

GCO.

Sistemas de recomendación.

Métodos de filtrado colaborativo.

Víctor Rodríguez Dorta alu0101540153

ALEJANDRO RODRÍGUEZ MEDEROS alu0101413938

MARIO GUERRA PÉREZ alu0101395036

1. Introducción	3
2. Matriz de Utilidad empleada para ejemplo	3
3. Proceso de Filtrado Colaborativo	4
a. Métricas	4
b. Selección de vecinos	4
c. Tipos de predicción	5
4. Ejemplo de ejecución	5
4.1 Ejemplo con matriz 10-25	7
4.2 Ejemplo con matriz de 100x1000	8
4.3 Ejemplo con matriz de 25x100	9
4.3 Ejemplo con matriz de 5x10	10
4.4 Ejemplo con matriz de 50x250	11
5. Conclusiones	13

1. Introducción

En esta práctica hemos desarrollado un sistema de recomendación basado en filtrado colaborativo, una técnica muy utilizada en inteligencia artificial y en la gestión del conocimiento. El filtrado colaborativo parte de una idea sencilla: si dos personas han valorado de forma parecida ciertos ítems, es probable que también coincidan en sus opiniones sobre otros ítems que aún no han probado. Básicamente, el sistema intenta predecir qué cosas podrían gustarle a un usuario basándose en las valoraciones de otros usuarios con gustos similares. Lo interesante de este método es que no necesita saber nada sobre las características de los ítems (como el género, el autor o el tipo), sino que se centra únicamente en los patrones de comportamiento de los usuarios.

Para implementar el proyecto, decidimos crear una aplicación web con Vue.js, un framework de JavaScript que facilita bastante el desarrollo de interfaces dinámicas y reactivas. Vue nos ha venido genial para organizar el código, separando la parte visual de la lógica y los datos. Además, nos ha permitido mostrar la información de forma clara e interactiva, como la matriz de utilidad, las similitudes entre usuarios o las predicciones calculadas.

También usamos Bun como entorno de ejecución de JavaScript. Bun es una alternativa más moderna a Node.js, pensada para ser más rápida y sencilla de usar. Trae integrado su propio gestor de paquetes, bundler y servidor, lo que nos permitió levantar el proyecto y gestionar las dependencias sin complicarnos demasiado. Con un solo comando podíamos instalar, ejecutar o compilar.

2. Matriz de Utilidad empleada para ejemplo

La matriz que se ha empleado para probar como ejemplo en el proyecto es la siguiente. Cabe destacar que cada fila representa un usuario y que cada columna un item, teniendo en cuenta que el símbolo "-" indica una valoración desconocida.

	Item 1	Item 2	Item 3	Item 4	Item 5	Mean
User 1	5	3	4	4	-	4.00
User 2	3	1	2	3	3	2.40
User 3	4	3	4	3	4	3.60
User 4	3	3	1	5	4	3.20
User 5	1	4	4	2	1	2.40
Mean	3.20	2.80	3.00	3.40	3.00	

3. Proceso de Filtrado Colaborativo

a. Métricas

Para predecir la valoración faltante, primero se calcula la similaridad entre el usuario objetivo y los demás usuarios. Las métricas más comunes son:

 Correlación de Pearson: mide la relación lineal entre las valoraciones de dos usuarios, ajustando por la media de cada uno.

$$sim(u,v) = \frac{\sum_{i \in S_{uv}} (r(u,i) - \overline{r}(u)) \cdot (r(v,i) - \overline{r}(v))}{\sqrt{\sum_{i \in S_{uv}} (r(u,i) - \overline{r}(u))^2} \sqrt{\sum_{i \in S_{uv}} (r(v,i) - \overline{r}(v))^2}}$$

 <u>Distancia Coseno:</u> mide el ángulo entre los vectores de valoraciones, ignorando la magnitud.

$$sim(u,v) = \frac{\sum_{i \in S_{uv}} r(u,i) \cdot r(v,i)}{\sqrt{\sum_{i \in S_{uv}} (r(u,i))^2} \sqrt{\sum_{i \in S_{uv}} (r(v,i))^2}}$$

 Distancia Euclídea: mide la distancia geométrica directa entre los vectores de valoraciones.

