Movie Recommender System

2024-01-18

PROJECT OVERVIEW

In this project, I developed a movie recommender system similar to those used by popular streaming platforms like Netflix, Hotstar, and Amazon Prime. I utilized a dataset from MovieLens, which consists of approximately 100,000 ratings given by 943 users on 1,664 movies.

First lets create top recommendations, demo it another top users. First we will create a userbased then an itembased recommendation systems and understand the patterns.

```
#This is user-based
library(recommenderlab)
```

```
## Warning: package 'recommenderlab' was built under R version 4.3.2
```

```
## Loading required package: Matrix
```

```
## Warning: package 'Matrix' was built under R version 4.3.3
```

```
## Loading required package: arules
```

```
## Warning: package 'arules' was built under R version 4.3.2
```

```
##
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
##
##     abbreviate, write
```

```
## Loading required package: proxy
```

```
##
## Attaching package: 'proxy'
```

```
## The following object is masked from 'package:Matrix':
##
##     as.matrix
```

```
## The following objects are masked from 'package:stats':
##
##     as.dist, dist
```

```
## The following object is masked from 'package:base':
##
##     as.matrix
```

```
## Registered S3 methods overwritten by 'registry':
##   method               from
##   print.registry_field proxy
##   print.registry_entry proxy
```

```
data(MovieLense)

R <- as(MovieLense, "realRatingMatrix")

recommender_userbased <- Recommender(R, method = "UBCF")

#Creating top 10 recommendations demo it on 5 users
recom <- predict(recommender_userbased, R[1:5], n=10)
recom
```

```
## Recommendations as 'topNList' with n = 10 for 5 users.
```

```
as(recom, "list")
```

```
## $`0`
##  [1] "Boot, Das (1981)"
##  [2] "Matilda (1996)"
##  [3] "Winter Guest, The (1997)"
##  [4] "She's the One (1996)"
##  [5] "Manchurian Candidate, The (1962)"
##  [6] "City of Lost Children, The (1995)"
##  [7] "Double vie de Veronique, La (Double Life of Veronique, The) (1991)"
##  [8] "187 (1997)"
##  [9] "Leaving Las Vegas (1995)"
## [10] "Wag the Dog (1997)"
## 
## $`1`
##  [1] "Con Air (1997)"
##  [2] "Paradise Lost: The Child Murders at Robin Hood Hills (1996)"
##  [3] "Forbidden Planet (1956)"
##  [4] "Great Dictator, The (1940)"
##  [5] "Rob Roy (1995)"
##  [6] "M (1931)"
##  [7] "Game, The (1997)"
##  [8] "Local Hero (1983)"
##  [9] "C'est arrive pres de chez vous (1992)"
## [10] "In the Name of the Father (1993)"
## 
## $`2`
##  [1] "Heavy Metal (1981)"
##  [2] "Mystery Science Theater 3000: The Movie (1996)"
##  [3] "Fear of a Black Hat (1993)"
##  [4] "Serial Mom (1994)"
##  [5] "Brady Bunch Movie, The (1995)"
##  [6] "Forbidden Planet (1956)"
##  [7] "39 Steps, The (1935)"
##  [8] "Crying Game, The (1992)"
##  [9] "Paths of Glory (1957)"
## [10] "Adventures of Priscilla, Queen of the Desert, The (1994)"
## 
## $`3`
##  [1] "It Happened One Night (1934)"      "Jungle Book, The (1994)"
##  [3] "Deer Hunter, The (1978)"           "Man Who Would Be King, The (1975)"
##  [5] "39 Steps, The (1935)"              "Innocents, The (1961)"
##  [7] "Old Man and the Sea, The (1958)"   "Delta of Venus (1994)"
##  [9] "Big Lebowski, The (1998)"          "Ed Wood (1994)"
## 
## $`4`
##  [1] "Paradise Lost: The Child Murders at Robin Hood Hills (1996)"
##  [2] "Leaving Las Vegas (1995)"
##  [3] "Primal Fear (1996)"
##  [4] "Anna Karenina (1997)"
##  [5] "Wallace & Gromit: The Best of Aardman Animation (1996)"
##  [6] "Wizard of Oz, The (1939)"
##  [7] "Clockwork Orange, A (1971)"
##  [8] "Field of Dreams (1989)"
##  [9] "Sling Blade (1996)"
## [10] "Night of the Living Dead (1968)"
```

```
#Otherwise
Recom10 <- bestN(recom, n = 10)
Recom10
```

```
## Recommendations as 'topNList' with n = 10 for 5 users.
```

```
as(Recom10, "list")
```

```
## $`0`
##  [1] "Boot, Das (1981)"
##  [2] "Matilda (1996)"
##  [3] "Winter Guest, The (1997)"
##  [4] "She's the One (1996)"
##  [5] "Manchurian Candidate, The (1962)"
##  [6] "City of Lost Children, The (1995)"
##  [7] "Double vie de Veronique, La (Double Life of Veronique, The) (1991)"
##  [8] "187 (1997)"
##  [9] "Leaving Las Vegas (1995)"
## [10] "Wag the Dog (1997)"
##
## $`1`
##  [1] "Con Air (1997)"
##  [2] "Paradise Lost: The Child Murders at Robin Hood Hills (1996)"
##  [3] "Forbidden Planet (1956)"
##  [4] "Great Dictator, The (1940)"
##  [5] "Rob Roy (1995)"
##  [6] "M (1931)"
##  [7] "Game, The (1997)"
##  [8] "Local Hero (1983)"
##  [9] "C'est arrive pres de chez vous (1992)"
## [10] "In the Name of the Father (1993)"
##
## $`2`
##  [1] "Heavy Metal (1981)"
##  [2] "Mystery Science Theater 3000: The Movie (1996)"
##  [3] "Fear of a Black Hat (1993)"
##  [4] "Serial Mom (1994)"
##  [5] "Brady Bunch Movie, The (1995)"
##  [6] "Forbidden Planet (1956)"
##  [7] "39 Steps, The (1935)"
##  [8] "Crying Game, The (1992)"
