

Combinación de diferentes fuentes de datos en la recomendación de puntos de interés

Jaime Pascual Francés





Introducción



Sistemas de recomendación



Diseño e implementación



Pruebas y resultados



Conclusiones y trabajo futuro

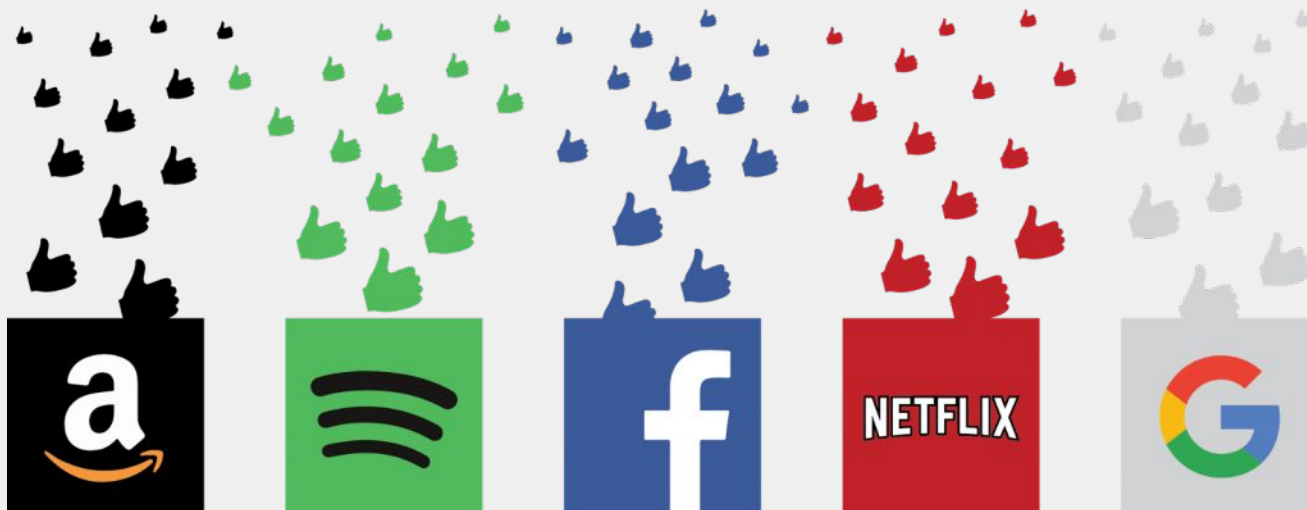




Introducción



Sistemas de recomendación



Puntos de interés



A scenic view of a coastal town, likely in Norway, featuring a large, rugged mountain in the background and a body of water in the foreground. The town is built on a hillside, and the water is calm with some small boats visible. The sky is overcast.





