Package 'CFilt'

June 19, 2020

Type Package

Title Collaborative Filtering by Reference Classes

Version 0.1.0

Author Thiago Lima, Jessica Kubrusly

Maintainer Thiago Lima <thiagoaugusto@id.uff.br>

Description

The collaborative Filtering methodology has been widely used in recommendation systems, which uses similarities between users or items to make recommendations. A class called CF was implemented,
where it is structured by matrices and composed of recommendation and database manipulation functions.
See Aggarwal (2016) <doi:10.1007/978-3-319-29659-3> for an overview.

Depends R (>= 3.5.0)

Imports methods, utils

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2020-06-19 09:50:06 UTC

R topics documented:

CF-class *A Reference Class to represent a object CF*

Description

A class of objects created structured with the following objects: the MU - Utility Matrix, the SU - A Matrix of Similarity between Users, the SI - A Matrix of Similarity between Items, and the vectors averages_u, averages_i, n_aval_u and n_aval_i. The class contains methods, general functions with the objectives of manipulating the data and making recommendations, from the structures present in the class. The data manipulation methods comprise addnewuser, addnewemptyuser, addnewitem, addnewemptyitem and newrating, while the recommendations methods recommend, kclosestitems, topkusers, topkitems are created through choices available in the Collaborative Filtering methodology. All objects and methods are accessed through the "$" character. A CF class object is created through the CFbuilder function.

Fields

`MU` A utility matrix, matrix that contains all the users ratings. The rows comprise users and the columns, itens.

`SU1` A superior triangular user similarity matrix that contains the similarities between users, calculated using Cosine similarity

`SU2` A superior triangular user similarity matrix that contains the similarities between users, calculated using Pearson Correlation.

`SI1` A superior triangular item similarity matrix that contains the similarities between items, calculated using Cosine similarity.

`SI2` A superior triangular item similarity matrix that contains the similarities between items, calculated using Adjusted Cosine similarity.

`averages_u` A vector that contains the averages of users ratings.

`averages_i` A vector that contains the averages of item ratings.

`n_aval_u` A vector that contains the numbers of ratings performed by each user.

`n_aval_i` A vector that contains the numbers of ratings received for each item.

Methods

`addnewemptyitem(Id_i)` Adds a new item without ratings. The object CF matrices and vectors will be updated. Id_i : a character, an item ID.

`addnewemptyuser(Id_u)` Adds a new user without ratings. The object CF matrices and vectors will be updated. Id_u : a character, a user ID.

`addnewitem(Id_i, Ids_u, r)` Adds a new item that has been rated by one or more users. The object CF matrices and vectors will be updated. Id_i : a character, an item ID; Ids_u : a character vector, a user IDs; r : a vector with its respective ratings.

`addnewuser(Id_u, Ids_i, r)` Adds a new user who rated one or more items. The object CF matrices and vectors will be updated. Id_u : a character, a user ID; Ids_i : a character vector, item IDs; r : a vector with its respective ratings.

`changerating(Id_u, Id_i, r)` Changes the rating from user Id_u to item Id_i. The object CF matrices and vectors will be updated. Id_u : A character, a user ID; Id_i : A character, an item ID; r : The new rating.

`deleterating(Id_u, Id_i)` Deletes the rating from user Id_u to item Id_i. The object CF matrices and vectors will be updated. Id_u : A character, a user ID; Id_i : A character, an item ID.

`estimaterating( Id_u, Id_i, type, neighbors = 5, similarity = ifelse(is_binary == TRUE, "cos", "adjcos"),` A function that returns the estimated rating for the evaluation of item Id_i by user Id_u. The recommendation can be made through similarity between users, when type = 'user', and also through the similarity between items, when type = 'item'. Id_u: A character, a user ID; Id_i: A character, an item ID; type: A character string, 'user' or 'item'; neighbors: Number of similarities used for the estimates.(Default=5); similarity: The methodology used to estimate the rating. Must be one of 'cos', for cosine similarity, or 'adjcos' (default), for adjusted cosine similarity. When type='user', 'adjcos' comprises the Pearson similarity. This choice can alter the way the estimate is calculated; is_binary: it's TRUE if your dataset contains binary ratings or FALSE (default), otherwhise. If it's TRUE, similarity will be 'cos', because 'adjcos' isn't recommended for binary ratings.

`kclosestitems( Id_i, k = 5, similarity = ifelse(is_binary == TRUE, "cos", "adjcos"), is_binary = FALSE )` A function that returns the k items most similar to an item. Id_i : A Character, a Item ID; k : Number of items most similar to item Id_i (deafult = 5); similarity: The methodology used to estimate the rating. Must be one of 'cos', for cosine similarity, or 'adjcos' (default), for adjusted cosine similarity. When type='user', 'adjcos' comprises the Pearson similarity. This choice can alter the way the estimate is calculated; is_binary: it's TRUE if your dataset contains binary ratings or FALSE (default), otherwhise. If it's TRUE, similarity will be 'cos', because 'adjcos' isn't recommended for binary ratings.

