

# INFORME PRÁCTICA MÉTODO DE FILTRADO COLABORATIVO

| Sistema de recomendación                                      |  |  |  |  |  |  |  |  |  |  |  |
|---|--|--|--|--|--|--|--|--|--|--|--|
| Sube la matriz de utilidad: Seleccionar archivo problemas.txt |  |  |  |  |  |  |  |  |  |  |  |
| Selecciona un fichero .txt con la matriz de utilidad.         |  |  |  |  |  |  |  |  |  |  |  |
| Métrica de similitud: Correlación de Pearson ∨                |  |  |  |  |  |  |  |  |  |  |  |
| Número de vecinos:2   |  |  |  |  |  |  |  |  |  |  |  |
| Tipo de predicción: Diferencia con la media v                 |  |  |  |  |  |  |  |  |  |  |  |
| Calcular predicciones   |  |  |  |  |  |  |  |  |  |  |  |
| Resultados de la predicción                                   |  |  |  |  |  |  |  |  |  |  |  |
| Usuario Ítem 1 Ítem 2 Ítem 3 Ítem 4 Ítem 5                    |  |  |  |  |  |  |  |  |  |  |  |
| U1 5.00 3.00 4.00 4.00 4.51                                   |  |  |  |  |  |  |  |  |  |  |  |
| U2 3.00 1.00 2.00 3.00 3.00                                   |  |  |  |  |  |  |  |  |  |  |  |
| U3 4.00 3.00 4.00 3.00 4.00                                   |  |  |  |  |  |  |  |  |  |  |  |
| U4 3.00 3.00 1.00 5.00 4.00                                   |  |  |  |  |  |  |  |  |  |  |  |
| U5 1.00 4.00 4.00 2.00 1.00                                   |  |  |  |  |  |  |  |  |  |  |  |

## Gestión del Conocimiento para las Organizaciones

Alberto Antonio Hernández Hernández Eduardo Socas Luis PE101



## ÍNDICE

| Matrices 10x25    | 2  |
|-------------------|----|
| Matrices 100x1000 | 5  |
| Matrices 25x100   | 8  |
| Matrices 5x10     | וו |
| Matrices 50x250   | 14 |



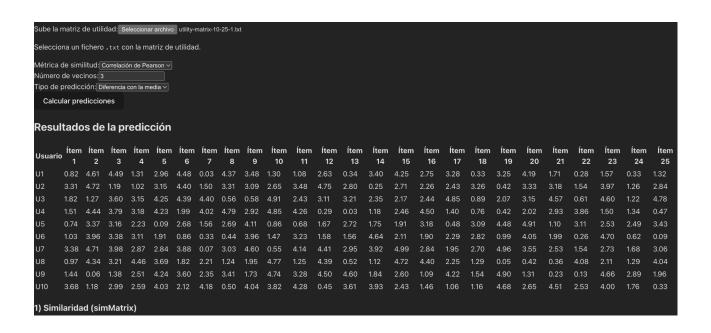
#### Matrices 10x25

Para este primer tipo de matrices, hemos hecho uso del ejemplo del siguiente fichero del repositorio de pruebas (<u>utility-matrix-10-25-1.txt</u>) con los siguientes parámetros:

MÉTRICA: CORRELACIÓN DE PEARSON

**VECINDAD**: 3

TIPO DE PREDICCIÓN: DIFERENCIA CON LA MEDIA





#### **VECINOS SELECCIONADOS**

```
2) Vecinos seleccionados (top-k por usuario)
Primeros 3 usuarios:
     "neighborIndex": 6,
     "similarity": 0.607983113735571
     "neighborIndex": 5,
     "similarity": 0.36910551993159035
     "neighborIndex": 4,
     "similarity": 0.2225263277311905
     "neighborIndex": 0,
     "similarity": 0.21346867965251612
     "neighborIndex": 6,
     "similarity": 0.13169025756813987
     "neighborIndex": 7,
      "similarity": 0.0929072894220118
     "neighborIndex": 8,
     "similarity": 0.33066718855005045
     "neighborIndex": 9,
     "similarity": 0.1440699715934266
    },
     "neighborIndex": 7,
      "similarity": 0.054459685168500936
```