##  [9] "Paths of Glory (1957)"
## [10] "Adventures of Priscilla, Queen of the Desert, The (1994)"
##
## $`3`
##  [1] "It Happened One Night (1934)"      "Jungle Book, The (1994)"
##  [3] "Deer Hunter, The (1978)"           "Man Who Would Be King, The (1975)"
##  [5] "39 Steps, The (1935)"              "Innocents, The (1961)"
##  [7] "Old Man and the Sea, The (1958)"   "Delta of Venus (1994)"
##  [9] "Big Lebowski, The (1998)"          "Ed Wood (1994)"
##
## $`4`
##  [1] "Paradise Lost: The Child Murders at Robin Hood Hills (1996)"
##  [2] "Leaving Las Vegas (1995)"
##  [3] "Primal Fear (1996)"
##  [4] "Anna Karenina (1997)"
##  [5] "Wallace & Gromit: The Best of Aardman Animation (1996)"
##  [6] "Wizard of Oz, The (1939)"
##  [7] "Clockwork Orange, A (1971)"
##  [8] "Field of Dreams (1989)"
##  [9] "Sling Blade (1996)"
## [10] "Night of the Living Dead (1968)"
```

```
#this is item based
recommender_itembased <- Recommender(R, method = "IBCF")

#Creating top 10 recommendations demo it on 5 users
recom <- predict(recommender_itembased, R[1:5], n=10)
recom
```

```
## Recommendations as 'topNList' with n = 10 for 5 users.
```

```
as(recom, "list")
```

```
## $`0`
## [1] "Entertaining Angels: The Dorothy Day Story (1996)"
## [2] "Big Bang Theory, The (1994)"
## [3] "They Made Me a Criminal (1939)"
## [4] "Cyclo (1995)"
## [5] "King of New York (1990)"
## [6] "Very Natural Thing, A (1974)"
## [7] "Walk in the Sun, A (1945)"
## [8] "Hush (1998)"
##
## $`1`
##  [1] "They Made Me a Criminal (1939)"
##  [2] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [3] "Big Bang Theory, The (1994)"
##  [4] "Other Voices, Other Rooms (1997)"
##  [5] "Further Gesture, A (1996)"
##  [6] "Cyclo (1995)"
##  [7] "Very Natural Thing, A (1974)"
##  [8] "Walk in the Sun, A (1945)"
##  [9] "Visitors, The (Visiteurs, Les) (1993)"
## [10] "Hush (1998)"
##
## $`2`
##  [1] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [2] "Legal Deceit (1997)"
##  [3] "Other Voices, Other Rooms (1997)"
##  [4] "Hush (1998)"
##  [5] "Big Bang Theory, The (1994)"
##  [6] "Visitors, The (Visiteurs, Les) (1993)"
##  [7] "Mirage (1995)"
##  [8] "Further Gesture, A (1996)"
##  [9] "Cyclo (1995)"
## [10] "Tokyo Fist (1995)"
##
## $`3`
##  [1] "Men of Means (1998)"
##  [2] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [3] "Mirage (1995)"
##  [4] "Further Gesture, A (1996)"
##  [5] "Big Bang Theory, The (1994)"
##  [6] "Other Voices, Other Rooms (1997)"
##  [7] "Star Kid (1997)"
##  [8] "Very Natural Thing, A (1974)"
##  [9] "Walk in the Sun, A (1945)"
## [10] "Tokyo Fist (1995)"
##
## $`4`
##  [1] "They Made Me a Criminal (1939)"
##  [2] "Coldblooded (1995)"
##  [3] "Visitors, The (Visiteurs, Les) (1993)"
##  [4] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [5] "King of New York (1990)"
##  [6] "Shopping (1994)"
##  [7] "Nemesis 2: Nebula (1995)"
##  [8] "Very Natural Thing, A (1974)"
##  [9] "Walk in the Sun, A (1945)"
## [10] "Cyclo (1995)"
```

```
Recom10 <- bestN(recom, n = 10)
Recom10
```

```
## Recommendations as 'topNList' with n = 10 for 5 users.
```

```
as(Recom10, "list")
```

```
## $`0`
## [1] "Entertaining Angels: The Dorothy Day Story (1996)"
## [2] "Big Bang Theory, The (1994)"
## [3] "They Made Me a Criminal (1939)"
## [4] "Cyclo (1995)"
## [5] "King of New York (1990)"
## [6] "Very Natural Thing, A (1974)"
## [7] "Walk in the Sun, A (1945)"
## [8] "Hush (1998)"
##
## $`1`
##  [1] "They Made Me a Criminal (1939)"
##  [2] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [3] "Big Bang Theory, The (1994)"
##  [4] "Other Voices, Other Rooms (1997)"
##  [5] "Further Gesture, A (1996)"
##  [6] "Cyclo (1995)"
##  [7] "Very Natural Thing, A (1974)"
##  [8] "Walk in the Sun, A (1945)"
##  [9] "Visitors, The (Visiteurs, Les) (1993)"
## [10] "Hush (1998)"
##
## $`2`
##  [1] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [2] "Legal Deceit (1997)"
##  [3] "Other Voices, Other Rooms (1997)"
##  [4] "Hush (1998)"
##  [5] "Big Bang Theory, The (1994)"
##  [6] "Visitors, The (Visiteurs, Les) (1993)"
##  [7] "Mirage (1995)"
##  [8] "Further Gesture, A (1996)"
##  [9] "Cyclo (1995)"
## [10] "Tokyo Fist (1995)"
##
## $`3`
##  [1] "Men of Means (1998)"
##  [2] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [3] "Mirage (1995)"
##  [4] "Further Gesture, A (1996)"
##  [5] "Big Bang Theory, The (1994)"
##  [6] "Other Voices, Other Rooms (1997)"
##  [7] "Star Kid (1997)"
##  [8] "Very Natural Thing, A (1974)"
##  [9] "Walk in the Sun, A (1945)"
## [10] "Tokyo Fist (1995)"
##
## $`4`
##  [1] "They Made Me a Criminal (1939)"
##  [2] "Coldblooded (1995)"
##  [3] "Visitors, The (Visiteurs, Les) (1993)"
##  [4] "Entertaining Angels: The Dorothy Day Story (1996)"
##  [5] "King of New York (1990)"
##  [6] "Shopping (1994)"
##  [7] "Nemesis 2: Nebula (1995)"
##  [8] "Very Natural Thing, A (1974)"
##  [9] "Walk in the Sun, A (1945)"
## [10] "Cyclo (1995)"
```