Sistemas de recomendación



Ejemplos



Popularidad

Se basa en la cantidad de usuarios que se han interesado en un artículo



Vecinos

Se basa en la similitud entre usuarios



Híbridos

Combina diferentes técnicas para mejorar la recomendación



Evaluación



Métricas de evaluación offline

Relevancia

Precision

Recall

Diversidad

Aggregate Diversity

Novedad

Expected
Popularity
Complement

Cobertura

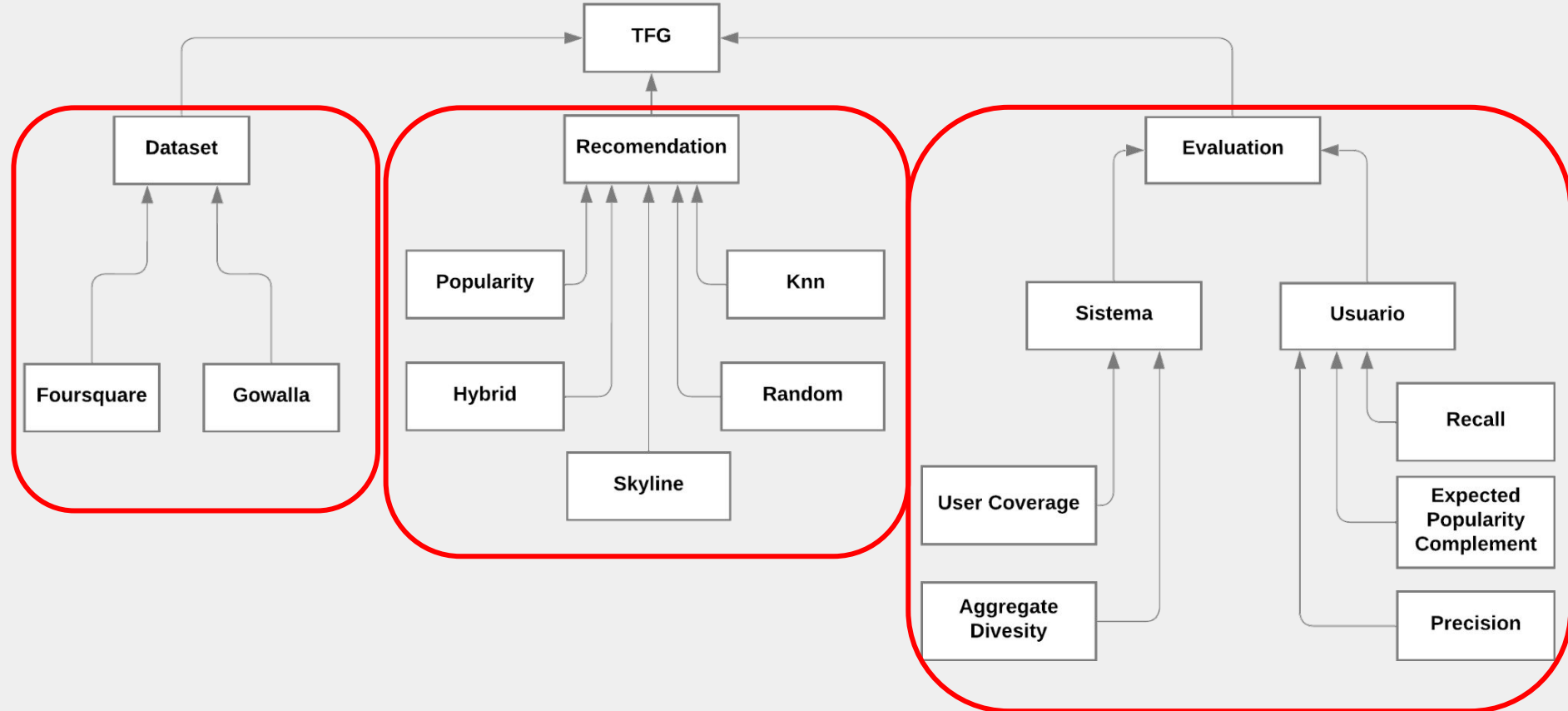