`newrating(Id_u, Id_i, r)` Adds a new rating from user Id_u to item Id_i.The object CF matrices and vectors will be updated. Id_u : a character, a user ID; Id_i : a character, an item ID; r : the rating.

`recommend( Id_u, Id_i, type, neighbors = 5, cuts = 3.5, similarity = ifelse(is_binary == TRUE, "cos", "adjc` A function that returns True if user Id_u will like item Id_i or returns FALSE, otherwise. The recommendation can be made through similarity between users, when type = 'user', as well as through the similarity between items, when type = 'item'. Id_u : a character, a User ID; Id_i : a character, an Item ID; type: a character string, 'user' or 'item'; neighbors: number of similarities used for to estimates (default = 5); cuts: cut score designated to determine if it is recommended (default=3.5); similarity: the methodology used to estimate the rating. Must be one of 'cos', for cosine similarity, or 'adjcos' (default), for adjusted cosine similarity. When type='user', 'adjcos' comprises the Pearson similarity. This choice can alter the way the estimate is calculated; is_binary: it's TRUE if your dataset contains binary ratings or FALSE (default), otherwhise. If it's TRUE, similarity will be 'cos', because 'adjcos' isn't recommended for binary ratings.

`topkitems( Id_u, k = 5, type, neighbors = 5, cuts = 3.5, similarity = ifelse(is_binary == TRUE, "cos", "adjc` A function that recommends k items for an Id_u user. The recommendation can be made through similarity between users, when type = 'user', as well as through similarity between items, when type = 'item'. Id_u : A character, a User ID; k : Number of recommendations (default=5); type: A character string, 'user' or 'item'; neighbors: Number of similarities used for the estimates(default=5); cuts: Cut score designated to determine if it is recommended (default = 3.5); similarity: The methodology used to estimate the rating. Must be one of 'cos', for

cosine similarity, or 'adjcos' (default), for adjusted cosine similarity. When type='user', 'adjcos' comprises the Pearson similarity. This choice can alter the way the estimate is calculated; is_binary: it's TRUE if your dataset contains binary ratings or FALSE (default), otherwhise. If it's TRUE, similarity will be 'cos', because 'adjcos' isn't recommended for binary ratings.

`topkusers( Id_i, k = 5, type, neighbors = 5, cuts = 3.5, similarity = ifelse(is_binary == TRUE, "cos", "adjc`
A function that indicates the k users who will like the item Id_i.The recommendation can be made through similarity between users, when type = 'user', as well as through similarity between items, when type = 'item'. Id_i : A Character, a Item ID; k : Number of recommendations (default=5); type: A character string, 'user' or 'item'; neighbors: Number of similarities used for the estimates (default=5); cuts: Cut score designated to determine if it is recommended (Default=3.5); similarity: The methodology used to estimate the rating. Must be one of 'cos', for cosine similarity, or 'adjcos' (default), for adjusted cosine similarity. When type='user', 'adjcos' comprises the Pearson similarity. This choice can alter the way the estimate is calculated; is_binary: it's TRUE if your dataset contains binary ratings or FALSE (default), otherwhise. If it's TRUE, similarity will be 'cos', because 'adjcos' isn't recommended for binary ratings.

Author(s)

Thiago Lima, Jessica Kubrusly.

References

- LINDEN, G.; SMITH, B.; YORK, J. Amazon. com recommendations: Item-toitem collaborative filtering. Internet Computing, IEEE, v. 7, n. 1, p. 76-80,2003
- Aggarwal, C. C. (2016). Recommender systems (Vol. 1). Cham: Springer International Publishing.
- Leskovec, J., Rajaraman, A., & Ullman, J. D. (2020). Mining of massive data sets. Cambridge university press.