Now to compare A and B, we see the following: For User 0 we see that User-Based system recommends movies like Boot, Das and Matilda. While Item-Based system duggests movies such as The Dorothy Day Story and Big Bang Theory.

For User 1, the User-Based system Provides recommendations like Con Air and Paradise Lost: The Child Murders at Robin Hood Hills. While Item-Based system suggests films like “They Made Me a Criminal and The Dorothy Day Story.

For User 2, the User-Based system recommends movies including Heavy Metal and Mystery Science Theater 3000: The Movie (1996). While Item-Based suggests films such as The Dorothy Day Story (1996) and Legal Deceit (1997).

Similarly for User 3 the User-Based system provides recommendations like It Happened One Night and Jungle Book. Wjile Item-Based suggests movies such as Men of Means and The Dorothy Day Story.

Now, for User 4, the User-Based system recommends movies including Paradise Lost: The Child Murders at Robin Hood Hills and Leaving Las Vegas. While Item-Based system suggests films like They Made Me a Criminal and Coldblooded.

Therefore we can say that, User based recommendations use collaborative filtering by identifying users with similar preferences. They personalizes recommendations by considering the historical interactions of a user’s peers.Thereby, provides a mix of movie genres, as seen in suggestions like Con Air and Heavy Metal. The recommendations are influenced by the collective behavior of users having similar taste, and this enhance variety.