User Coverage



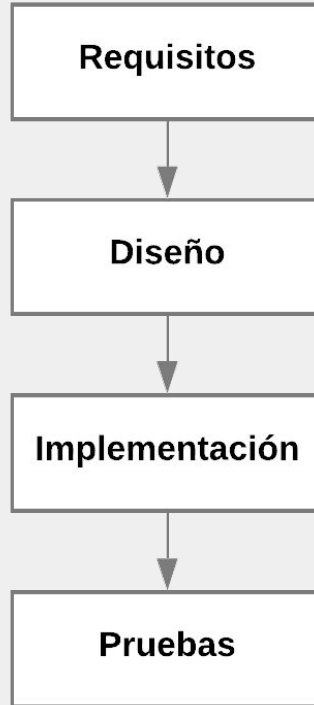
Diseño e implementación



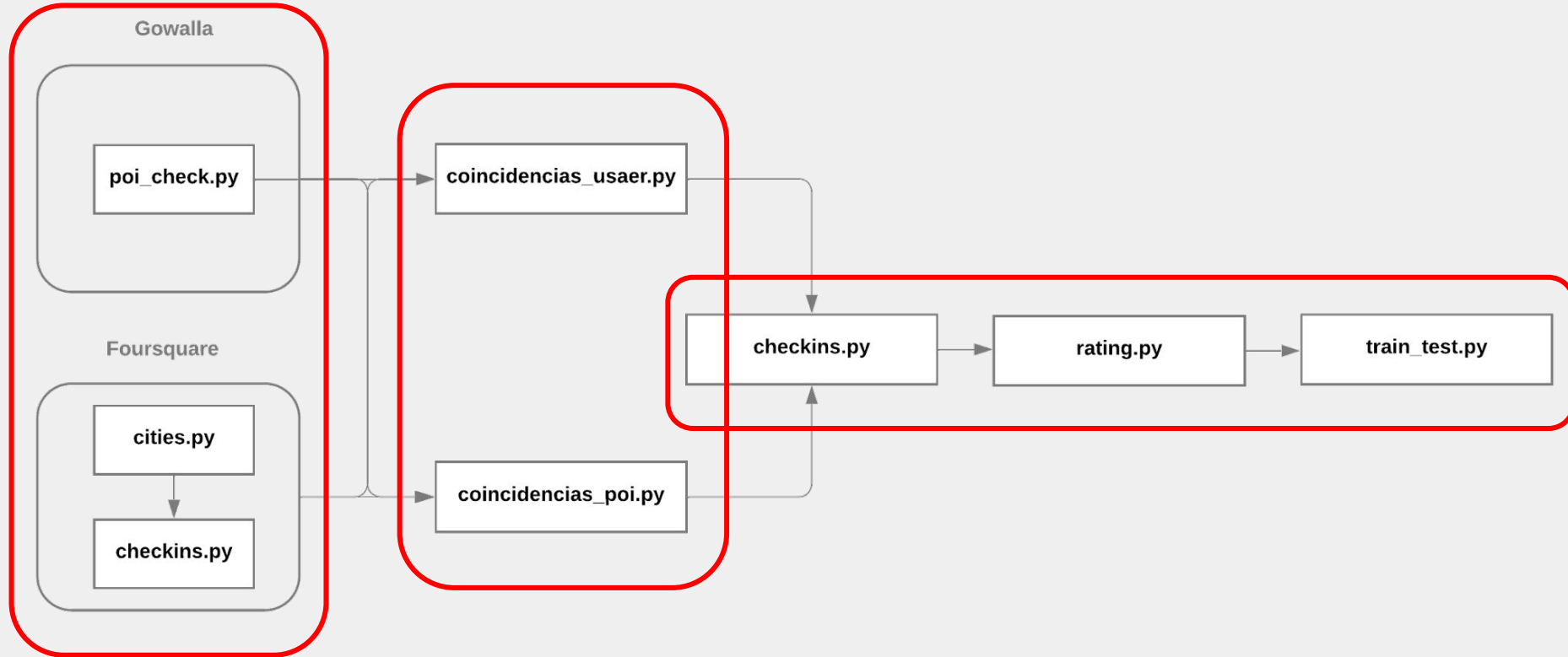
Estructura general



Ciclo de vida



Flujo de ejecución (Procesado de datos)





Pruebas y resultados





Datos de cada aplicación

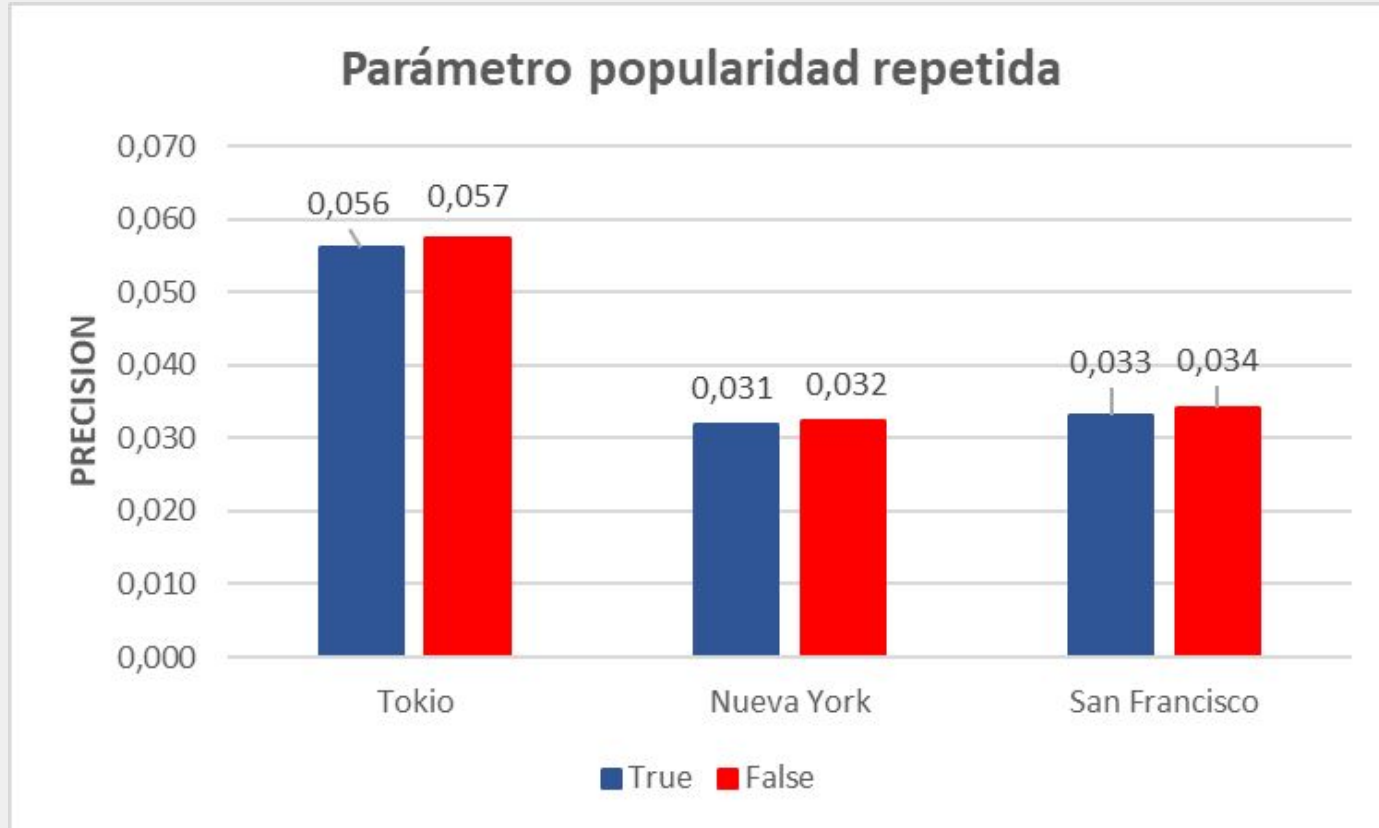
	Usuarios	POIs	Check-ins	Ciudades
Foursquare	266,909	3,680,126	33,263,633	415
Gowalla	196,586	1,280,969	6,442,892	-