$$d(u,v)_{euc} = \sqrt{\sum_{p \in P} (r(u,i) - r(v,i))^2}$$

Cada métrica puede dar lugar a diferentes vecinos más cercanos, por lo que puede generar diferentes predicciones.

b. Selección de vecinos

Se selecciona los K usuarios más similares que hayan valorado el ítem que se quiere predecir. Por ejemplo, si se toma un k=2, eso quiere decir que se seleccionan los dos usuarios con mayor similaridad respecto al usuario objetivo.

c. Tipos de predicción

En este caso, se puede realizar de dos maneras:

 Predicción simple: se calcula como la media ponderada de las valoraciones de los vecinos, usando la similaridad como peso

$$\hat{r}(u,i) = \frac{\sum_{v \in N_u^k} sim(u,v) \cdot r(v,i)}{\sum_{v \in N_u^k} |sim(u,v)|}$$

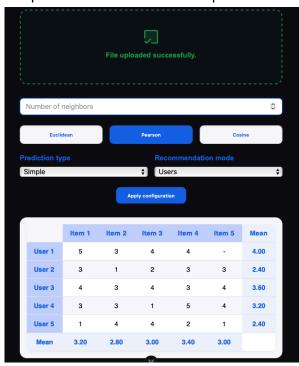
 Diferencia con la media: se ajusta a predicción sumando la media del usuario objetivo y la media ponderada de las restas respecto a la media de los vecinos

$$\hat{r}(u,i) = \overline{r}(u) + \frac{\sum_{v \in N_u^k} sim(u,v) \cdot (r(v,i) - \overline{r}(v))}{\sum_{v \in N_u^k} |sim(u,v)|}$$



4. Ejemplo de ejecución

Lo primero que se hace es introducir el ".txt" con el ejemplo. Seguidamente se indica el número de vecinos con el que se quiere llevar a cabo la predicción, la métrica con la que se quiere calcular y el tipo de predicción para realizar el ajuste. Cabe destacar que el último apartado, que es el modo de recomendación, está implementado para una posible escalabilidad futura, por lo que se debe de marcar siempre como "users".



Una vez se presiona ejecutar, lo primero que se realiza es la identificación de la valoración a recibir.

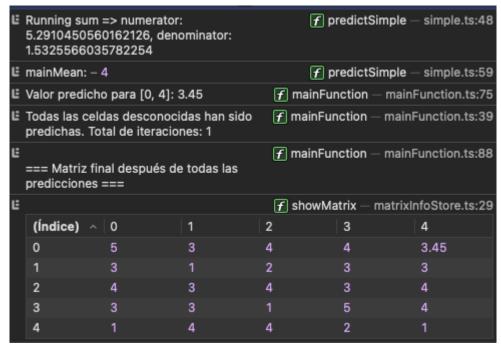
Seguido de esto, se comparan las valoraciones del Usuario1 (en este caso) con las de los otros usuarios usando en este caso, la métrica Pearson. Después de realizar esta comparación, se obtiene una tabla de similaridades.

Ahora se eligen los K usuarios con mayor similaridad que hayan valorado Item5, en este caso, ya que es donde se encuentra la incógnita. Como en este caso se ha seleccionado dos vecinos, estos son los más próximos:

```
{\color{red} {f}} getNeighbor... — getNeighborRating.ts:19
                                                                                                 ▼ 0 handler: MutableReactiveHandler
O handler: MutableReactiveHandler
                                                                                                       B _isReadonly: false
     B _isReadonly: false
                                                                                                       B _isShallow: false
    B _isShallow: false
                                                                                                    ▶ Prototipo de MutableReactiveHandler
   ▶ Prototipo de MutableReactiveHandler
                                                                                                  o target: Array (5)
▼ 0 target: Array (5)
                                                                                                          ▶ {value: 4, row: 2, col: 0}
      ▶ {value: 3, row: 1, col: 0}
                                                                                                      1 ▶ {value: 3, row: 2, col: 1}
2 ▶ {value: 4, row: 2, col: 2}
                                                                                                         ▶ {value: 3, row: 2, col: 3}
▶ {value: 4, row: 2, col: 4}
      ▶ {value: 3, row: 1, col: 3}
▶ {value: 3, row: 1, col: 4}
                                                                                                    ▶ Prototipo de Array
  ▶ Prototipo de Array
```

Como se observa en las siguientes capturas, estos dos vecinos tienen este valor de similaridad respectivamente:

Seguidamente, se calcula la predicción. En este caso, se calcula con la predicción simple, tal y como se indicó en un principio. Para ello, se multiplica la valoración de cada vecino sobre Ítem5 por su similaridad con Usuario1 y luego, se suman estos valores y se dividen por la suma de las similaridades.



