



**Escuela Superior
de Ingeniería y Tecnología**
Universidad de La Laguna

Sistema Recomendación en filtrado colaborativo

Distancia Euclídea

Diego Hernández Toledo

JOSÉ SUÁREZ FELIPE

C/ Padre Herrera s/n
38207 La Laguna
Santa Cruz de Tenerife. España

T: 900 43 25 26

ull.es



Índice

Introducción.....	2
Descripción del sistema desarrollado.....	3
Ejemplos y Análisis.....	4
Ejemplo 1.....	4
Ejemplo 2.....	5
Ejemplo 3.....	6
Ejemplo 4.....	7
Ejemplo 5.....	8
Conclusiones.....	9
Referencias.....	10



Introducción

Conocemos a los sistemas de recomendación como herramientas muy importantes en plataformas digitales, ya que nos predicen las preferencias de los usuarios y ayuda a ofrecerles sugerencias personalizadas.

En esta práctica implementamos un sistema de recomendación basado en filtrado colaborativo, este utiliza las valoraciones de otros usuarios con gustos similares para poder calcular puntuaciones desconocidas. Para medir la similitud entre usuarios nosotros hemos escogido la distancia euclídea.

Por facilidad y conocimiento hemos escogido HTML, CSS y JavaScript, con lo que cargamos una matriz, posteriormente calculamos las similitudes y así predecimos los valores faltantes y generamos las recomendaciones.



Descripción del sistema desarrollado

El sistema desarrollado consiste en una plataforma web que establece el método de filtrado colaborativo que mediante la distancia euclídea es capaz de generar recomendaciones personalizadas.

La aplicación pide en primer lugar cargar un fichero del ordenador que tenga de contenido una matriz de utilidad, donde las filas representan usuarios y las columnas ítems. Los números representan valores que conocemos mientras que los guiones representan valoraciones desconocidas los cuales posteriormente debemos calcular.

El sistema utiliza la distancia euclídea para medir la similitud entre usuarios y seleccionar a los vecinos más cercanos. A partir de ellos, se calculan las predicciones para los valores faltantes mediante un método de promedio ponderado según la similitud.

Finalmente, mostramos los resultados en tablas que incluyen la matriz original, las similitudes calculadas, los vecinos seleccionados, las predicciones generadas y las recomendaciones finales para cada usuario.

Sistema de Recomendación

Matriz:

Seleccionar archivo Ningún archivo seleccionado

Vecinos (k):

3

Tipo de predicción:

Simple

Calculo de recomendaciones

Resultados

Componentes:

- Diego Hernández Toledo
- Jose Suarez Felipe



Ejemplos y Análisis

Ejemplo 1

Matriz elegida: utility-matrix-10-25-1

Características: Datos muy dispersos, muchas ausencias, pero con bastantes coincidencias.

Vecinos seleccionados:

Usuario	Vecinos
U1	U7 (0.14), U6 (0.11)
U2	U7 (0.12), U1 (0.11)
U3	U9 (0.11), U2 (0.10)
U4	U8 (0.11), U1 (0.10)
U5	U7 (0.11), U2 (0.11)
U6	U1 (0.11), U5 (0.11)
U7	U1 (0.14), U2 (0.12)
U8	U2 (0.11), U4 (0.11)
U9	U3 (0.11), U2 (0.10)
U10	U6 (0.10), U3 (0.10)

Recomendación:

Recomendaciones

- Usuario 1 → Ítem 19 (predicción: 3.18)
- Usuario 2 → Ítem 5 (predicción: 2.90)
- Usuario 3 → Ítem 23 (predicción: 4.32)
- Usuario 4 → Ítem 9 (predicción: 2.68)
- Usuario 5 → Ítem 23 (predicción: 3.34)
- Usuario 6 → Ítem 16 (predicción: 2.96)
- Usuario 7 → Ítem 9 (predicción: 3.30)
- Usuario 8 → Ítem 23 (predicción: 2.74)

Conclusión

El sistema logra identificar similitudes entre usuarios pese a la dispersión de datos, lo que permite generar recomendaciones razonables. Sin embargo, la gran cantidad de valores ausentes limita la precisión de las predicciones. Se observa que el modelo sigue siendo útil incluso con matrices poco densas.



Ejemplo 2

Matriz elegida: utility-matrix-10-25-2

Características: Matriz más completa

Vecinos seleccionados:

Usuario	Vecinos
U1	U6 (0.12), U2 (0.11)
U2	U6 (0.12), U7 (0.12)
U3	U6 (0.14), U5 (0.11)
U4	U2 (0.12), U8 (0.11)
U5	U6 (0.13), U3 (0.11)
U6	U7 (0.14), U3 (0.14)
U7	U6 (0.14), U2 (0.12)
U8	U4 (0.11), U6 (0.10)
U9	U3 (0.11), U2 (0.11)
U10	U3 (0.10), U6 (0.10)

Recomendación:

Recomendaciones

- Usuario 1 → Ítem 15 (predicción: 4.42)
- Usuario 3 → Ítem 19 (predicción: 2.27)
- Usuario 4 → Ítem 11 (predicción: 3.42)
- Usuario 5 → Ítem 2 (predicción: 3.88)
- Usuario 6 → Ítem 14 (predicción: 2.64)
- Usuario 9 → Ítem 6 (predicción: 3.25)
- Usuario 10 → Ítem 2 (predicción: 3.79)

Conclusión:

Al tratarse de una matriz más completa, las distancias euclídeas reflejan mejor las preferencias reales. Las predicciones son más coherentes y los vecinos más consistentes. Esto demuestra que una mayor densidad de datos mejora notablemente la calidad de las recomendaciones.