Selección de ciudades

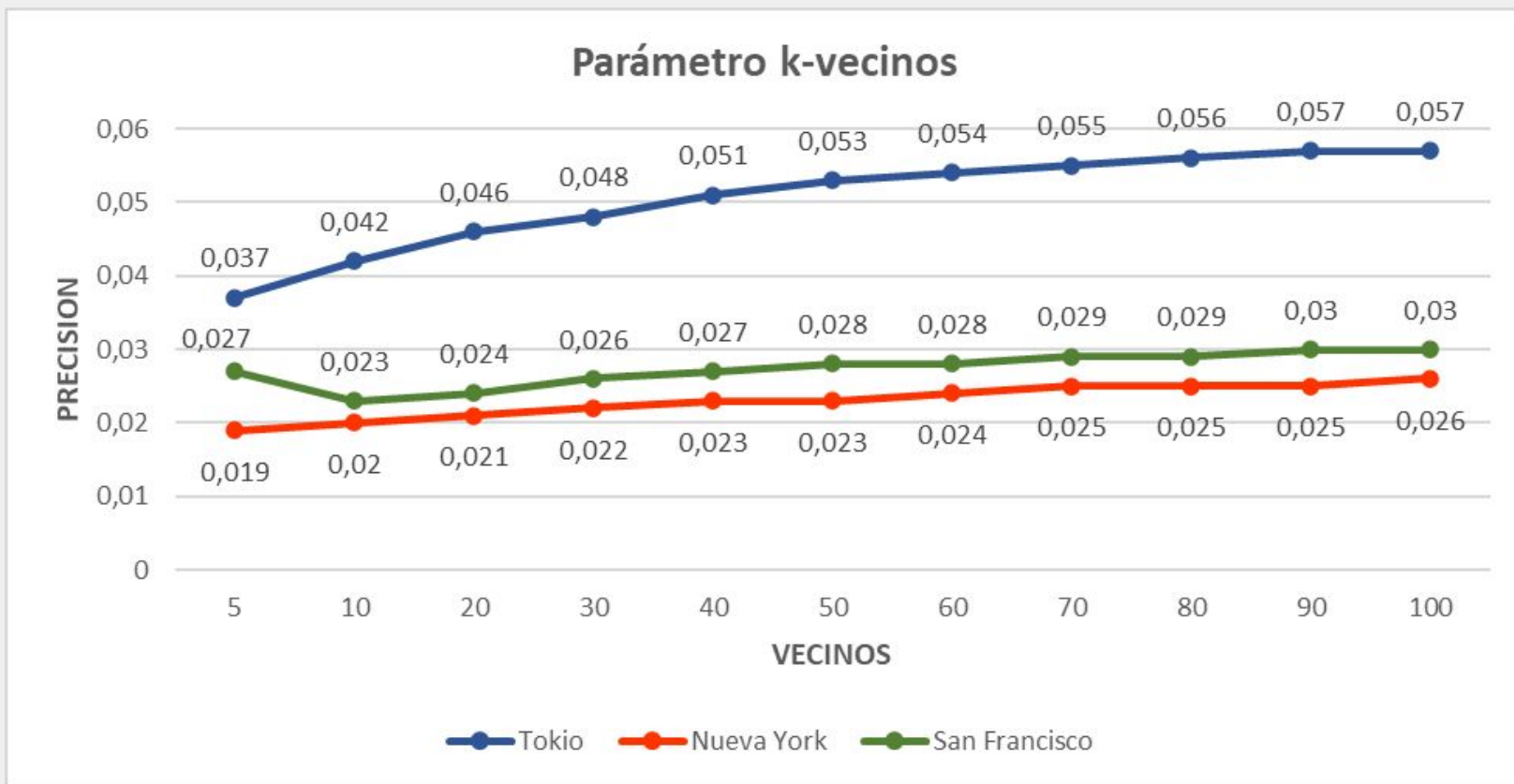
	Usuarios	POIs	Check-ins	Rating
Tokio	12,464	83,189	1,030,090	474,150
Nueva York	15,785	41,386	380,247	210,192
San Francisco	6,682	13,104	106,648	65,425

