation System Using Machine Learning" (2019)
3. Nirav Raval, Vijayshri Khedar, "Collaborative Filtering System Based Movie Recommendation System" (2018)