

<u>Gestión del Conocimiento en</u> <u>las Organizaciones</u>

Sistemas de recomendación. Métodos de filtrado colaborativo

Alberto Oria Martín (alu0101415598@ull.edu.es)

Izan Pérez González (alu0101543345@ull.edu.es)

Airam Prieto González

(alu0101546377@ull.edu.es)



Matrices empleadas	3
Comportamiento general	3
Matrices pequeñas (5×10 y 10×25)	3
Matrices medianas y grandes (25×100, 50×250, 100×1000)	3
Ejemplos de ejecución	4
Matriz 5x10	5
Pearson	5
Distancia coseno	6
Distancia euclídea	7
Matriz 10x25	8
Pearson	8
Distancia coseno	9
Distancia euclídea	10
Comparativa entre métodos de predicción	11
Resultados y tendencias generales	12
Tabla descriptiva	12
Conclusión final	13



Matrices empleadas

- <u>utility-matrix-5-10-1.txt</u>
- <u>utility-matrix-10-25-1.txt</u>
- <u>utility-matrix-25-100-1.txt</u>
- utility-matrix-50-250-1.txt
- utility-matrix-100-1000-1.txt

Comportamiento general

Matrices pequeñas (5×10 y 10×25)

En estas matrices con pocos ítems por usuario, el **coeficiente de correlación de Pearson** mostró resultados inestables: similitudes con signo negativo y
predicciones fuera de rango si no se filtran las correlaciones negativas. Esto
ocurre porque **la correlación requiere un número suficiente de ítems comunes**;
con pocos valores compartidos, la varianza es poco representativa y el resultado
se vuelve ruidoso.

La **distancia coseno** produjo similitudes positivas y más coherentes con los valores originales. Al basarse en la orientación de los vectores, ignora las diferencias de escala entre usuarios, lo que genera predicciones más estables ante la falta de datos o valores atípicos.

La **distancia euclídea**, al transformarse en similitud inversa, arrojó similitudes bajas y homogéneas. Sus predicciones fueron conservadoras y amortiguadas, evitando extremos y reduciendo el efecto de ratings aislados, aunque con menor capacidad de discriminación.

Matrices medianas y grandes (25×100, 50×250, 100×1000)

A medida que la matriz crece, la **distancia coseno** se mantiene estable y ofrece similitudes coherentes entre usuarios con comportamientos parecidos. El **coeficiente de Pearson** mejora su fiabilidad cuando hay densidad moderada (por ejemplo, en 25×100), pero se degrada con el aumento de sparsity en 50×250 y 100×1000, mostrando correlaciones cercanas a cero debido a la falta de coincidencias suficientes.

La distancia euclídea, aunque estable, sufre el fenómeno de concentración de



distancias en espacios de alta dimensión, lo que reduce su capacidad de distinguir entre usuarios realmente distintos.

Nota: Las matrices 25×100, 50×250 y 100×1000 no se incluyen visualmente por limitaciones de espacio, pero sus resultados siguieron las mismas tendencias observadas: mejora progresiva de Pearson con densidad moderada y estabilidad de Coseno y Euclídea en entornos dispersos.

Ejemplos de ejecución

Las capturas mostradas a continuación muestran los resultados de cada métrica, la primera con la predicción simple y la segunda usando mean difference.

Matriz 5x10

Pearson

Matriz de uti	ilidad complet									
 					Ítem 4					
Usuario 0		2.65	1.65	-1.51	1.12	0.88	+======= 0.42	3.98 3.98	1.8	3.14
Usuario 1	3.41	0.31	3.8	 4.23	2.16	4.51	2.39	0.87	2.47	-0.44
Usuario 2	4.41	4.5	2.05	-1.45	0.05	0.42	3.35	3.74	4.08	2.86
Usuario 3	1.73	1.18	-0.19	1.51	4.87	2.22	 3	2.9	2.11	-1.32
Usuario 4	0.56	4.89	1.22	0.8	3.8	4.88	2.83	0.99	4.49	0.44
Tabla de simi	ilitudes entre	usuarios se	gún la r	nétrica pe				, ,		
	Usuario 0	Usuario	1 U:	suario 2	Usuario 3	Usuari	o 4			
Usuario 0	0	-0.4	+ 8 	0.56	-0.33	-0	.58			
Usuario 1	-0.48	i 0	i	-0.26	-0.55	-0	.24			
Usuario 2	0.56	-0.2	6		-0.97	-0	.05			
Usuario 3	-0.33	-0.5	5	-0.97		. 0	.34			
Usuario 4	-0.58	-0.2	+ 4	-0.05	0.34	[0	į			



	Ítem 0	Ítem 1	Ítem 2	Ítem 3	Ítem 4	Ítem 5	Ítem 6	Ítem 7	Ítem 8	Ítem 9
======================================	3.14	2.65	1.65	3.23	1.12	0.88	0.42	3.98	3.01	3.14
Usuario 1	3.41	0.31	3.8	4.23	2.16	4.51	2.39	0.87	2.47	4.74
Usuario 2	4.41	4.5	2.05	4.14	0.05	1.78	3.35	3.74	4.08	4.04
Usuario 3	1.73	3.68	2.44	1.51	4.87	2.22	3	2.9	2.11	1.27
Usuario 4	0.56	4.89	1.22	0.8	3.8	4.88	2.83	0.99	4.49	0.44
abla de simi 	litudes entre Usuario 0		+	étrica pea uario 2		+ Usuario				
	Usuario 0		+ 1 Us ==+====		Usuario 3					
	Usuario 0 ========0	+	+ 1 Us ==+==== 8 +	uario 2	Usuario 3 	+======== -0. +	===+ 58 +			
	Usuario 0 ======= 0 -0.48	+	+ 1 Us ==+==== 8 +	uario 2 ===================================	Usuario 3 	-0. -0.	58 + 24			
	Usuario 0 0 -0.48 0.56	+	+ 1 Us + 8 + 6 +	uario 2 =======+= 0.56 -0.26 0	Usuario 3 -0.33 -0.55	-0. -0. -0.	58 + 24			

- Predicciones a veces negativas.
- Alta dispersión entre usuarios.
- Inestable por escasez de ítems en común.