#### **ALGUNAS RECOMENDACIONES**

```
4) Recomendaciones por usuario (top)
Primeros 3 usuarios:
   "user": 0,
   "recommendations": [
       "item": 18,
       "predicted": 3.253447847269392
       "item": 11,
       "predicted": 2.631102704545359
     },
       "item": 20,
       "predicted": 1.7122170136370407
   "user": 1,
    "recommendations": [
       "item": 4,
       "predicted": 3.1515691150789125
       "item": 21,
       "predicted": 1.5357848390650939
   "user": 2,
    "recommendations": [
       "item": 22,
       "predicted": 4.602351829783861
```



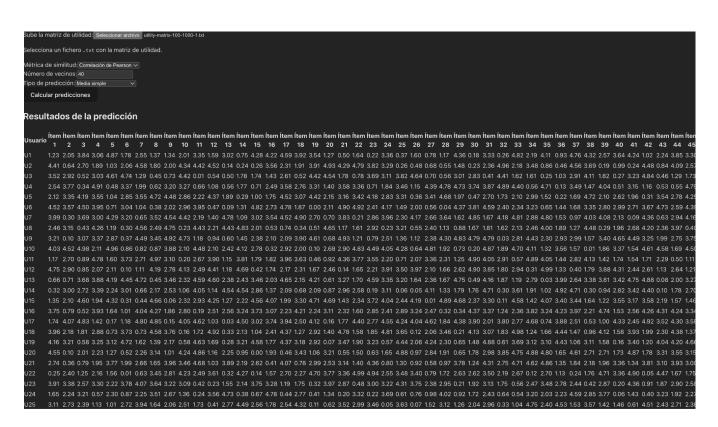
#### **Matrices 100x1000**

En este segundo tipo de matrices hemos usado el siguiente fichero (<u>utility-matrix-100-1000-1.txt</u>) y dado que la matriz es excesivamente grande, no se ha podido plasmar entera en la captura. Sin embargo, en la aplicación web se puede deslizar la página para verla por completo. Hemos usado estos parámetros:

MÉTRICA: CORRELACIÓN DE PEARSON

VECINDAD: 40

TIPO DE PREDICCIÓN: MEDIA SIMPLE





#### **ALGUNOS DETALLES DE LAS PREDICCIONES**

```
3) Detalle del cálculo de predicciones (muestras)
Cada entrada incluye vecinos usados, fórmula, predicción cruda y final.
   "user": 0,
   "item": 447,
   "neighborsUsed": [
       "neighborIndex": 39,
       "similarity": 0.08282019435505752,
       "rating": 3.158
        "neighborIndex": 80,
       "similarity": 0.07125731840791229,
       "rating": 2.689
       "neighborIndex": 5,
       "similarity": 0.06598596225326163,
       "rating": 4.601
       "neighborIndex": 33,
       "similarity": 0.05924457605154232,
       "rating": 4.87
       "neighborIndex": 68,
       "similarity": 0.05010437047751404,
       "rating": 1.555
        "neighborIndex": 13,
       "similarity": 0.04926476430024568,
       "rating": 1.122
       "neighborIndex": 86,
       "similarity": 0.04858604730620318,
       "rating": 0.574
        "neighborIndex": 54,
       "similarity": 0.04654103591788334,
        "rating": 0.958
```



#### **RECORTE MATRIZ DE SIMILARIDAD**

```
1) Similaridad (simMatrix)
Matriz usuario×usuario. Mostrando un recorte 5×5:
   -0.00028124840487621193,
   -0.03041503929035448,
   -0.03675357969311844,
   0.012861700079614423
 ],
   -0.00028124840487621193,
   0.005573191137041729,
   0.05750807521261387,
   -0.04758053270582271
 ],
   -0.03041503929035448,
   0.005573191137041729,
   0.0315107484116241,
   -0.03870512389595767
 ],
   -0.03675357969311844,
   0.05750807521261387,
   0.0315107484116241,
   -0.04600357371830146
 ],
   0.012861700079614423,
   -0.04758053270582271,
   -0.03870512389595767,
   -0.04600357371830146,
   0
```



#### Matrices 25x100

Para esta prueba se usó el fichero <u>utility-matrix-25-100-2.txt</u> del repositorio de pruebas. Nuevamente no se puede capturar la imagen de toda la matriz por sus dimensiones, pero en la aplicación web se puede deslizar la página para ver todos los valores.

