



**UNIVERSIDAD
DE GRANADA**

Gestión de Información en la Web

**Práctica 4. Desarrollo de un Sistema de
Recomendación basado en Filtrado Colaborativo**

2019-2020

Máster en Ingeniería Informática

Lidia Sánchez Mérida

lidiasm96@correo.ugr.es

76065456-V

Índice

Descripción del problema	3
Desarrollo del sistema	3
Manual de usuario	5

Descripción del problema

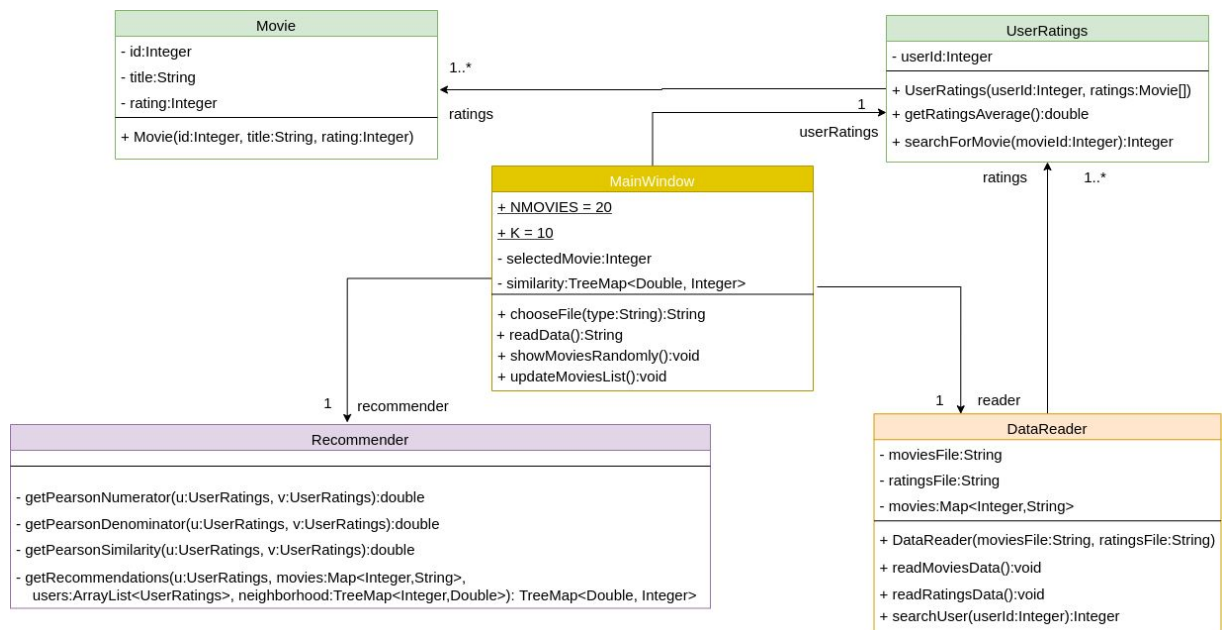
El objetivo de esta práctica consiste en implementar un **Sistema de Recomendación basado en Filtrado Colaborativo** desde cero. Para ello el sistema deberá, en primer lugar, proponer un conjunto de 20 películas al usuario activo para que este sea capaz de valorarlas de 1 a 5 estrellas, representando que no le ha gustado nada hasta que es su película favorita, respectivamente.

Por otro lado también dispondremos de valoraciones de diversas películas realizadas por otro conjunto de usuarios utilizando la misma representación con estrellas. A partir de ellos se podrá extraer cuán similares son sus gustos con respecto a los del usuario activo mediante el estudio de las valoraciones que han realizado a las mismas películas.

Finalmente, dentro de las películas que el usuario activo no haya valorado, se le recomendarán aquellas que sean más afines a su gusto. Para ello nos basaremos en las películas que han valorado los usuarios pertenecientes a su vecindario y solo mostraremos aquellas que tengan de 4 a 5 estrellas, para aumentar el número de posibilidades de acertar con las películas recomendadas.