Distancia coseno

atriz de util	lidad completa	ada:							
!	 	Ítem 1 Ít	em 2 Ítem 3		 				Ítem 9
Usuario 0	3.14	2.65	1.65 1.51	1.12	0.88	0.42	3.98	3.23	3.14
Usuario 1	3.41	0.31	3.8 4.23	2.16	4.51	2.39	0.87	2.47	-
Usuario 2	4.41	4.5	2.05 1.1 9	0.05	2.71	3.35	3.74	4.08	1.9
Usuario 3	1.73	2.7	2.45 1.51	4.87	2.22	3	2.9	2.11	0.44
Usuario 4	0.56	4.89	1.22 0.8	3.8	4.88	2.83			0.44
abla de simil	litudes entre		n la métrica co						
	Usuario 0	Usuario 1	Usuario 2	Usuario 3	Usuario 4	i			
Usuario 0	0	0.62	0.9	0.69	0.54	• [
Usuario 1	0.62	0	0.75	0.78	0.69)			
Usuario 2	0.9	0.75	0	0.69	0.76	5 [
Usuario 3	0.69	+ 0.78	0.69	0	0.85	5]			
Usuario 4	0.54	0.69	0.76	0.85	0				



	Ítem 0	Ítem 1	 Ítem 2		Ítem 4				+- Ítem 8	Ítem 9
+: Usuario 0	======+== 3.14	2.65	+ 1.65	1.01	+======+= 1.12	0.88	0.42	3.98	2.3	3.14
 Usuario 1	3.41	0.31	3.8	4.23	2.16	4.51	2.39	0.87	2.47	2.68
 Usuario 2	4.41	4.5	2.05	1.85	+ 0.05	3.48	3.35	3.74	4.08	2.72
Usuario 3	1.73	2.73	2.5	1.51	+ 4.87	2.22	3	2.9	2.11	1.55
 Usuario 4	0.56	4.89	1.22	0.8	3.8	4.88	2.83	0.99	4.49	0.44
ota de Simi 	Usuario 0				Usuario 3					
Usuario 0		0	.62 +	0.9 +-	0.69	0.5	64			
				0.9 +- 0.75						
 Usuario 1 					0.78	0.6				
	0.62 0.9			 0.75 +-	0.78	0.6				

- Predicciones positivas y consistentes con el rango original.
- Suaviza el efecto de valores extremos.
- Mide la orientación de los ratings, no su magnitud.



Distancia euclídea

Matriz de uti +	lidad completa	ada:			+	+			4				-+
i i	Ítem 0	Ítem 1		m 2			Ítem 5	Ítem		Ítem 7		Ítem 9 	i
Usuario 0	3.14	2.65		.65	1.51	1.12	0.88	0.4	42	3.98	3.23		i
Usuario 1	3.41			.8		2.16	4.51	2.	39			-	j
Usuario 2	4.41	4.5		2.05		0.05		3.3		3.74		3.14	Ĭ
Usuario 3	1.73			.42							2.11	1.7	
Usuario 4	0.56	4.89	1	.22	0.8		4.88	2.	83	0.99	4.49	0.44	Ĭ
Tabla de simi	litudes entre	usuarios	s según	la r	nétrica euc	lidean:							
	Usuario 0					Usuario 3							
	0	I				0.2							
Usuario 1	0.16		0		0.16		0.	14					
Usuario 2	0.26		0.16		0 <u> </u>		0.	16					
Usuario 3	0.2		0.19		0.17	0	0.	23					
Usuario 4	0.13	+ 	0.14		0.16	0.23	0.23 0						
Matriz de uti	lidad completa	ada:											
+	 Ítem 0		Íte	 m 2	+ Ítem 3	++ Ítem 4	+ Ítem 5	Ítem		 Ítem 7	 Ítem 8	+ Ítem 9	
++ +======+ Usuario 0		Ítem 1	-=====:		+=======		+						
	3.14	Ítem 1 ===================================	1	====- .65 	+======== 1.01 + 4.23	1.12 2.16		0.4	+ 42	3.98	2.3	3.14 	
Usuario 0 ++	3.14	Ítem 1 2.65 0.31	1 3	.65 .8 	1.01 1.01 	1.12 1.12 	0.88 0.88 4.51 	2	===+ 42 +	3.98	2.3 2.47	3.14 2.68	
Usuario 0 ++ Usuario 1 ++	3.14	Îtem 1 2.65 0.31 4.5	1 3 2 2	.65 .8 	1.01 +	1.12 2.16 0.05 4.87	0.88 	2.3	+ 42 + 39 + 35 +	3.98 0.87 3.74 2.9	2.3 2.47 4.08	3.14 2.68 4.06	
Usuario 0 ++ Usuario 1 ++ Usuario 2 ++	3.14	Îtem 1 2.65 0.31	1 3 2 1 1	 .65 .8 	1.01 4.23 2.12 1.51	1.12 2.16 0.05 4.87	0.88 	233	+ 42 + 39 + 35 +	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 	3.41 3.41 4.41 1.73	1tem 1 2.65 0.31 4.5 4.19 4.89	1 3 2 1	.65 .8 .05 .05 	1.01 4.23 2.12 1.51	1.12 2.16 0.05 4.87 3.8	4.51 2.19	233	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 	3.14 3.41 4.41 1.73 0.56	1tem 1 2.65 0.31 4.5 4.19 4.89 usuarios	1 3 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	.65 .8 .05 .96	1.01 4.23 2.12 1.51 0.8	1.12 2.16 0.05 4.87 3.8	0.88 4.51 2.19 2.22 4.88	3.3	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 Usuario 1 Usuario 2 Usuario 3 Usuario 4 Tabla de simi	3.41 3.41 4.41 1.73 0.56 litudes entre	2.65 0.31 4.5 4.19 4.89 Usuarios	1 3 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	.05 .96 .22	1.01 4.23 2.12 1.51 0.8 métrica euc	1.12 2.16 0.05 4.87 3.8 3.8 Usuario 3	0.88 4.51 2.19 2.22 4.88	3.:	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 Usuario 1 Usuario 2 Usuario 3 Usuario 4 Tabla de simi	3.14 3.41 4.41 1.73 0.56 litudes entre	1 tem 1 2.65 0.31 4.5 4.89 usuarios	1 3 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	.05 .96 .22	1.01 4.23 2.12 1.51 0.8 métrica euc	1.12 2.16 0.05 4.87 3.8 3.8 Usuario 3	0.88 4.51 2.19 2.22 4.88 Usuario	33	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0	3.14 3.41 4.41 1.73 0.56 litudes entre Usuario 0	1 tem 1 2.65 0.31 4.5 4.89 4.89 Usuarios	1 3 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	6589696	1.01 4.23 2.12 1.51 0.8 métrica euc	1.12 2.16 0.05 4.87 3.8 3.8 Usuario 3	0.88 4.51 2.19 2.22 4.88 4.88	2.: 3.: 3.: 2.: 4	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 Usuario 1 Usuario 2 Usuario 3 Usuario 4 Tabla de simi	3.14 3.41 4.41 1.73 0.56 litudes entre Usuario 0 0.16 0.26	1 tem 1 2.65 0.31 4.5 4.89 4.89 Usuarios	1 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	 .65 .8 .05 .05 .22 .22	1.01 4.23 2.12 1.51 0.8 0.8 0.8	1.12 2.16 0.05 4.87 3.8 Usuario 3 0.2 0.19 0.17	0.88 4.51 2.19 2.22 4.88 Usuario	2.: 3.: 3.: 2.: 4	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	
Usuario 0 Usuario 1 Usuario 2 Usuario 3 Usuario 4 Tabla de simi Usuario 0 Usuario 1 Usuario 2	3.14 3.41 4.41 1.73 0.56 litudes entre Usuario 0 0.16 0.26	1 tem 1 2.65 0.31 4.5 4.89 4.89 Usuarios	1 3 3 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	==== .65 	1.01 4.23 2.12 1.51 0.8 0.8 0.8	1.12 2.16 0.05 4.87 3.8 3.8 0.10 0.17 0.17	0.88 4.51 2.19 2.22 4.88 Usuarice	2.i	===+ 42 + 39 + 35 + 	3.98 0.87 3.74 2.9	2.47 4.08 2.11	3.14 2.68 4.06	