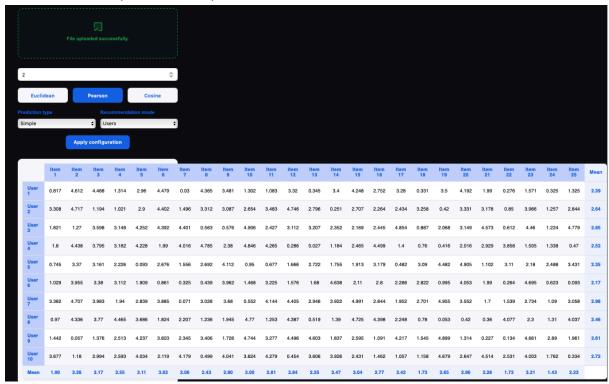
Como se observa, se calcula el numerador y seguidamente el denominador, dando como resultado que el valor predicho para este caso es <u>3.45</u>.

Para concluir, ya que se especificó que los valores en este ejemplo van entre 1 y 6, al haber predicho un 3.45 se concluye que se recomienda Item5 al Usuario1.

4.1 Ejemplo con matriz 10-25



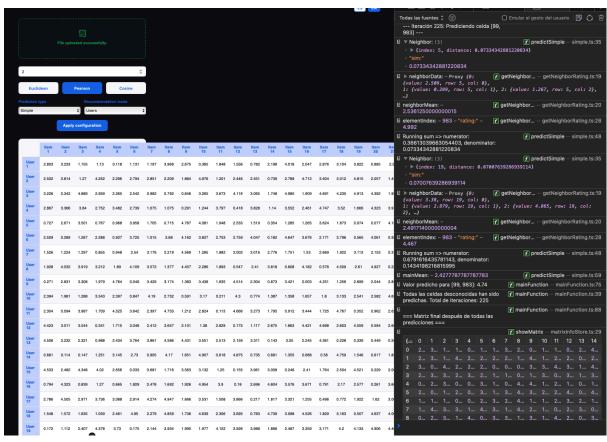
Esta es la matriz que se va a implementar, la cual es de 10x25.



Una vez aplicado el algoritmo, este es el resultado de la matriz.

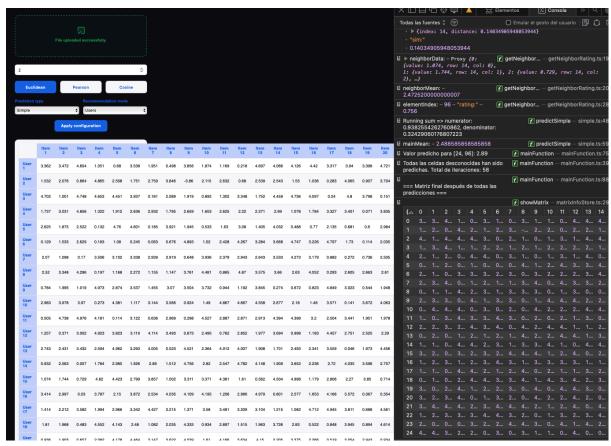
En esta captura se contempla el total de interacciones hasta completar todas las incógnitas

4.2 Ejemplo con matriz de 100x1000



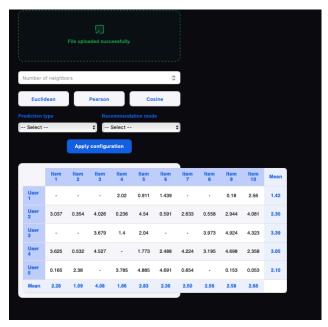
En este ejemplo se puede observar el resultado de haber evaluado una matriz de 100x1000. Como se contempla en la captura, ha calculado un total de 225 incógnitas hasta llegar al resultado final.

