

UNIVERSIDAD DE GRANADA
E.T.S.I. INFORMÁTICA Y TELECOMUNICACIÓN



Departamento de Ciencias de la
Computación e Inteligencia Artificial

Gestión de Información en la Web

Guión de Prácticas

Práctica 4:
Desarrollo de un Sistema de Recomendación basado
en Filtrado Colaborativo

Curso 2016-2017

Máster en Ingeniería Informática

Práctica 4

Desarrollo de un Sistema de Recomendación basado en Filtrado Colaborativo

1. Objetivos

1. Entender el proceso de recomendación basado en filtrado colaborativo.
2. Ser capaz de plasmarlo en un programa de ordenador.

2. Trabajo a Realizar

En esta práctica se construirá, partiendo de cero y en el lenguaje preferido, un sistema de recomendación de películas basado en filtrado colaborativo (usuario - usuario). Para tal fin, se desarrollará una aplicación que muestre al usuario 20 películas al azar y éste las evalúe asignándole un valor de 1 estrella (*) , indicando que no le gusta nada, hasta 5 estrellas (es una de sus películas favoritas).

Una vez obtenidas las valoraciones del usuario, el sistema procederá a calcular el vecindario, es decir, el grupo de usuarios que más se parecen a él en cuanto a las películas vistas y a las valoraciones dadas.

Finalmente, la aplicación predecirá la valoración para el usuario activo de todas las películas que han visto sus vecinos y que no ha visto el usuario activo, y le mostrará aquellas predichas con cuatro o cinco estrellas.

3. Documentación y Ficheros a Entregar

La práctica se subirá a la plataforma docente de Prado2 en forma de fichero comprimido (en formato zip o rar), con el nombre “practica4_apellidos_nombre.zip”. Este fichero contendrá a su vez todo el código fuente, dos ficheros jar, uno para el indexador y otro para el recuperador, y por último, un archivo pdf con la documentación.

La **documentación** de la práctica será un fichero *pdf* que deberá incluir, al menos, el siguiente contenido:

- a) Portada con el número y título de la práctica, el curso académico y el nombre, DNI y dirección e-mail del alumno.

- b) La documentación de la práctica contendrá una breve descripción de la misma, explicando cómo se ha hecho, los elementos innovadores y un pequeño manual de usuario.
- c) Referencias bibliográficas u otro tipo de material distinto del proporcionado en la asignatura que se haya consultado para realizar la práctica (en caso de haberlo hecho).

La fecha de entrega será el miércoles, 18 de mayo, como máximo justo antes de la hora de clase. En esa clase de prácticas procederéis a mostrar vuestras aplicaciones al resto de compañeros y a explicar cómo la habéis desarrollado.

Aunque lo esencial es el contenido, también debe cuidarse la presentación y la redacción.

4. Evaluación

La evaluación de esta práctica tendrá en cuenta la calidad de los siguientes aspectos:

- Adecuación a los requerimientos de las aplicaciones propuestas.
- Inclusión de aspectos innovadores no vistos en clase.
- Documentación de la práctica.

5. Logística

La práctica se realizará con la colección MovieLens, compuesta por 100.000 valoraciones de 943 usuarios sobre 1682 películas. En Prado2 está alojado un fichero denominado `ml-data.tar.gz`, el cual contiene los ficheros `u.data`, con todas las valoraciones (formato: `idUser idMovie valoración timestamp`), y `u.item`, con información de las películas (`idMovie | título | ...`). El resto de ficheros no nos hará falta, aunque para conocer más detalles de la colección, véase el fichero `README`.

Las películas que se presentarán al azar al usuario se obtendrán del fichero `u.item` y los vecindarios se crearán a partir de las valoraciones que ha hecho el usuario activo y de las contenidas en `u.data`. Se pueden considerar vecindarios de tamaño 10.