- Predicciones más conservadoras.
- Similitudes bajas pero consistentes.
- Reduce el impacto de outliers individuales.

Análisis:

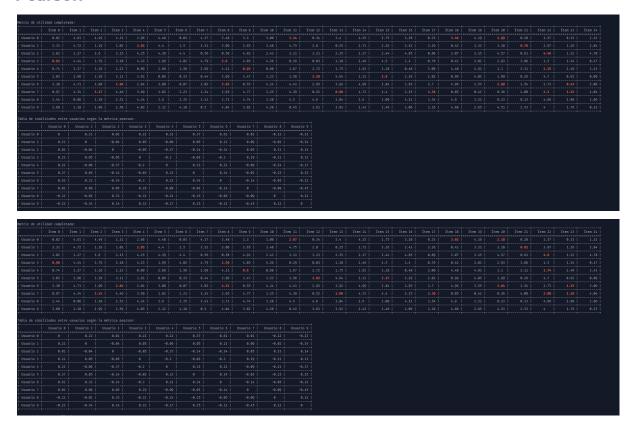
Pearson muestra su limitación en conjuntos pequeños; Coseno ofrece



predicciones razonables y Euclídea mantiene estabilidad pero poca sensibilidad entre usuarios.

Matriz 10x25

Pearson



- Muy sensible a valores atípicos o vecinos con datos faltantes.
- Puede dar valores indefinidos o erráticos (altos o negativos).
- Útil sólo cuando existe suficiente solapamiento entre usuarios.
- Funciona mejor con la fórmula de predicción basada en desviaciones respecto a la media.