See Also

CFbuilder

Examples

```
ratings<-movies[1:1000,]
objectCF<-CFbuilder(Data = ratings)
objectCF$MU
objectCF$SU1
objectCF$SU2
objectCF$SI1
objectCF$SI2
objectCF$averages_u
objectCF$averages_i
objectCF$n_aval_u
objectCF$n_aval_i
objectCF$addnewuser(Id_u = "Thiago",Ids_i = "The Hunger Games: Catching Fire",r = 5)
objectCF$addnewemptyuser(Id_u = "Jessica")
objectCF$addnewitem(Id_i = "Avengers: Endgame",Ids_u = c("1","2"),r = c(5,3))
```

```
objectCF$addnewemptyitem(Id_i = "Star Wars")
objectCF$newrating(Id_u = "1", Id_i = "Till Luck Do Us Part 2",r = 2)
objectCF$recommend(Id_u = "2", Id_i = "Iron Man 3", type = "user")
objectCF$recommend(Id_u = "2", Id_i = "Thor: The Dark World", type = "item")
objectCF$kclosestitems(Id_i = "Iron Man 3", k = 3)
objectCF$topkitems(Id_u = "3",k = 3, type = "user")
objectCF$topkitems(Id_u = "3",k = 3, type = "item")
objectCF$topkusers(Id_i = "Thor: The Dark World", k = 3,type = "user")
objectCF$topkusers(Id_i = "Thor: The Dark World", k = 3,type = "item")
objectCF$estimaterating(Id_u = "2",Id_i = "Iron Man 3", type = "user")
objectCF$deleterating("1","Brazilian Western")
objectCF$changerating("1","Wreck-It Ralph",2)
```

`CFbuilder` *A function to create and build a CF class object*

Description

A CF class object constructor. This function can perform two procedures: If Data is a data frame, style: User Id - Item Id - Ratings, it creates and builds an FC class object from the data frame, containing a Utility Matrix, a User Similarity Matrix, an Item Similarity Matrix, a vector with the number of user ratings, a vector with the number of ratings received for the items, a vector with the average ratings of each user and another vector with the average ratings received from each item. If Data is the Utility Matrix, it also constructs all matrices and vectors. When building the object, the progress percentage is shown. Step 1: Building the MU and vectors. Step 2: Building the SU. Step 3: Building the SI.

Usage

```
CFbuilder(Data)
```

Arguments

`Data`	A data frame by style: User ID - Item ID - Ratings, or a Utility Matrix.

Value

a CF class object.

Author(s)

Thiago Lima, Jessica Kubrusly.

References

LINDEN, G.; SMITH, B.; YORK, J. Amazon. com recommendations: Item-to-item collaborative filtering. Internet Computing, IEEE, v. 7, n. 1, p. 76-80,2003

See Also

CF-class

Examples

```
ratings<-movies[1:1000,]
objectCF<-CFbuilder(Data = ratings)
```

CFilt *CFilt: A package about Collaborative Filtering by RC in R.*

Description

The CFilt package provides one builder function CFbuilder and one class CF with methods that serve to change objects and recommend items or users.

Details

Two main goals:

- Structure the database so that changes can be made in a practical way through object-oriented programming.
- Make recommendations through choices by the Collaborative Filtering methodology in a practical, fast and efficient manner.

Author(s)

Authors:

- Jessica Quintanilha Kubrusly - jessicakubrusly@id.uff.br
- Thiago Augusto Santos Lima - thiagoaugusto@id.uff.br

---

movies *Movie ratings by users*

Description

A dataset containing 7276 ratings for 50 movies by 526 users. This database was created by Giglio (2014).

Usage

```
movies
```

Format

A data frame with 7276 rows and 3 variables:

Id Users Users identifier. Numbers 1 to 526.

Id Items Movies identifier. Movies list:

1. Iron Man 3
2. Despicable Me 2
3. My Mom Is a Character
4. Fast & Furious 6
5. The Wolverine
6. Thor: The Dark World
7. Hansel & Gretel: Witch Hunters
8. Wreck-It Ralph
9. Monsters University
10. The Hangover Part III
11. Vai Que Dá Certo
12. Meu Passado me Condena
13. We're So Young
14. Brazilian Western
15. O Concurso
16. Mato sem Cachorro
17. Cine Holliudy
18. Odeio o Dia dos Namorados
19. Argo
20. Django Unchained
21. Life of Pi
22. Lincoln
23. Zero Dark Thirty
24. Les Miserables
25. Silver Linings Playbook
26. Beasts of the Southern Wild
27. Amour
28. A Royal Affair
29. American Hustle
30. Capitain Phillips
31. 12 Years a Slave
32. Dallas Buyers Club
33. Gravity
34. Her
35. Philomena
36. The Wolf of Wall Street
37. The Hunt

38. Frozen
39. Till Luck Do Us Part 2
40. Muita Calma Nessa Hora 2
41. Paranormal Activity: The Marked Ones
42. I, Frankenstein,
43. The Legend of Tarzan
44. The Book Thief
45. The Lego Movie, , ,
46. Walking With Dinosaurs
47. The Hunger Games: Catching Fire
48. Blue Is The Warmest Color
49. Reaching for the Moon
50. The Hobbit: The Desolation of Smaug

Ratings Movie ratings by users. The ratings follows the Likert scale: 1 to 5.

References

Giglio , J. C. (2014). Recomendação de Filmes Utilizando Filtragem Colaborativa [Recommending Films Using Collaborative Filtering]. Undergraduate thesis - Universidade Federal Fluminense.

Index