**Previous Work and Analysis**

|  |
| --- |
| **Link to previous work:**  http://datamining-r.blogspot.in/2014/09/movie-recommender-system-in-r.html |
| **Finding in previous work:**  In this work, they had used the recommender package available in R. They had built a recommender system based on collaborative filtering.They had build the User Based Collaborative Filtering (UBCF) Model. One of the most popular collaborative filtering algorithms is User-based Collaborative Filtering, also known as Resnick’s algorithm. It operates by first calculating similarity values between users in a system, then finding the k most similar. These k users constitute a so called neighbourhood. The recommender system then uses the neighbourhood’s ratings to compute and return a list of n recommendations.In this work,It initially preprocesses the data using the preProcess function available in R.They create the rating matrix and then create a model based on the matrix. The create evaluation model and finally using the UBCF method and recommendation function available in R, they provide the list of top five movies. |
| **Movie-Recommender System(Need/Scope)**  Recommender systems are used to provide personalized recommendations according to user profile and previous behavior.Recommender systems are widely used in the Internet Industry.Services like Amazon,Netflix, and YouTube are typical examples of recommender system users.Recommender systems cannot only help the users find their favorite products, but also bring potential profit to online service providers.  In our project we have:   * Analysed the MovieLens Data set * Content Based Filtering Method * Collaborative Filtering Method * Evaluation of the created Model * An interactive User Interface to display the list of recommended movies bsed on top three choices of the user |
| **Differences in methods/techniques applied between previous work and Movie-Recommender System:**  In the previous work only Collaborative filtering technique was ued but in our work we have build a recommender system using both content based and collaborative based filtering taking into account three best choices of the user. We had also analysed the data set and evaluated our model and displayed the result in an interactive user interface using the shiny package available in R  In Content based filtering we had  The code does the following,   * Data preprocessing, pipe separated genres had to split * Created matrix with columns representing every unique genre * Convert into Data Frame for processing user profile matrix in the next step * ( Temporary Conversion for faster processing , I first convert the ratings into a binary format to keep things simple. ratings of 4 and 5 are mapped to 1, representing likes, and ratings of 3 and below are mapped to -1, representing dislikes. ) * Dcast function is used to build user profile matrix ( change format to wide from long ) * To generate simple user profile matrix, calculated the dot product of the movie genre matrix and the binaryratings matrix. **This user profiles shows the aggregated inclination of each user towards movie genres.** * Generate two separate matrix, where in the rows,i.e, movies which has no ratings are removed and another vice- versa, not necessary just for easier processing. (Also since there was observed differences in numbers between both, coz of null sets) * In order to generate user profile similarity, we are used **Jaccard distance** method of finding from vegan library * **Recommend movies - Simple content based method**   **In Collaborative filtering:**   * We first create the rating matrix * Notes :   + Cosine Similarity Method in UBCF is used. Nearest Neighbours considered 30.   + The predicted movie ratings of the user will be derived from the 5 nearest neighbors in its neighborhood.   + When the predicted item ratings are obtained, the top 10 most highly predicted ratings will be returned as the recommendations. * First converted rating matrix into a recommenderlab sparse matrix. * Normalize the data and obtain recommendations   **In the previous work**   * It initially preprocesses the data using the preProcess function available in R. * They create the rating matrix and then create a model based on the matrix. * They create evaluation model * Finally using the UBCF method and recommendation function available in R, they provide the list of top five movies. |
| **Strength**  User-based Collaborative Filtering gives us recommendations that can be complements to the item the user was interacting with. This might be a stronger recommendation along with what a content-based recommender can provide as users might not be looking for direct substitutes to a movie they had just viewed or previously watched. |
| **Weakness**  **If there were millions of users, this computation would be very time consuming.**  User-based Collaborative Filtering is a type of Memory-based Collaborative Filtering that uses all user data in the database to create recommendations. Comparing the pairwise correlation of every user in your dataset is not scalable. Possible ways to get around this would be to implement some form of dimensionality reduction or to use a model-based algorithm instead. |