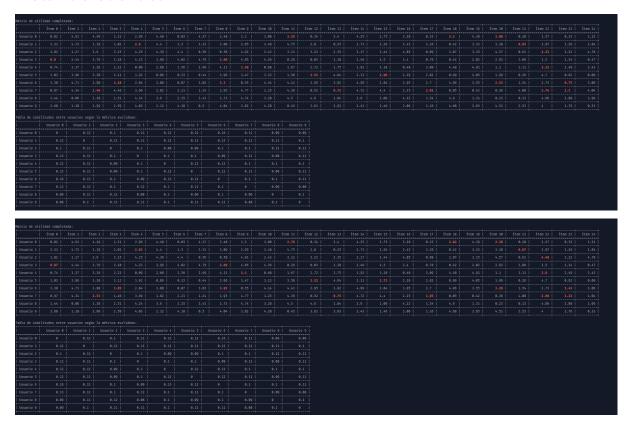
Distancia coseno

sario 0	0.82	4.61	4.49	1.31	2.95	4.48	0.03	4.37	3.48		1.08	4.57	0.34		4.25	2.75	3.28	0.33		4.19	3.18	0.28	1.57	0.33	1.32
uario 6																									3.00
																									1.96
					Usuario 3			20 5 Usi					9												
				0																					
				0.75				0.7																	
				0.71																					
uario 7 uario 8	0.73			0.77	0.81			0.09																	
wario 9	0.08		•	0.82	0.78	0.75	ļ	0.8	0.8	0.63	0.8														
suario 9 riz de utili	dad complete	ada:		0.82 iten 3	0.78 Iten 4		iten 0	0.8 item 7	0.8 Iten 8	0.63		iten 11 i	ften 12	Íten 13	îten 14	İten 15	Ítem 10	ften 17	Íten 18	ften 19	Íten 20	ften 21	ften 22	item 23	Íten 24
uario 9 iz de utili	dad complete	ada:	iten 2 1	İten 3	Íten 4	İten 5				Ítem 9	îten 10	iten 11					Îtem 10				iten 20	Íten 21 0.28	item 22	Îten 23 0.33	
uario 9 iz de utili uario 0	dad complete	ada: Íten 1					Íten 6	Íten 7	Íten 8				Íten 12	iten 13 3.4 0.25	ften 24 4.25 2.71	ften 15 2.75 2.26		Îten 17 0.33 3.20	item 18 2.25 0.42	Íten 19 4.19 3.23					ftem 24 1.32 2.84
uario 9 iz de utili	dad complete Item 0	ada: Íten 1 4.61	item 2 1	iten 3	Îten 4	Îten 3 4.48	Ítem 6 0.03	Íten 7	Íten 8 3.48	Ítem 9	Îten 10	Íten 11 3.95	ften 12 0.34												1.32
iz de utili uario 0 uario 1 uario 2	dad complets item 0 0.82 3.31	ada: item 1 4.01 4.72	item 2 1 4.49 1.19	iten 3 1.31 1.02	item 4 2.90 2.81	item 5 4.48 4.4	ften 6 0.03 1.5	item 7 4.37 3.31	iten 8 3.48 3.09	item 9 1.3 2.65	item 10 1.08 3.48	iten 11 3.95 4.75	İten 12 0.34 2.8												1.32 2.84
iz de utili usrio 0 usrio 1 usrio 2 usrio 3	dad complets Item 0 0.82 3.31 1.82	ada: ften 1 4.01 4.72 1.27	iten 2 1 4.49 1.19 3.6	iten 3 1.31 1.02	iten 4 2.90 2.81 4.25	iten 5 4.48 4.4 4.39	iten 6 0.03 1.5	iten 7 4.37 3.31 0.36	Îten 8 3.48 3.09 0.58	item 9 1.3 2.05 4.91	Iten 10 1.08 3.48 2.43	iten 11 3.95 4.75	Îten 12 0.34 2.8 3.21										1.57 3.97 4.46		1.32 2.84 4.78
uario 9 iz de utili uario 0 uario 1	dad complete iten 0 0.82 3.31 1.82 2.2	ada: ften 1 4.61 4.72 1.27 4.44	iten 2 1 4.49 1.29 3.0 3.79	iten 3 1.31 1.02 3.15	ften 4 2.90 2.81 4.25 4.23	iten 5 4.48 4.4 4.39 1.99	ften 6 0.03 1.5 4.4 4.02	iten 7 4.37 3.31 0.56 4.79	Item 8 3.48 3.49 0.58 0.58	ftem 9 1.3 2.05 4.91 4.85	Iten 10 1.65 3.45 2.43 4.20	iten 11 3.95 4.75 3.11 0.29	iten 12 0.34 2.8 3.21 0.03												1.32 2.84 4.78 0.47
uario 9 iz de utili uario 0 uario 1 uario 2 uario 3 uario 4 uario 5	dad complets ften 0 0.82 3.31 1.82 2.2 0.74	ada: ftem 1 4.01 4.72 1.27 4.44 3.37	iten 2 1 4.49 1.19 3.6 3.79 3.10	iten 3 1.31 1.02 3.15 3.18 2.23	iten 4 2.90 2.81 4.25 4.23 0.09	iten 5 4.48 4.4 4.39 1.99 2.08	iten 0 0.03 1.5 4.4 4.02 1.30	iten 7 4.37 3.31 9.30 4.79 2.69	Item 8 3.48 3.99 0.58 2.86 4.11	ftem 9 2.3 2.05 4.91 4.85 1.88	item 10 1.05 3.45 2.43 4.26 9.05	iten 11 3.95 4.75 3.11 0.29 1.07	iten 12 0.34 2.8 3.21 0.03 2.72												1.32 2.84 4.78 0.47 3.43
uario 9 iz de utili sario 0 uario 1 uario 2 uario 3 uario 4 uario 5 uario 6	ded complete Iten 0 0.82 3.31 1.82 2.2 0.74 1.03	ada: item 1 4.01 4.72 1.27 4.44 3.37 3.90	item 2 1 A.49 1.19 3.6 3.79 3.16 3.18	ften 3 1.31 1.02 3.15 3.18 2.23	iten 4 2.96 2.81 4.25 4.23 6.09 1.91	Item 5 4.48 4.4 4.39 1.99 2.08 0.86	item 0 0.03 1.5 4.4 4.02 1.50 0.33	item 7 4.37 3.32 0.56 4.79 2.69 0.44	item 8 3.48 3.48 3.69 0.38 2.86 4.11 3.90	item 9 1.3 2.65 4.91 4.85 1.08	iten 10 1.05 3.45 2.43 4.25 9.65 3.23	item 11 3.95 4.75 3.11 0.29 1.07	item 12 0.34 2.8 3.21 0.03 2.72 2.5												1.32 2.84 4.78 0.47 3.43 0.09
ario 9 z de utili ario 0 ario 1 ario 2 ario 3 ario 4 ario 5 ario 6 ario 7	dad complets ften 0 0.82 3.31 1.82 2.2 0.74 1.03 3.38	ada: ften 1 4.61 4.72 2.27 4.44 3.37 3.90 4.71	iten 2 1 4.49 1.29 3.6 3.79 3.16 3.38 3.98	item 3 1.31 1.02 3.15 1.18 2.23 3.11	iten 4 2.90 2.81 4.25 4.23 0.09 1.91 2.84	Iten 5 4.48 4.4 4.39 1.99 2.08 9.80 3.88	iten 6 0.03 1.5 4.4 4.02 1.50 0.33 0.07	Iten 7 4,37 3,23 0,56 4,79 2,69 0,44 3,03	item 8 3.48 3.48 3.09 0.58 2.86 4.11 3.90 3.86	ftem 9 2.3 2.65 4.91 4.85 1.68 1.47 9.25	Iten 10 1.08 3.48 2.43 4.26 0.08 3.23 4.34	item 11 3.95 4.75 3.11 0.29 1.07 1.56 4.41	item 12 0.14 2.8 3.21 0.03 2.72 2.5 2.95												1.32 2.84 4.78 9.47 3.43 9.09 3.00