Where as, Item-Based Recommendations utilizes collaborative filtering by emphasizing item similarity and recommends movies based on the intrinsic characteristics of items the user has previously liked. They highlight similarities based on the content, and offers suggestions like Legal Deceit and Coldblooded. This mainly focuses on the user’s individual history, tailoring suggestions to specific movie attributes. This also captures a more focused set of preferences, considering the user’s past interactions with specific movie traits.

What are our observation

Relevance: The recommended movies include titles such as Boot, Das and Matilda. The system appears to provide relevant recommendations, as these movies align with the user’s potential preferences based on their historical interactions as it corresponds to a known interest in historical and war movies and movies with depth unline normal genres like rom-com.

Novelty: Here, the recommendations include well-known titles like Boot, Das but also potentially introduce lesser-known titles. Therefore, we can say that the system introduces a level of novelty by suggesting movies that the user may not have already seen. Also, suggesting movies like Winter Guest, which introduces novelty, as this film may not be as widely known or seen compared to mainstream releases. However, the degree of novelty may depend on the user’s familiarity with the suggested titles.

Serendipity: The recommended movies, such as Winter Guest and City of Lost Children, may offer unexpected and diverse choices. Here, the system demonstrates serendipity by suggesting movies that may be less predictable based on the user’s past interactions, potentially leading to pleasant surprises.

Diversity: We see that the system offered a mix of genres like Matilda (family), Winter Guest, (drama), and City of Lost Children (fantasy) shows diversity in catering to different movie preferences within the user’s potential interests.The recommended list includes movies of different genres, such as drama, fantasy, and adventure. Therefore, we can say that the system achieves diversity by suggesting movies with varied themes and genres, catering to different aspects of the user’s potential interests.

Therefore, we can say that the goals of the recommender system - relevance, novelty, serendipity, diversity are met. However, since we only have limited information a more detailed analysis, including user feedback and additional metrics, would provide a more comprehensive evaluation of the recommender system’s performance in meeting these goals.

Implementation of User-Based Recommendation System

I began by implementing a user-based recommendation system using the recommenderlab library in R. The system relies on collaborative filtering to identify users with similar preferences. Here's a brief overview of the process:

Data Preparation

I converted the MovieLens data into a realRatingMatrix, which is a format suitable for recommendation algorithms. Model Training: I used the UBCF (User-Based Collaborative Filtering) method to create the recommender model.

Generating Recommendations

I generated top 10 movie recommendations for a subset of 5 users. The recommendations were outputted and formatted to show the top suggestions.

Implementation of Item-Based Recommendation System

Next, I developed an item-based recommendation system. This approach also utilizes collaborative filtering but focuses on item similarity:

Model Training

I applied the IBCF (Item-Based Collaborative Filtering) method to the same realRatingMatrix. Generating Recommendations: Similar to the user-based system, I generated top 10 recommendations for the same subset of 5 users.

Comparison of User-Based and Item-Based Recommendations

I compared the recommendations from both systems to understand their differences:

User-Based Recommendations

This method suggested movies such as Boot, Das, and Matilda for User 0. The recommendations were influenced by the preferences of users with similar tastes.

Item-Based Recommendations

For the same user, this method suggested movies like The Dorothy Day Story and Big Bang Theory. The recommendations were based on the similarity between items that the user had previously rated highly. Similar patterns were observed for other users, illustrating that user-based recommendations often focus on broader genre preferences, while item-based recommendations are more specific to individual movie attributes.

Observations

Relevance The recommendations were generally relevant, with suggestions aligning with known user interests. For instance, recommendations such as Boot and Das correspond to historical and war genres, reflecting the user's prior ratings.

Novelty The system introduced a level of novelty by suggesting lesser-known movies like Winter Guest. This aspect adds value by potentially exposing users to new content they might not have encountered otherwise.

Serendipity Recommendations such as Winter Guest and City of Lost Children showcased an element of serendipity. The system provided unexpected choices that might pleasantly surprise users, based on their past interactions.

Diversity The recommendations exhibited diversity by including movies from different genres like drama, fantasy, and family. This diversity ensures a broader range of suggestions, catering to various aspects of user interests.

Conclusion

Overall, the recommender system successfully met its goals of relevance, novelty, serendipity, and diversity. The user-based system provided a broader mix of genres based on user similarities, while the item-based system offered more specific recommendations based on item attributes. Further analysis with user feedback and additional metrics would enhance the evaluation of the system's performance and its alignment with user preferences.