**MÉTRICA**: COSENO

**VECINDAD**: 8

TIPO DE PREDICCIÓN: MEDIA SIMPLE

#### MATRIZ DE UTILIDAD COMPLETADA

Resultados de la predicción 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 3.30 2.56 2.51 2.78 1.07 2.03 0.09 1.31 4.91 1.94 0.06 0.63 4.15 3.74 2.08 4.84 4.46 0.34 3.36 2.92 0.76 2.32 1.26 2.31 3.20 1.72 0.80 4.21 2.42 0.39 1.69 2.81 1.26 3.58 1.13 4.48 1.69 0.72 4.26 0.08 2.27 0.68 0.69 4.58 3.66 1.81 1.28 2.25 0.59 0.33 1.32 4.11 2.46 4.97 0.69 3.96 2.13 3.65 1.91 2.52 1.17 1.22 2.32 1.35 4.53 2.08 1.29 0.68 2.96 2.48 417 457 0.37 4.56 1.49 3.28 4.88 2.83 3.75 2.57 4.06 4.42 0.41 1.84 2.06 4.77 1.16 4.97 1.23 1.92 4.61 2.85 2.27 4.05 1.28 0.15 3.10 3.75 2.03 2.52 2.38 1.15 4.52 2.96 3.26 0.11 0.57 1.24 0.39 2.56 0.54 2.31 1.00 0.66 4.80 1.79 3.13 1.78 2.16 1.99 2.37 2.73 0.65 3.19 1.20 2.69 2.84 1.41 1.61 2.81 1.07 3.16 1.14 4.72 1.54 1.44 0.89 2.83 0.35 1.61 3.86 2.65 0.07 2.89 3.28 4.23 3.45 1.73 0.45 3.23 3.18 0.99 4.42 4.01 0.23 2.54 4.50 3.79 2.99 1.93 3.14 2.96 2.86 2.14 0.44 4.31 1.68 0.82 3.90 3.38 4.98 1.28 4.57 0.38 3.44 3.97 3.65 4.40 3.65 2.43 3.11 3.56 0.75 2.79 3.99 0.17 2.14 3.65 4.36 0.16 0.04 3.58 2.25 1.24 0.23 2.89 3.52 4.33 1.43 1.37 0.50 1.91 4.21 1.83 2.71 2.27 2.17 1.18 1.15 2.32 4.06 2.18 2.55 1.11 0.19 0.62 1.54 1.48 1.15 2.97 3.81 2.43 4.99 1.29 3.08 3.54 4.52 4.81 4.51 3.53 2.62 2.68 1.08 4.89 3.34 2.34 3.46 3.86 1.30 2.20 4.29 3.91 0.62 4.69 3.76 4.65 3.50 3.23 4.79 1.89 2.82 0.71 3.53 3.55 4.69 2.19 2.61 0.63 2.04 0.24 4.39 0.64 3.12 3.86 1.76 2.84 3.41 1.96 4.41 3.58 3.63 0.36 2.46 1.36 4.97 4.90 0.80 4.96 3.61 2.27 2.67 0.85 3.38 0.21 1.90 1.56 0.12 1.03 2.22 4.45 1.94 2.15 1.03 0.31 1.22 3.14 4.44 2.82 1.70 2.23 1.56 2.44 1.89 3.95 2.23 1.89 3.72 2.77 2.07 2.85 0.40 0.60 0.62 4.74 3.95 3.62 2.85 2.25 3.02 1.14 2.85 0.40 4.54 4.06 3.34 2.10 2.99 2.34 0.84 1.55 2.42 4.25 3.17 2.30 0.87 3.13 3.94 3.71 0.06 4.68 0.17 3.48 1.82 2.24 0.15 3.52 2.46 3.23 2.69 3.82 3.30 4.37 0.10 0.93 1.18 2.44 1.44 0.15 4.66 2.40 3.92 2.18 3.00 0.76 0.64 0.08 2.83 3.84 0.42 4.89 4.46 2.12 1.88 4.51 1.60 1.59 4.44 0.44 4.79 1.02 4.94 2.45 2.23 0.38 3.10 0.89 0.82 3.47 1.39 1.02 2.98 1.41 2.65 3.76 0.30 2.40 3.71 4.25 3.30 3.30 3.22 4.34 4.56 0.19 0.77 2.60 4.33 1.04 3.70 1.80 0.87 4.46 3.84 0.21 1.45 0.94 4.06 2.29 3.94 0.24 0.48 1.32 1.69 0.71 3.38 0.99 1.85 0.64 3.85 4.70 3.85 3.21 0.83 1.93 1.69 1.84 0.63 3.92 4.96 1.57 4.15 0.54 4.41 2.56 1.91 2.97 0.18 1.59 2.38 4.65 0.84 0.09 1.83 1.45 1.59 0.34 2.96 0.21 3.77 2.84 4.24 4.26 1.38 3.86 3.91 1.96 4.47 2.74 4.02 1.16 3.24 1.24 1.61 2.23 4.66 1.85 1.54 4.66 4.60 2.13 2.39 3.21 2.28 3.77 1.80 0.04 4.14 4.94 3.54 4.81 3.59 3.40 0.86 3.23 2.47 1.03 3.97 2.72 0.48 3.68 4.93 0.20 2.35 0.67 1.13 2.63 4.46 4.79 0.92 1.04 0.78 2.80 1.90 4.88 4.95 3.75 2.20 3.68 2.65 4.41 1.75 3.69 3.93 2.96 2.55 2.75 1.36 0.87 3.31 3.06 2.25 4.82 2.87 4.47 4.23 0.30 4.03 2.13 4.43 0.97 1.56 4.32 2.22 3.88 3.23 4.74 0.43 2.06 1.33 4.78 2.42 0.82 4.73 0.95 1.83 4.39 2.18 3.22 1.23 1.89 2.75 0.95 4.91 3.65 0.26 3.73 0.52 1.25 1.64 3.99 2.42 4.53 0.90 1.64 2.49 3.52 0.04 2.57 4.96 1.89 4.24 3.15 4.31 4.75 2.50 3.77 4.49 0.63 3.49 4.15 4.10 2.67 0.89 2.31 0.16 3.40 3.77 3.59 3.97 1.22 0.52 4.97 2.74 0.48 4.89 1.97 3.40 4.48 4.67 2.75 1.35 2.83 2.28 1.38 4.00 0.67 1.36 1.56 1.46 4.31 3.91 3.83 3.69 0.77 2.92 1.53 4.43 3.37 4.73 0.23 3.69 0.00 1.06 1.67 0.27 0.44 2.18 3.29 1.27 4.80 3.05 4.71 2.45 1.46 2.52 1.47 4.75 1.85 0.68 3.71 3.81 3.44 3.31 2.16 2.55 1.60 0.32 3.06 2.48 3.12 4.25 3.63 4.71 1.30