ario 9 iz de utili mario 0 mario 1 mario 2 mario 3 mario 4 mario 5 mario 6 mario 6 mario 7 mario 8	dad complets item 0 0.82 1.31 1.82 2.2 0.74 1.03 3.38 0.97	ada: item 1 4.61 4.72 2.27 4.44 3.37 3.96 4.71 4.34	iten 2 1 4.49 1.39 3.6 3.79 3.16 3.38 3.98 2.52	ften 3 1.31 1.02 3.15 3.18 2.23 3.11 1.81	item A 2.96 2.81 4.25 4.23 0.09 1.91 2.84 3.09	ften 5 4.48 4.4 4.39 1.99 2.68 9.86 3.88 1.82	iten 0 0.03 1.5 4.4 4.02 1.50 0.33 0.07 2.21	Iten 7 4.37 3.31 0.56 4.79 2.69 0.44 3.03 1.24	item 8 3.48 3.48 3.09 0.38 2.86 4.11 3.90 3.86 1.95	iten 9 1.3 2.65 4.92 4.85 1.47 0.25 4.77	Item 10 1.05 3.45 2.43 4.26 9.65 1.23 4.34 1.25	3.95 4.73 3.11 0.29 1.67 1.58 4.41 4.39	iten 12 0.34 2.8 3.21 0.03 2.72 2.5 2.95 0.52												
ir de utili uario 0 uario 0 uario 1 uario 2 uario 3 uario 3 uario 5 uario 6 uario 7 uario 9 uario 9	dad complete Iten 0 0.82 1.11 1.82 2.2 0.74 1.03 1.18 0.97 1.44 1.05	ada: Item 1 4.61 4.72 1.27 4.44 3.37 3.96 4.71 4.24 9.06 1.18	iten 2 1 4.49 1.19 1.19 1.10 1.1	item 3 1.31 1.02 3.15 3.18 2.23 3.11 1.61 4.46 2.51 2.59	ften 4 2-90 2-81 4-25 4-22 6-80 1-91 2-84 3-90 4-24 4-83	iten 5 4.48 4.4 4.39 1.99 2.68 0.80 1.88 1.82 3.6	ften 6 0.03 1.5 4.4 4.02 1.50 0.33 0.07 2.21 2.25	Iten 7 4.27 3.21 0.50 4.79 2.69 0.44 3.03 1.24 3.41	Item 8 3.48 3.48 3.48 3.69 0.58 2.86 4.11 3.66 1.95 1.73	1ten 9 1.3 2.65 4.92 4.65 1.47 0.55 4.77 4.74	item 10 1.05 3.46 2.43 4.26 9.65 3.23 4.34 1.25 3.28	item 11 3.95 4.75 3.11 0.29 1.67 1.58 4.41 4.39 4.5	iten 12 0.34 2.8 3.21 0.03 2.72 2.5 2.95 0.52 4.0												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
ir de utili uario 0 uario 0 uario 1 uario 2 uario 3 uario 3 uario 5 uario 6 uario 7 uario 9 uario 9 u ario 9 u ario 9	dad complets Item 0 0.52 1.52 2.2 0.74 1.03 3.18 0.97 1.44 3.08 tudes entre	ada: Item 1 4.01 4.01 4.72 1.27 4.44 3.37 3.90 4.73 4.34 0.00 1.18 usuarios se	iten 2 1 4.49 1.39 3.6 3.79 3.10 3.10 3.18 1.98 2.52 1.38 2.97 3.99	item 3 1.31 1.02 3.15 3.15 3.15 3.11 1.02 3.15 3.11 1.03 3.11 1.03 3.11 1.03 3.11 1.03 3.11	Item 4 2.90 2.81 4.25 4.25 1.90 1.91 1.90 1.91 1.284 1.20 1.425 4.25	Item 5 A.48 A.4 A.39 I.99 I.90 I.88 I.82 J.60 I.82 J.61 I.82 J.62 I.83 I.82 I.83 I.84 I.85 II.85 II.85 II.85 II.85 III.85 IIII.85 III.85 I	item 6 0.03 1.5 4.4 4.02 1.50 0.33 0.07 2.21 2.35 4.18	ften 7 4,27 3,31 0,50 4,79 2,09 0,44 3,63 1,24 3,43 0,5	Item 8 1.48 1.48 1.09 0.38 2.86 4.11 1.90 1.93 1.93 1.73 4.04	item 9 1.3 2.65 4.91 4.85 1.47 4.77 4.77 4.74 3.82	11cm 10 1.65 2.43 4.25 4.25	item 11 3.95 4.75 3.11 9.29 1.67 1.38 4.41 4.39 4.5	iten 12 0.34 2.8 3.21 0.93 2.77 2.5 2.95 0.52 4.6 3.61												1.32 2.84 4.78 9.47 3.43 9.09 3.00 4.04
ir de utili sario 0 sario 0 sario 1 sario 1 sario 3 sario 4 sario 5 sario 6 sario 7 sario 8 sario 9 sario 9 sario 9 s de simili	dad complets Item 0 0.52 1.11 1.52 2.2 0.74 1.03 1.15 0.97 1.44 1.08 tudes entre	ada: Item 1 4.01 4.72 1.27 3.37 3.37 3.90 4.71 4.34 9.00 1.18 usuarios se	item 2 1 4.49 1.39 1.09 3.0 3.10 3.15 3.15 2.52 1.38 2.99 gin la métri 1 Unvari	Item 3 1.31 1.02 3.15 3.15 3.15 3.11 1.02	Item 4 2.90 2.81 4.22 4.23 6.00 1.91 2.84 4.23 4.25	Item 5 4.48 4.49 4.39 1.99 2.68 6.86 1.82 1.82 2.12 Usuario 4	item 6 0.03 1.5 4.4 4.02 1.50 0.33 0.07 2.21 2.35 4.18	Iten 7 4.37 3.21 0.20 1.24 3.41 0.5 0.5 0.5	item 8 3.48 3.09 0.38 2.60 4.11 3.60 1.93 1.73 4.04	item 9 1.3 2.05 4.91 4.85 1.47 0.35 4.77 4.74 3.82 Usuario 7	Item 20 1.00 3.40 3.45 4.25 4.25 4.25 4.25 4.25 4.25 4.25 4.25 4.25 4.25	item 11 3.95 4.75 1.11 9.29 1.67 1.38 4.41 4.39 4.5 9.45	2 2 2 3 3 3 3 3 3 3												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
