

# *Práctica nº 4(Voluntaria)*

## *Inteligencia de Negocio y en la Web*

Universidad de Jaén  
Departamento de Informática

**Implementación elementos de un Sistema de  
Recomendación basado en Filtrado Colaborativo**

## OBJETIVO

En esta práctica se pretende que el alumno estudie y conozca las distintas fases de la técnica de filtrado colaborativo basado en factorización de matrices. Como existen distintos modelos para construir las matrices factorizadas para el sistema de recomendación. En esta práctica vamos a utilizar el modelo más ampliamente utilizado como es el *gradiente estocástico descendiente* (una explicación detallada de su implementación puede encontrarse en <http://sifter.org/~simon/journal/20061211.html>).

## CONSIDERACIONES GENERALES

En esta práctica los alumnos trabajarán de forma individual.

Al igual que en prácticas anteriores el alumno utilizará la base de datos:

- *rs-movie-cour*

Dicha BD se utilizará para construir el sistema de recomendación.

## EJERCICIOS PROPUESTOS

Realizar los siguientes ejercicios:

1. Acceso a los datos
  - a) Crear los objetos para almacenar los datos proporcionados en los ficheros (usuarios, productos, tags y ratings)
  - b) Crear las interfaces que acceden a los ficheros para obtener los datos de los ficheros y así crear los objetos para usuarios, productos, tags y ratings
  - c) Comprobar que el acceso a los datos es correcto.
2. Implementación de un módulo que calcule las matrices factorizadas con 40 factores y los parámetros descritos en <http://sifter.org/~simon/journal/20061211.html>
3. Implementación del Sistema de Recomendación Colaborativo que devuelva una lista de N predicciones.
4. Prueba y evaluación del sistema:
  - a) Para comprobar el funcionamiento del sistema, los valores a recomendar que se deberían obtener son los siguientes:

**Top 10 recommendations for user 1:**

8,Gladiator (2000),4.925  
862,Ty Story (1995),4.400  
807,Seven (a.k.a. Se7en) (1995),4.321  
1422,The Departed (2006),4.306  
550,Fight Club (1999),4.257  
77,Memento (2000),4.235  
11,Star Wars: Episode IV - A New Hope (1977),4.167  
1892,Star Wars: Episode VI - Return of the Jedi (1983),4.134  
155,The Dark Knight (2008),4.132  
1891,Star Wars: Episode V - The Empire Strikes Back (1980),4.109.

**Top 10 recommendations for user 500:**

238,The Godfather (1972),4.522  
77,Memento (2000),4.411  
107,Snatch (2000),3.959  
1422,The Departed (2006),3.954  
585,Monsters Inc. (2001),3.855  
752,V for Vendetta (2006),3.760  
12,Finding Nemo (2003),3.753  
9806,The Incredibles (2004),3.694  
194,Amelie (2001),3.659  
680,Pulp Fiction (1994),3.640

**IMPORTANTE:** Leer detenidamente tanto este apartado como el de “RESULTADOS A ENTREGAR”.  
antes de comenzar a resolver los ejercicios.

# RESULTADOS A ENTREGAR

Cada alumno debe entregar:

- Un archivo zip que contenga:
  - o Un documento que explique cómo ejecutar el sistema de recomendación
    - En el caso de utilizar un sistema de gestión de base de datos (no obligatorio), es necesario explicar su configuración para la carga de datos y acceso.
  - o Los archivos fuente del proceso de optimización del ejercicio 2 y un archivo ejecutable del mismo.
  - o Los archivos fuente del sistema de recomendación implementado en 3 y un archivo ejecutable del mismo.
  - o Un documento que indique cómo funciona el sistema y especifique las pruebas del sistema realizadas.
- Una explicación de las mayores dificultades encontradas en el desarrollo del sistema