Desarrollo del sistema

A continuación adjunto una imagen con el diagrama de clases simplificado que representa el diseño del sistema.



Tal y como se puede observar el diagrama consta de hasta cinco clases. En primer lugar se encuentran aquellas que representan las estructuras de datos en las que he organizado la información de los ficheros. La primera denominada **Movie** es la que representa los datos de las películas y por ende, dispone de un campo para almacenar el identificador y otro para

guardar el título de la película. No obstante, como todas las películas van a disponer de una puntuación también he añadido un tercer campo para guardarla. De este modo puedo plantear la clase **UserRatings** de una forma más sencilla, almacenando en un campo su identificador y relacionando esta clase con la anterior para que **cada usuario disponga de las películas que ha valorado**. Dentro de esta última clase, además, se añaden dos métodos: el primero para calcular la media de las valoraciones de un usuario y el segundo para encontrar la posición de una determinada película, si ha sido valorada por el usuario. Este último método será utilizado al obtener las recomendaciones tanto para comprobar que el usuario no ha visto la película como para verificar que sus vecinos sí que la han visto y así predecir la puntuación que el usuario activo le otorgaría para recomendarla o no.

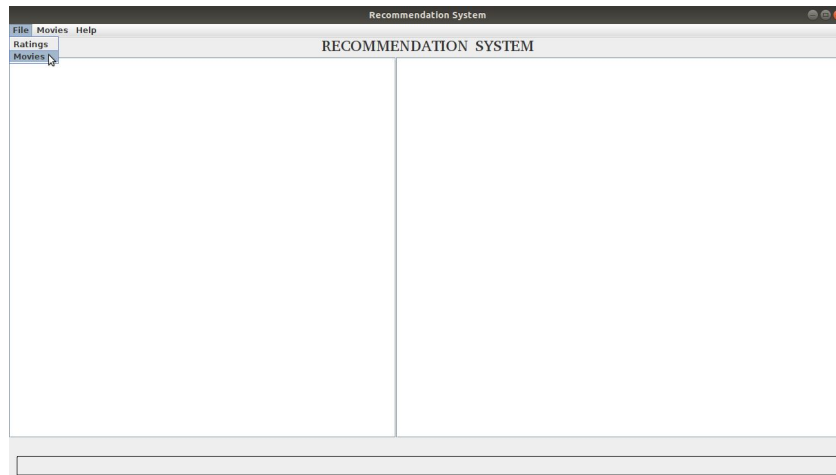
En segundo lugar nos encontramos con la clase **DataReader** que es la encargada de leer los ficheros de películas y valoraciones y almacenarlos en las dos estructuras explicadas anteriormente. En primer lugar se leen las películas, almacenándolas en un mapa cuya clave es su identificador y el valor es el título. Esta estructura de datos será utilizada al leer, a continuación, el fichero con las valoraciones de los usuarios. De esta forma se irán añadiendo a cada usuario las películas que ha valorado almacenando su puntuación y la información que se encuentra en el mapa anterior.

La penúltima clase denominada **Recommender** es la encargada de calcular la similitud entre el usuario activo y el resto así como proponer películas al usuario. Para el primer objetivo se ha utilizado la correlación de *Pearson*, tal y como se explica en las transparencias de clase, es decir, codificando la fórmula asociada a la similitud entre dos usuarios. Mientras que para proponer las películas se ha hecho uso de la fórmula capaz de predecir la puntuación del usuario activo frente a una película que no ha visto pero que su vecindario sí.