iz de utili sario 0 sario 0 sario 1 sario 1 sario 3 sario 2 sario 5 sario 5 sario 5 sario 6 sario 6 sario 7 sario 8 sario 9 sario 9	dad complete item 0 0.82 3.31 1.82 2.2 0.74 1.03 3.38 3.48 3.50 tudes entre	ada: ftem 1 4.01 4.72 1.27 4.44 3.17 3.90 4.71 6.00 1.18 usuarios se Usuario se Usuario se	item 2 1 4.49 1.39 3.6 3.79 3.10 3.35 3.98 2.52 1.38 2.99 1.00 1 Unuari	ften 3 1.31 1.92 3.15 3.18 2.23 3.11 1.61 4.40 2.51 2.59 ica cosin io 2 0.72	ften 4 2.50 2.85 4.25 4.25 4.23 6.09 1.91 2.84 4.25 4.23 4.23 4.23 4.23 4.23 4.25	Item 5 A.48 A.4 A.5 A.	item 6 0.01 1.5 4.4 4.02 1.30 0.33 0.07 2.21 2.35 4.18	ften 7 4.37 3.31 0.50 4.79 2.69 0.44 3.03 1.24 0.5 0.5 0.6	Item 8 1.48 1.49 0.18 1.49 0.18 1.40 1.40 1.41 1.96 1.41 1.96 1.40	Iten 9 1.13 2.05 4.91 4.85 1.88 1.47 0.25 4.77 4.74 3.62 Usuario 7 0.73 0.73	1ten 10 1.68 3.45 2.43 4.26 4.26 3.22 4.31 1.22 3.28 4.28 Usuario 8	1tem 11 3.95 4.75 3.91 4.75 3.91 4.75 4.41 4.39 4.5	item 12 0.34 2.8 3.21 0.93 2.77 2.5 2.95 0.52 4.0 3.01 0.93 0.93 0.95 0.95 0.95 0.95 0.95												1.32 2.84 4.78 9.47 3.43 9.09 3.00 4.04 1.90
ario 9 ir de utili ario 0 ario 1 ario 2 ario 3 ario 3 ario 5 ario 5 ario 6 ario 7 ario 8 ario 9 4 de simili	dad complets Item 0 0.82 1.11 1.82 2.2 0.74 1.03 1.18 0.97 1.44 1.08 tudes entre Usuario 0 0.81	add: Item 1	iten 2 1 A.49 1 A.49 1 3.0 3.79 3.10 1 3.26 3.38 3 3.98 3.98 3 2.52 1 3.38 3 2.52 1 3.38 3 2.99 3 gin la nétri 1 Ussari 1 Ussari 1 Constitution Constitu	item 3 1.31 1.62 1.32 1.62	ften 4 2.90 2.81 4.25 4.25 4.23 0.09 1.91 2.84 4.28 4.28 4.28 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 1.92 4.28 4.03 4.28	Iten 5 A.46 A.4 A.4 A.9 A.	item 6 0.01 1.5 4.4 4.02 1.30 0.33 0.07 2.21 2.35 4.18	ften 7 4.37 3.31 0.50 4.79 2.69 0.44 3.63 1.24 0.5 0.5 0.6 0.77	Item 8 1.48 1.69 0.18 1.69 0.18 1.90 1.86 1.90 1.73 1.73 1.73 1.74 1.90 0.89 0.87	Item 9 1-3 2-05 4-91 4-85 1-08 1-47 6-25 4-77 4-77 3-82 0-79 0	2ten 30 1.08 2.43 4.26 0.08 1.22 1.22 1.22 1.22 1.22 1.22 1.23 4.26 0.08	1tem 11 3.95 4.75 1.11 0.29 1.67 1.36 4.41 4.39 4.5 0.45	Item 12 0.14 2.8 3.21 0.03 2.72 2.5 2.95 0.51 3.01 3												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
ario 9 ario 9 ario 0 ario 1 ario 2 ario 3 ario 3 ario 5 ario 6 ario 7 ario 8 ario 9 ario 9 ario 9 ario 9 ario 9	dad complets item 0 0.82 3.31 1.82 2.2 0.74 1.83 1.84 3.06 1.44 1.06 1.44 1.06 1.45 0.97 1.44 0.97 0.85 0.82 0.82 0.82 0.83 0.72	ada: Item 1 4.01 4.72 2.27 4.44 3.37 4.73 4.34 9.60 1.18 Usuario 50 Usuario 9.8	A.49 1.39 3.00 3.79 3.15 3.	1.31 1.02 3.15 3.15 3.15 3.15 3.15 3.15 3.15 3.15 3.17 3.	ften 4 2.96 2.85 4.23 4.23 6.99 1.91 2.65 4.28 4.28 4.28 4.28 4.28 4.28 4.27 6.77 6.77 6.77 6.75 6.75	Iten 5 A.48 A.4 A.39 I.99 I.99 I.65 I.62 I.62 I.62 I.62 I.63 I.64 I.65 I.65 I.65 I.65 I.67 I.68 I	item 0 0.03 1.5 1.5 1.5 1.5 1.50 0.33 0.07 2.21 2.35 4.18 Usuar	1ten 7 4.37 3.21 0.50 4.79 2.00 1.03 1.24 3.41 0.5 1.05 1	Item 8 1.46 1.40	Item 9 1.3 2.05 4.91 4.85 2.08 1.47 4.77 4.74 2.02 0.73 6.79 6.77 6.79 6.77	2 tem 20 1	Item 11 1 2-95 4-72 1-11 1-12 1-	Item 12 0.14 2.8 3.21 0.03 2.72 2.9 2.95 0.02 4.0 3.01												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
ir de utili uario 0 uario 0 uario 1 uario 2 uario 3 uario 3 uario 5 uario 6 uario 7 uario 8 uario 9 u ario 1 u ario 1 u ario 2 u ario 3	dad complets item 0 0.82 1.31 1.82 2.2 0.74 1.03 1.03 1.04 1.05 tudes entre Usuario 0 0.81 0.72 0.77	ada:	iten 2 1 1 4.49 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1 1.19 1.19 1 1 1.19 1 1 1 1 1 1 1 1 1	1.31 1.92 3.15 3.15 3.15 3.15 3.15 3.16 3.17 3.	tten A 2.90 4.25 4.25 4.25 4.25 4.26	21cm 5 4.48 4.4 4.39 2.08 0.86 1.52 2.08 0.86 2.12 Usuario A 0.79 0.8 0.71 0.72 0.73	Item 0 0.03 1.5 4.4 1.5 4.62 1.20 0.33 0.07 2.21 2.23 4.18	Item 7 A.37 A.37 A.37 A.37 A.37 A.37 A.79 A.79 A.79 A.79 A.79 A.74	Item 8 3.48 3.99 0.38 2.86 4.11 3.90 3.86 1.93 1.93 4.04	Item 9 1.3 2.65 4.92 4.85 1.08 1.47 0.53 4.74 3.62 0.73 0.73 0.77 0.61	21cm 20 1.08 3.45 4.26 4.26 4.25 4.26	Item 11 1 3.95 1 4.73 1 1 1 1 1 1 1 1 1	ften 12 0.14 2.8 3.21 0.93 2.72 2.72 3.21 0.93 3.21 0.93 0.52 6.5 0.52 0												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