Ejemplo 3

Matriz elegida: utility-matrix-10-25-3

Características: Menor densidad, muchos valores faltantes, pero diferencias marcadas en algunos ítems.

Vecinos seleccionados:

Usuario	Vecinos
U1	U3 (0.12), U4 (0.12)
U2	U9 (0.15), U10 (0.12)
U3	U1 (0.12), U10 (0.11)
U4	U7 (0.12), U1 (0.12)
U5	U7 (0.12), U4 (0.11)
U6	U10 (0.12), U7 (0.11)
U7	U4 (0.12), U5 (0.12)
U8	U1 (0.11), U2 (0.11)
U9	U2 (0.15), U10 (0.11)
U10	U2 (0.12), U6 (0.12)

Recomendación:

Recomendaciones

- Usuario 1 → Ítem 4 (predicción: 2.88)
- Usuario 2 → Ítem 2 (predicción: 2.51)
- Usuario 3 → Ítem 24 (predicción: 2.10)
- Usuario 4 → Ítem 20 (predicción: 3.41)
- Usuario 5 → Ítem 20 (predicción: 3.73)
- Usuario 6 → Ítem 22 (predicción: 1.88)
- Usuario 7 → Ítem 3 (predicción: 4.71)
- Usuario 8 → Ítem 22 (predicción: 2.68)
- Usuario 9 → Ítem 7 (predicción: 2.95)
- Usuario 10 → Ítem 5 (predicción: 3.28)

Conclusión:

La baja densidad y las diferencias marcadas entre usuarios generan resultados más variables. Aun así, el algoritmo logra detectar patrones de afinidad útiles. Se evidencia que la distancia euclídea puede adaptarse incluso en escenarios con escasa información, aunque con menor fiabilidad.



Ejemplo 4

Matriz elegida: utility-matrix-10-25-4

Características: Matriz bastante completa, con valores moderados

Vecinos seleccionados:

Usuario	Vecinos
U1	U2 (0.13), U3 (0.12)
U2	U1 (0.13), U4 (0.12)
U3	U10 (0.15), U7 (0.15)
U4	U6 (0.13), U10 (0.13)
U5	U3 (0.11), U8 (0.11)
U6	U4 (0.13), U10 (0.12)
U7	U3 (0.15), U1 (0.12)
U8	U2 (0.12), U4 (0.12)
U9	U2 (0.12), U1 (0.12)
U10	U3 (0.15), U4 (0.13)

Recomendación:

Recomendaciones

- Usuario 2 → Ítem 16 (predicción: 3.22)
- Usuario 3 → Ítem 10 (predicción: 3.18)
- Usuario 6 → Ítem 11 (predicción: 3.75)
- Usuario 8 → Ítem 11 (predicción: 3.86)
- Usuario 9 → Ítem 10 (predicción: 3.27)
- Usuario 10 → Ítem 3 (predicción: 3.35)

Conclusión:

Al tener una matriz mucho más completa podemos observar como los resultados son más fiables, con los valores a completar son más precisos debido a la cantidad de datos que tiene para calcular y eso nos demuestra que el prototipo también funciona cuando no hay muchos campos a completar.



Ejemplo 5

Matriz elegida: utility-matrix-10-25-5

Características: Bastante completo, valores medios-altos

Vecinos seleccionados:

Usuario	Vecinos
U1	U6 (0.13), U2 (0.12)
U2	U9 (0.13), U4 (0.13)
U3	U1 (0.11), U6 (0.11)
U4	U2 (0.13), U9 (0.12)
U5	U2 (0.12), U1 (0.11)
U6	U1 (0.13), U10 (0.12)
U7	U6 (0.11), U2 (0.10)
U8	U9 (0.13), U1 (0.10)
U9	U2 (0.13), U8 (0.13)
U10	U6 (0.12), U1 (0.11)

Recomendación:

Recomendaciones

- Usuario 1 → Ítem 21 (predicción: 3.77)
- Usuario 2 → Ítem 24 (predicción: 4.65)
- Usuario 3 → Ítem 18 (predicción: 2.76)
- Usuario 4 → Ítem 21 (predicción: 3.34)
- Usuario 5 → Ítem 16 (predicción: 2.36)
- Usuario 6 → Ítem 18 (predicción: 1.50)
- Usuario 7 → Ítem 6 (predicción: 1.88)
- Usuario 9 → Ítem 13 (predicción: 3.37)
- Usuario 10 → Ítem 14 (predicción: 2.21)

Conclusión:

Al tratarse de una matriz medianamente completa, las distancias reflejan las preferencias realmente. Podemos observar unas predicciones más coherentes y unos vecinos con más presencia. Esto demuestra que cuanto más datos nos aporta mejor son los resultados de las recomendaciones.



Conclusiones

Hemos realizado los estudios de cada uno de los 5 ejemplos usando como valor de k de números de vecinos 2 y estos han sido los resultados existen otras matrices como la de similitudes, la inicial y la final pero son resultados demasiado extensos que generaron que el documento no muestre lo importante.

Con lo que hemos mostrado podemos observar como los resultados han sido bastante parejos ya que los valores han sido con el mismo número de usuarios y valores en intervalo de 0 a 5.



Referencias

Enlace a repo

https://github.com/DiegoHdezToledo/GCO_Sistema_Recomendacion.git

Enlace repo matrices

<https://github.com/ull-cs/gestion-conocimiento/tree/main/recommeder-systems/examples-utility-matrices>

Enlace a diapositivas

https://docs.google.com/presentation/d/1G9dtDODvXBQpYzOCBI_7VCBKcOD4SV_ZZZrVUBLxGHnE/edit