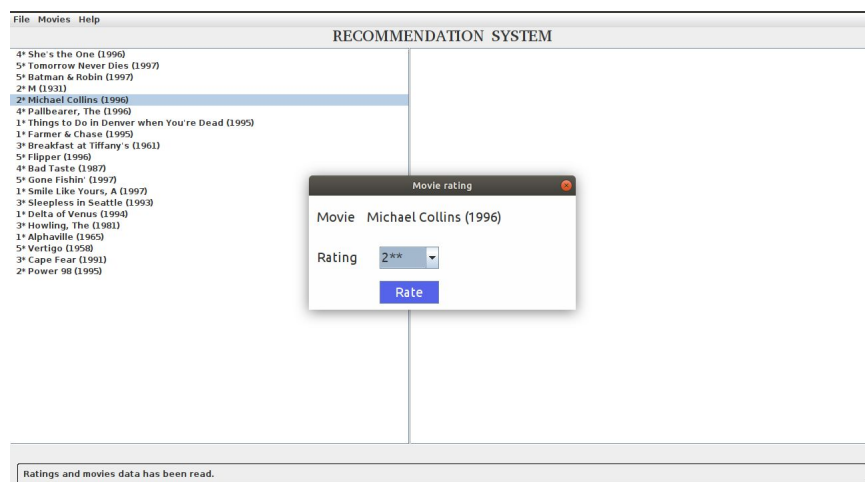
Por último se encuentra la clase **MainWindow** que es la interfaz del sistema y por la cual se pueden llevar a cabo todos los métodos descritos anteriormente, desde la visualización de las veinte películas a valorar hasta mostrar las películas sugeridas.

Manual de usuario

Una vez ejecutado el *jar* aparecerá la ventana principal del sistema. En ella podemos visualizar un menú superior con tres opciones. Para cargar los ficheros con los datos de las películas y sus respectivas valoraciones realizadas por los usuarios, debemos especificar sendos archivos mediante las opciones *Movies* y *Ratings*, respectivamente.



Una vez hayan sido leídos los datos, de forma automática se mostrará en la división izquierda las veinte películas que el usuario puede valorar. Si bien este proceso se puede realizar manualmente pulsando sobre cualquiera de las películas y cambiando su valoración con la ventana de diálogo que se muestra en la siguiente captura, para agilizar las pruebas del sistema se ha simulado la valoración del usuario asignando una puntuación aleatoria a cada película. La valoración de cada una se puede observar al lado de su título.



Por último, para activar el sistema de recomendación debemos pulsar sobre la opción oculta dentro del segundo ítem del menú superior, tal y como se muestra en la siguiente captura. A continuación aparecerán las películas recomendadas automáticamente en la división derecha. Del mismo modo que anteriormente, al lado del título de cada película aparecerá la puntuación predicha que el usuario podría otorgarle a la misma.

Recommendation System

File

Movies

Help

Suggest movies

RECOMMENDATION SYSTEM

4* She's the One (1996)

5* Tomorrow Never Dies (1997)

5* Batman & Robin (1997)

2* M (1931)

2* Michael Collins (1996)

4* Pallbearer, The (1996)

1* Things to Do in Denver when You're Dead (1995)

1* Farmer & Chase (1995)

3* Breakfast at Tiffany's (1961)

5* Flipper (1996)

4* Bad Taste (1987)

5* Gone Fishin' (1997)

1* Smile Like Yours, A (1997)

3* Sleepless in Seattle (1993)

1* Delta of Venus (1994)

3* Howling, The (1981)

1* Alphaville (1965)

5* Vertigo (1958)

3* Cape Fear (1991)

2* Power 98 (1995)

4* Old Lady Who Walked in the Sea, The (Vieille qui marchait dans la mer, La) (1991)

4* City Slickers II: The Legend of Curly's Gold (1994)

4* Godfather: Part II, The (1974)

4* Once Were Warriors (1994)

4* Stargate (1994)

4* Homeward Bound II: Lost in San Francisco (1996)

4* Robin Hood: Prince of Thieves (1991)

4* American in Paris, An (1951)

4* One Flew Over the Cuckoo's Nest (1975)

4* Silence of the Lambs, The (1991)

4* Arsenic and Old Lace (1944)

4* Everest (1998)

4* Casablanca (1942)

4* Braveheart (1995)

4* Annie Hall (1977)

4* Three Colors: Red (1994)

4* Stand by Me (1986)

4* Titanic (1997)

Ratings and movies data has been read.