ario 9 ario 0 ario 0 ario 0 ario 1 ario 2 ario 2 ario 3 ario 3 ario 5 ario 6 ario 6 ario 7 ario 6 ario 7 ario 8 ario 9 ario 9 ario 9 ario 1 ario 0 ario 1 ario 1 ario 1 ario 1 ario 3 ario 3	dad complete Item 0 0.52 1.31 1.82 2.2 1.74 1.03 1.44 1.05 1.45 1.45 1.45 1.65 1.46 1.65 1.65 1.65 1.67 0.77 0.79	ada: Item 1 4.01 4.01 4.72 1.27 4.44 3.37 3.50 4.71 4.34 9.80 1.18 Usuario se Usuario 9.8 0.8	Stem 2 1	2 ten 3 1.31 1.02 1.32 1.03		21cm 5 4.48 4.4 4.39 1.99 1.99 1.99 1.95 1.55 1.65 2.16 2.12 2	0.03 1	iten 7 4.37 3.31 6.50 4.79 2.09 6.44 3.63 1.24 3.43 6.5 6.5 6.77 6.77 6.70	Item 8 3.48 3.48 3.48 3.48 4.5 3.48 4.5 3.48 4.5 3.48 4.5 4.	Item 9 1.3 2.65 4.92 4.65 1.08 1.47 0.25 4.77 4.77 4.77 0	1ten 30	1.05 1.	3 tem 12 0.14 2.8 3.21 0.03 2.72 2.5 2.90 3.01												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
artio 9	dad complete 2 ten 0 0.82 1.82 2.22 0.74 1.03 3.18 0.97 1.44 1.68 1.68 1.68 0.72 0.73 0.75 0.77 0.77 0.79	ada: Item 1 4.01 4.01 4.01 4.72 1.27 1.27 1.27 3.90 4.44 9.80 1.18 9.80 1.18 9.80 9.80 9.80 9.80 9.80 9.80 9.80 9.8	1 1 1 1 1 1 1 1 1 1	1.31 1.02 1.32	Item 4 2-50 2-81 4-22 4-23 1-91 1-91 1-92 4-24 4-03 4-24 1-91 4-25	2ton 5 4.46 4.45 4.45 4.47 4.49 2.66 2.66 2.66 2.66 2.67 4.67	Etem 6 1.5	Item 7 4.27 3.31 0.50 4.79 2.69 0.44 3.43 0.5 1.24 3.43 0.5 0.6 0.71 0.67 0.67 0.67 0.67 0.67 0.67 0.77 0.67 0.7	Item 8 3.46	Item 9 2.3 2.00 4.91 4.85 2.08 1.47 4.74 3.82 4.77 4.74 3.82 4.77 6.79 6.77 6.81 6.77 6.81 6.77 6.81 6.79 6	Item 20	2 ten 21 3,95 4,75 9,21 1,07 1,07 1,07 1,05	1ten 12 0.34 2.8 3.21 0.03 2.72 2.5 2.95 0.52 4.6 3.01 0.03 0.												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
iz de utilia iz de	dad complete 2 ten 0 0.62 3.31 1.62 2.22 0.74 1.63 1.68 1.68 1.68 1.68 0.72 0.77 0.79 0.81 0.77 0.79 0.82 0.97	ada: Item 1 4.51 4.72 4.44 3.37 3.50 4.71 4.14 6.05 1.18 Usuario 0.8 0.8 0.8 0.8 0.7 0.8 0.8 0.8 0.8 0.7 0.8 0.8 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.7 0.8 0.8 0.8 0.7 0.8	tem 2 1 4.49 1.19 1.50 1.50 1.79 1.20 1.38 1.98 1.98 1.98 1.98 1.99 1.18 1.99 1 1 1 1 1 1 1 1 1 1	1ten 3 1.31 1.97 3.13 1.97 3.13 3.13 3.14 1.97 3.15	Item 4 2.90 2.81 4.23 4.23 4.23 4.23 4.23 4.23 4.24 4.25	2 ten 5		Iten 7 4.37 3.31 6.35 4.79 2.09 6.44 3.41 6.5 6.	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	Item 9 1.3 2.05 4.92 4.65 1.68 1.47 4.75 1.68 1.47 4.75 1.55 4.77 4.75 1.55 7.70 6.73 6.77 6	1tem 10	1,55 1,12 1,12 1,12 1,12 1,12 1,12 1,12 1,12 1,13	\$tem 12 0.14 2.8 3.21 0.03 2.72 2.5 2.95 3.01 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 4.0 0.02 0.02 4.0												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90
iz de utili usrio 9 usrio 6 usrio 2 usrio 2 usrio 3 usrio 4 usrio 5 usrio 6 usrio 6 usrio 7 usrio 8 usrio 8 usrio 9 a de simili	dad complete 2 ten 0 0.82 1.82 2.22 0.74 1.03 3.18 0.97 1.44 1.68 1.68 1.08 1.09	ada: Item 1 4.63 4.72 5.27 4.44 3.50 4.73 4.34 4.34 5.18 5.18 5.18 5.18 5.18 6.0 6.0 6.0 7.18 8.0 9.0 10.0	Item 2 1 A.49 A.49 A.59 A.59 A.59 A.59 A.59 A.50	1.31 1.02 1.32	Item 4 2-50 2-81 4-22 4-23 1-91 1-91 1-92 4-24 4-03 4-24 1-91 4-25	2ton 5 4.46 4.45 4.45 4.47 4.49 2.66 2.66 2.66 2.66 2.67 4.67	1:en 0	Item 7 4.27 3.31 0.50 4.79 2.69 0.44 3.43 0.5 1.24 3.43 0.5 0.6 0.71 0.67 0.67 0.67 0.67 0.67 0.67 0.77 0.67 0.7	Item 8 3.46	Item 9 2.3 2.00 4.91 4.85 2.08 1.47 4.74 3.82 4.77 4.74 3.82 4.77 6.79 6.77 6.81 6.77 6.81 6.77 6.81 6.79 6	Item 20	1 tem 13 3.95 4.72 3.121 4.72 3.121 4.72 3.121 4.72 3.121 4.72 3.121 4.72	1ten 12 0.34 2.8 3.21 0.03 2.72 2.5 2.95 0.52 4.6 3.01 0.03 0.												1.32 2.84 4.78 9.47 3.43 9.99 3.90 4.04 1.90