#### **ALGUNOS VECINOS SELECCIONADOS**

```
2) Vecinos seleccionados (top-k por usuario)
Primeros 3 usuarios:
     "neighborIndex": 9,
     "similarity": 0.8188216637345042
    },
      "neighborIndex": 10,
     "similarity": 0.7821738400372578
     "neighborIndex": 16,
      "similarity": 0.7693013753331867
      "neighborIndex": 17,
     "similarity": 0.7678742741444269
      "neighborIndex": 15,
     "similarity": 0.7611312100502314
      "neighborIndex": 21,
     "similarity": 0.7598957056456148
      "neighborIndex": 19,
     "similarity": 0.7565767219730521
    },
      "neighborIndex": 1,
      "similarity": 0.7547051011432021
    }
  ],
```



#### **ALGUNAS RECOMENDACIONES**

```
4) Recomendaciones por usuario (top)
Primeros 3 usuarios:
   "user": 0,
   "recommendations": [
       "item": 14,
       "predicted": 2.0784059284336416
  },
   "user": 1,
   "recommendations": [
       "item": 91,
       "predicted": 3.1858400223439416
     },
       "item": 59,
       "predicted": 2.9307461175661103
  },
   "user": 2,
   "recommendations": [
       "item": 60,
       "predicted": 2.8413673838061277
       "item": 7,
       "predicted": 2.826813832643643
     },
```



## **Matrices 5x10**

Para este tipo de matrices elegimos el ejemplo del fichero <u>utility-matrix-5-10-6.txt</u> con los siguientes parámetros:

**MÉTRICA**: COSENO

**VECINDAD**: 5

TIPO DE PREDICCIÓN: DIFERENCIA CON LA MEDIA

| Sistema de recomendación  |        |        |        |        |        |        |        |        |        |         |  |
|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|--|
| Sube la matriz de utilidad: Seleccionar archivo utility-matrix-5-10-6.txt                         |        |        |        |        |        |        |        |        |        |         |  |
| Selecciona un fichero .txt con la matriz de utilidad.   |        |        |        |        |        |        |        |        |        |         |  |
| Métrica de similitud: Coseno  Número de vecinos: 5  Tipo de predicción: Diferencia con la media v |        |        |        |        |        |        |        |        |        |         |  |
| Calcular predicciones   |        |        |        |        |        |        |        |        |        |         |  |
| Resultados de la predicción   |        |        |        |        |        |        |        |        |        |         |  |
| Usuario   | Ítem 1 | Ítem 2 | Ítem 3 | Ítem 4 | Ítem 5 | Ítem 6 | Ítem 7 | Ítem 8 | Ítem 9 | Ítem 10 |  |
| U1  | 0.38   | 4.88   | 3.86   | 3.18   | 4.92   | 3.12   | 0.13   | 2.17   | 2.64   | 4.75    |  |
| U2  | 2.22   | 3.35   | 2.11   | 0.53   | 2.06   | 1.78   | 2.13   | 4.84   | 4.24   | 3.24    |  |
| U3  | 3.90   | 3.60   | 1.61   | 0.35   | 0.21   | 3.01   | 4.33   | 3.49   | 1.81   | 2.16    |  |
| U4  | 4.82   | 2.52   | 0.15   | 1.08   | 2.29   | 4.54   | 2.35   | 3.18   | 2.66   | 0.78    |  |
| U5  | 0.60   | 1.30   | 1.91   | 0.19   | 2.90   | 1.76   | 0.67   | 1.13   | 1.72   | 0.30    |  |



#### MATRIZ DE SIMILARIDAD

```
1) Similaridad (simMatrix)
Matriz usuario×usuario. Mostrando un recorte 5×5:
   0.7692181340718326,
   0.5427343989249985,
   0.5874865083317209,
   0.8653885951064811
 ],
   0.7692181340718326,
   0.8320249652760822,
   0.7365698143681068,
   0.7576626473760124
  ],
   0.5427343989249985,
   0.8320249652760822,
   0.9439157589935002,
   0.6095871174394892
 ],
   0.5874865083317209,
   0.7365698143681068,
   0.9439157589935002,
   0.6092294491981591
   0.8653885951064811,
   0.7576626473760124,
   0.6095871174394892,
   0.6092294491981591,
```



### ALGUNOS DETALLES DE LA PREDICCIÓN

# 3) Detalle del cálculo de predicciones (muestras) Cada entrada incluye vecinos usados, fórmula, predicción cruda y final. "user": 1, "item": 1, "neighborsUsed": [ "neighborIndex": 2, "similarity": 0.8320249652760822, "rating": 3.6, "neighborMean": 2.41325 "neighborIndex": 0, "similarity": 0.7692181340718326, "rating": 4.875, "neighborMean": 3.0023 }, "neighborIndex": 4, "similarity": 0.7576626473760124, "rating": 1.301, "neighborMean": 1.4084285714285714 }, "neighborIndex": 3, "similarity": 0.7365698143681068, "rating": 2.52, "neighborMean": 2.438 "rawPrediction": 3.350784341082685, "finalPrediction": 3.350784341082685, "formula": "mean-diff"



#### Matrices 50x250

Para este tipo de matriz, hemos usado el ejemplo del fichero <u>utility-matrix-50-250-8.txt</u> del repositorio con los siguientes parámetros:

MÉTRICA: DISTANCIA EUCLÍDEA

**VECINDAD**: 20

TIPO DE PREDICCIÓN: MEDIA SIMPLE





#### **ALGUNOS VECINOS SELECCIONADOS**

```
2) Vecinos seleccionados (top-k por usuario)
Primeros 3 usuarios:
     "neighborIndex": 31,
     "similarity": 0.033443934113420226
   },
     "neighborIndex": 12,
     "similarity": 0.03293820270124858
     "neighborIndex": 47,
      "similarity": 0.03291062208880887
     "neighborIndex": 18,
     "similarity": 0.03269373559700928
     "neighborIndex": 28,
     "similarity": 0.032542491105275115
    },
     "neighborIndex": 26,
     "similarity": 0.03237141668829083
    },
     "neighborIndex": 42,
     "similarity": 0.03200862331787907
    },
```



#### **ALGUNAS RECOMENDACIONES**

```
"user": 1,
  "recommendations": [
      "item": 124,
     "predicted": 3.0696137927487688
    },
    {
     "item": 172,
     "predicted": 2.990874391374004
    },
    {
      "item": 118,
      "predicted": 2.5853299567305505
    },
      "item": 203,
      "predicted": 2.249280201423909
 ]
},
  "user": 2,
  "recommendations": [
      "item": 6,
      "predicted": 3.014017513162909
    },
      "item": 88,
     "predicted": 2.5677980684404482
    },
      "item": 132,
```