4.3 Ejemplo con matriz de 25x100

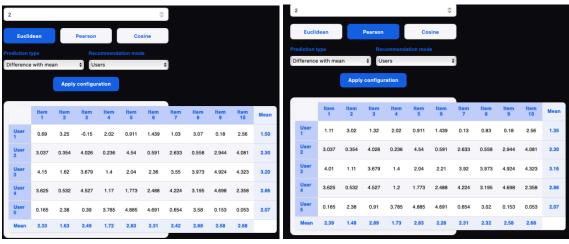


En este ejemplo se puede observar el resultado de haber evaluado una matriz de 25x100. Como se contempla en la captura, ha calculado un total de 58 incógnitas hasta llegar al resultado final.

4.3 Ejemplo con matriz de 5x10

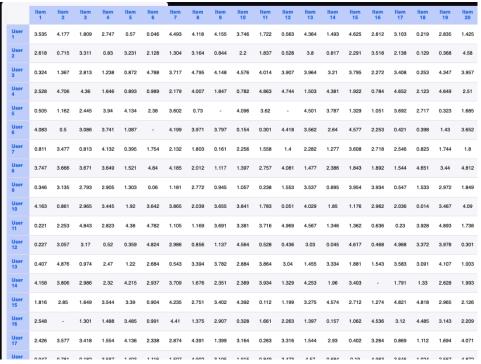


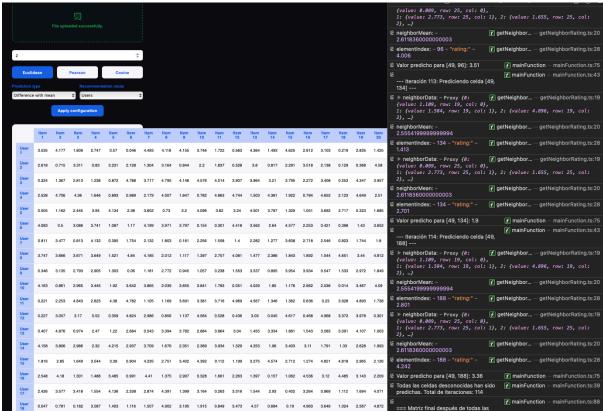
En este caso, se prueba con una matriz de 5x10.



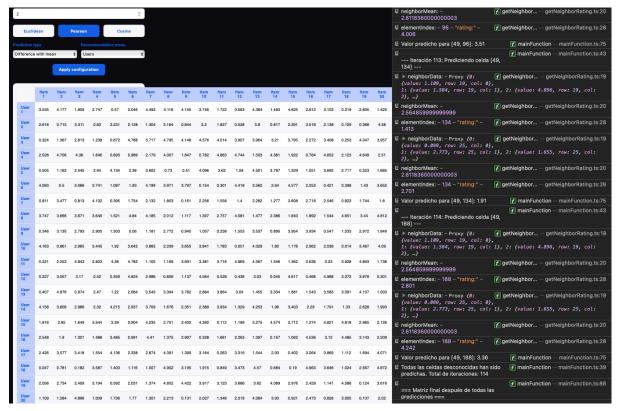
Como se observa, hay pequeñas variaciones dependiendo del tipo de métrica que se emplea en cada caso.

4.4 Ejemplo con matriz de 50x250





En este ejemplo, se evalúa una matriz de 50x250, dando como resultado que se han realizado 114 iteracciones.



Al implementar el algoritmo con otra métrica, se contempla que los resultados varían ligeramente.

5. Conclusiones

En conclusión, la implementación de un sistema de recomendación basado en filtrado colaborativo ha permitido predecir valoraciones y generar recomendaciones personalizadas de manera eficiente, demostrando la utilidad de este enfoque para mejorar la experiencia del usuario. El desarrollo se ha realizado utilizando la tecnología Vue.js junto con un framework moderno, lo que ha facilitado la creación de una interfaz interactiva y escalable, permitiendo gestionar de forma ágil tanto la visualización de resultados como la interacción con los datos. Esta combinación tecnológica ha resultado clave para integrar el algoritmo de recomendación en una aplicación web robusta y fácil de usar, mostrando el potencial de las herramientas actuales para resolver problemas reales de personalización y descubrimiento de información





Puedes ver el repositorio del proyecto aquí: https://github.com/Larzt/recomendation-system