- Ignora la magnitud absoluta, centrando la comparación en la dirección.
- Genera resultados más estables ante sparsity.
- No distingue patrones inversos (usuarios con preferencias opuestas).



Distancia euclídea



- Basada en magnitud, sensible a la escala de ratings.
- Estable incluso con pocos vecinos.
- Suaviza predicciones, pero pierde capacidad de discriminación en altas dimensiones.

Análisis:

Pearson es útil con datos densos y correlaciones claras, pero falla con datos dispersos o vecinos aislados. Distancia coseno y distancia euclídea proporcionan predicciones más estables y confiables en escenarios con sparsity (datos faltantes) o ratings extremos.

Comparativa entre métodos de predicción

En los experimentos se aplicaron dos métodos de predicción sobre las distintas métricas de similitud:



1. Predicción simple:

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

Se basa directamente en las puntuaciones de los vecinos. Como se ha podido observar, tiende a favorecer métricas como **distancia coseno** y **euclídea**, que producen valores siempre positivos y homogéneos. Con matrices pequeñas o dispersas, este método ofrece predicciones más estables y dentro del rango original.

2. Predicción con diferencia respecto a la media:

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

Corrige sesgos individuales (usuarios que puntúan sistemáticamente alto o bajo).

Este enfoque se ajusta mejor al **coeficiente de Pearson**, ya que ambas fórmulas se basan en desviaciones respecto a la media. En matrices densas, produce estimaciones más precisas y con menor error promedio.



Resultados y tendencias generales

- En matrices pequeñas (5×10, 10×25), la predicción simple funciona mejor con coseno y euclídea, porque Pearson generaba correlaciones poco fiables y, al aplicar la fórmula con media, amplificaba el ruido.
- En matrices medianas (25×100), la diferencia por media mejoró la coherencia de Pearson, reduciendo el sesgo y produciendo valores más realistas.
- En matrices grandes (50×250, 100×1000), la predicción con media estabiliza Pearson, pero su rendimiento se ve limitado por la falta de coincidencias entre usuarios. En cambio, coseno se mantiene robusto con el método simple, y euclídea conserva su consistencia, aunque con menor discriminación.
- En general, las métricas angulares (como el coseno) responden mejor a entornos dispersos, mientras que las centradas en desviaciones (como Pearson) se benefician de mayor densidad y del uso de la media.

Tabla descriptiva

Métrica	Predicción simple	Predicción con media
Pearson	Inestable en datos escasos; puede dar valores fuera de rango.	Más coherente; corrige sesgos individuales. Ideal en matrices densas.
Coseno	Predicciones estables y consistentes. Ideal en matrices pequeñas o dispersas.	Mejora poco con la media, ya que ignora magnitud.
Euclídea	Predicciones suaves y conservadoras.	Resultados similares; poco sensible a sesgos.



Conclusión final

- **Pearson** es el más adecuado en matrices densas o moderadamente completas, especialmente junto a la predicción con diferencia respecto a la media.
- **Distancia coseno** ofrece resultados más consistentes en entornos dispersos o con pocos ítems comunes, siendo ideal con la predicción simple.
- **Distancia euclídea** es la más estable, pero su capacidad de discriminación disminuye al aumentar la dimensionalidad.
- En conjunto, la elección óptima depende del nivel de sparsity:
 - Para datos **densos → Pearson + media**.
 - Para datos dispersos → Coseno + predicción simple.
 - Para entornos mixtos o de gran tamaño → Euclídea + simple.