



Grado en Ingeniería Informática, Universidad de León.

Cuaderno de trabajo

Sistema de recomendación de libros

Realizado por David González García

Diciembre de 2019

Semana 16 – 22 de Septiembre

- Creacion de cuenta en github
- Primera toma de contacto con Neo4J

Semana 23 – 29 de Septiembre

- Inicio curso de introducción a Neo4J

Semana 30 – 6 de Octubre

- Finalización curso de introducción a Neo4j
- Primer Brain-storm sobre que proyecto hacer:
 - Recomendador de libros
 - Recomendador de que visitar en una ciudad
- Búsqueda de información sobre grafos de conocimiento

Semana 7 – 13 de Octubre

- <https://thenextweb.com/podium/2019/06/11/what-is-a-knowledge-graph-and-how-does-one-work/>
- Búsqueda de información sobre grafos de conocimiento
- Búsqueda de información sobre IA

Semana 14 – 20 Octubre

- Primera presentación del trabajo: Recomendador de libros
 - Mediante las notas otorgadas por otros usuarios el programa deberá de ser capaz de devolver un libro a un usuario que le pida una recomendación.
- Creación de la base de datos a utilizar

Semana 21 – 27 Octubre

- Creación de los primeros pobladores de la base de datos
- Busqueda de información sobre el funcionamiento de los sistemas de recomendación
 - <https://promocionmusical.es/como-funcionan-algoritmos-recomendacion-spotify>
 - <https://www.analiticaweb.es/primeros-pasos-sistemas-recomendacion/>
 - <https://www.genbeta.com/web/asi-funcionan-las-recomendaciones-de-amazon>
 - <https://www.antevenio.com/blog/2018/12/algoritmo-seo-de-amazon/>

Semana 28 – 3 Noviembre

- Se decide como estará estructurada la base de datos
 - Nodos de Libros, Escritores y Lectores
 - Se descarta la idea de que Escritores y Lectores sean de la misma clase “persona”
 - Se decide introducir los géneros como Nodo y no como atributo de los libros
 - Se pospone la introducción de nodos de País
- Una vez diseñada la estructura, se procede a crear una pequeña base de datos de ejemplo y a poblarse con unos pocos nodos

Semana 4 – 10 Noviembre

- Con la base de datos de ejemplo se realizan pruebas con queries para comprobar cómo se navega por el grafo y se pueden alcanzar unos nodos desde otro. Estas queries serán utilizadas en el futuro por el algoritmo de recomendación.
- Se prueban otras queries de búsqueda que al final son descartadas para el algoritmo, pero que podrían añadirse en el futuro
- Búsqueda de algoritmos de filtrado colaborativo: como funcionan y como podría aplicarlos a mi base de datos

Semana 11 – 17 Noviembre

- Se acude a Kaggle para encontrar un dataset que se ajuste a lo que buscamos (que contenga al menos libros y escritores)
- Se encuentran varios datasets que se ajustan a lo que buscamos, pero que contienen errores o información incompleta
- Una vez tenemos nuestro dataset, este es depurado a mano para eliminar algunos libros duplicados, libros con varios autores (autor y alias) o colecciones de libros, las cuales no tenemos pensado introducir en la base de datos
- Se añaden los datos

Semana 18 -24 Noviembre

- Se crean constraints en la base de datos
- Se comprueba que no existan nodos duplicados
- Se inicia el desarrollo del sitio web que permita interactuar con la base de datos
- Una vez comprobada que la base de datos no contiene ningún elemento que nos pueda complicar en el futuro, pasamos a descargar las herramientas necesarias para comenzar con el desarrollo del sitio web

Semana 25 – 1 Diciembre

- Se diseña el aspecto principal del sitio web
- Se añaden las primeras funcionalidades de crear y borrar usuarios, así de localizar los vecinos del mismo
- Se crean las funciones de añadir y borrar cualquier nodo y relaciones entre los mismos
- Finalmente se decide la implementación del algoritmo de Slope One

Semana 2 – 8 Diciembre

- Comienza la implementación del algoritmo Slope One
- Se detectan las primeras debilidades en el algoritmo
- Se decide acotar el número de vecinos e ignorar a aquellos cuyas notas sean diferentes en más de +1 o -1 puntos
- Se crea el algoritmo de detección de comunidad por coeficiente de relación

Semana 9 – 15 Diciembre

- El algoritmo sigue dando resultados que si bien sobre el papel son correctos, no parecen lógicos
- Se estudia más a fondo los algoritmos de recomendación para intentar encontrar alguna manera de hacerlo más aceptado
- Se decide utilizar el Coeficiente de Relación calculado a la hora de decidir los vecinos más cercanos para influir en las notas, obteniendo un resultado mucho más parecido al que daría una persona normal leyendo los datos de la tabla
- Se estudian que otras funcionalidades y algoritmos podrían implementarse en el futuro.

Semana 15 – 18 Diciembre

- Se dan los últimos retoques al algoritmo
- Se comprueba que todo el sitio web funcione bien

Git Hub:

[https://github.com/DavidGG98/Book Recomendation-NEO4J-](https://github.com/DavidGG98/Book_Recomendation-NEO4J-)