MovieLens Rating Predictions

Agustin Gallo

2/4/2021

## Introduction

We can start this project by asking several questions about our surrounding world such as, How does Google give us options when we are typing a search, and we haven´t even search for anything like that before? How is it that Spotify create the famous “Your daily mix” for every single day? Or youTube and Netflix suggest a new video in order to keep you seeing more and more thing up to late hours at night?

Well, all of them are recommendation systems, and its name says the recommend YOU new things given a specific context, but, what is this context? Well, it could be how you rate previous things, or which videos you’ve been watching past week, or even if according to your profile there are items other users with a profile similar to yours have a great rating, then, is probably you will like them too, therefor its a good recommendation and a good chance to consume the suggested product.

In this project we will create a movie recommendation system using the dataset located in [linked phrase](http://files.grouplens.org/datasets/movielens). the data used for this consist in two tables, the first of them “movies” consist in a list of movies containing their ID, title and genre. The second table “rating” consist in the rating provided per user to a given movie and the time when this rating was issued.

Next you can see a glance of both tables:

#### movies table

## movieId title  
## 1 1 Toy Story (1995)  
## 2 2 Jumanji (1995)  
## 3 3 Grumpier Old Men (1995)  
## 4 4 Waiting to Exhale (1995)  
## 5 5 Father of the Bride Part II (1995)  
## 6 6 Heat (1995)  
## genres  
## 1 Adventure|Animation|Children|Comedy|Fantasy  
## 2 Adventure|Children|Fantasy  
## 3 Comedy|Romance  
## 4 Comedy|Drama|Romance  
## 5 Comedy  
## 6 Action|Crime|Thriller

#### rating table

## userId movieId rating timestamp  
## 1: 1 122 5 838985046  
## 2: 1 185 5 838983525  
## 3: 1 231 5 838983392  
## 4: 1 292 5 838983421  
## 5: 1 316 5 838983392  
## 6: 1 329 5 838983392

## Method

We will divide the structure of this project in parts \* Get movie and rating data \* First dummy attempt \* Data exploration \* Model declaration \*

## Results

## Conclusion

## Including Plots

You can also embed plots, for example:

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.