



Taller 3 – Últimas tendencias en Sistemas de Recomendación

Objetivos

- Desarrollar un sistema de recomendación utilizando las tendencias vistas en la última parte del curso según la modalidad de trabajo escogida por el grupo.

Prerrequisitos

- Conocimiento básico de programación en los lenguajes sugeridos: Java o Python
- Conocimiento básico de Unix
- Si se realiza análisis de grafos se sugiere el uso de bases de datos para manejo de grafos ([Neo4j](#))

Herramientas

- Máquina virtual Unix
- Ambiente de desarrollo de software
- *Dataset* de pruebas

Enunciado

1) Conocimiento de los *dataset* de trabajo

El taller se trabaja sobre el conjunto de datos MovieLens Latest, que corresponde a las valoraciones que realizan usuarios sobre películas, calificadas explícitamente y anotadas mediante etiquetas libres.

- a) Ubique los datos en <http://grouplens.org/datasets/movielens/>, siguiendo el enlace <http://grouplens.org/datasets/movielens/latest/> de septiembre de 2018, **full dataset** de 27 millones de ratings
- b) En la documentación complementaria encuentra la información técnica y práctica para el uso de los datos (<http://files.grouplens.org/datasets/movielens/ml-latest-README.html>).

2) Construcción de grafo de conocimiento

- a) Explore procesos de enriquecimiento ontológico sobre la información disponible sobre las películas y sobre atributos que encuentre descritos en dbpedia. Construya su propio grafo de información usando conceptos de *Linked Open Data*. El modelo debe incluir como mínimo los actores que participan en la película y los directores, sin embargo se espera que explore atributos que permitan incrementar la serendipia de las recomendaciones. Verifique cómo la información disponible en <http://sisinflab.poliba.it/semanticweb/lod/recsys/datasets/> le puede ayudar. Otra fuente que puede ser útil, es la descarga de información del sitio web IMDB.org para uso no comercial (<https://www.imdb.com/interfaces/>). Se sugiere el uso de la base de datos neo4j para representar el grafo de conocimiento.

3) Construcción del sistema de recomendación

- a) Se desea implementar un sistema de recomendación de películas utilizando la información del grafo de conocimiento y filtrado colaborativo. **El sistema debe, además de proveer recomendaciones para cualquier usuario presente en el dataset, proveer recomendaciones para un usuario nuevo.**
- b) El sistema desarrollado debe ser un modelo híbrido, que combine filtrado colaborativo (puede ser mediante grafos también) y filtraje ontológico, aplicando una estrategia de hibridación que sea apropiada para el objetivo planteado.
- c) Describa formalmente su modelo de recomendación (Framework, arquitectura del sistema híbrido y la de sus componentes).

4) Sintonización y evaluación del modelo

- a) Establezca un esquema de experimentación y de evaluación del modelo de recomendación propuesto, que permita identificar qué tan bien se cumple el objetivo de recomendación propuesto y, en particular, cómo se tienen en cuenta los aspectos de enriquecimiento semántico.
- b) Además de las métricas que considere relevantes para su objetivo de recomendación propuesto, la evaluación debe realizarse utilizando las medidas de ranking con listas de 5 y 10 *items* recomendados: P@5, P@10. Describa formalmente su manera de determinar si un elemento es relevante para un usuario.

5) Construya una aplicación Web sencilla que permita interactuar con su sistema de recomendación y con la experimentación solicitada.

- a) La aplicación debe permitir interactuar adecuadamente con la información, de manera que constituya una aplicación Web interesante para un usuario final. La aplicación debe permitir visualizar y evaluar la recomendación construida, con información suficiente para entender clara y eficientemente los resultados.

6) Análisis de resultados

Elabore un informe en el que describa el modelo desarrollado, los parámetros de sintonización, sus resultados de experimentación, los resultados de evaluación y concluya sobre los resultados obtenidos.

El informe debe tener una extensión máxima de 12 páginas y debe incluir todos los aspectos logrados y no logrados del trabajo, así como información suficiente para la evaluación integral de lo realizado. Debe incluir, al inicio del informe, el enlace al sitio Web de resultados.

👉 Recuerde incluir los créditos y referencias de uso de los dataset.

Metodología

- El taller se desarrolla en los grupos de estudiantes que están previstos para el desarrollo de trabajos prácticos.
- Se espera una participación equitativa de los integrantes del grupo

Entrega de taller

Fecha y hora límite: domingo 5 de junio, medianoche.

Archivo de entrega: **<Taller3_NN_login1_login2_login3 >.zip**. NN es el número del grupo y los *logins* corresponden a los integrantes participantes en el desarrollo.

Contenido del archivo:

- Informe de resultados, en formato .pdf, que sigue la plantilla de informes técnicos prevista para el curso.
 - Nombre del archivo: **<Taller3_NN_login1_login2_login3 >.pdf**
- Carpeta eval: Archivos adjuntos que permiten documentar y apoyar eventualmente la evaluación de resultados.
- Carpeta src: Archivos del código fuente desarrollado.
- Archivo README.txt con la descripción de integrantes y contenido del archivo ZIP.

El no seguimiento del formato de entrega del taller tiene una penalización de **0.5/5.0** en la nota final.

La evaluación de este taller es individual, presencial y exige demostración práctica de los resultados obtenidos. La no asistencia o no demostración efectiva de los resultados se ve reflejada en una nota de 0.0/5.0.