	Ratings Foursquare	Aumento	% de aumento
Tokio	379,320	4,100	1.08 %
Nueva York	168,153	10,698	6.36 %
San Francisco	52,340	15,994	30.55 %

Selección de parámetros (popularidad repetida)



Selección de parámetros (número de vecinos)



Tokio	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	7.5386355e-05	7.8790169e-05	0.9995846	74,380	10,612
Skyline	0.5794882	0.8377626	0.972587	24,351	10,472
Popularity	0.0574727	0.0904425	0.8536096	95	10,612
Knn	0.0573521	0.0887028	0.9045899	13,351	10,378
Hybrid	0.0504994	0.0817951	0.8915544	46,234	10,612
Random	1.3192612e-04	1.2714543e-04	0.9996105	74,489	10,612
Skyline	0.5799179	0.8385205	0.9738657	24,454	10,472
Popularity	0.0575009	0.0907488	0.8604287	95	10,612
Knn	0.0573562	0.0886455	0.9092653	13,464	10,379
Hybrid	0.0505089	0.0817766	0.8966666	46,284	10,612

Nueva York	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	9.0680101e-05	2.3381712e-04	0.9996865	36,604	9,925
Skyline	0.336668	0.9170562	0.9808087	11,278	9,676
Popularity	0.0323526	0.1390475	0.9251687	71	9,925
Knn	0.0253652	0.09184149	0.9683149	19,088	8,488
Hybrid	0.0273552	0.1189315	0.9481526	31,703	9,925
Random	5.0377834e-05	8.5660342e-05	0.9997322	36,718	9,925
Skyline	0.3370611	0.9181243	0.9841562	11,392	9,684
Popularity	0.0322921	0.1390092	0.9382898	72	9,925
Knn	0.0251765	0.0913275	0.974089	19,328	8,496
Hybrid	0.0272544	0.1185395	0.9574267	31,570	9,925

San Francisco	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	2.1727322e-04	3.5251896e-04	0.9992889	11,707	3,682
Skyline	0.2829165	0.9300513	0.9456978	4,065	3,559
Popularity	0.0340576	0.1760944	0.937286	63	3,682
Knn	0.0287088	0.1314895	0.9632036	6,831	2,859
Hybrid	0.0269419	0.1510288	0.9580159	11,594	3,682
Random	2.9875067e-04	9.1629808e-04	0.9994389	11,800	3,682
Skyline	0.284119	0.9332986	0.964609	4,158	3,559
Popularity	0.0314503	0.1658245	0.9443126	71	3,682
Knn	0.0274222	0.1271734	0.9726978	7,133	2,912
Hybrid	0.025774	0.1471707	0.9623470	11,534	3,682

Análisis de los resultados

Popularidad

Ciudad	Top
Tokio	20
Nueva York	16
San Francisco	3

Ciudad	Coincidencia
Tokio	93.33%
Nueva York	80%
San Francisco	33.33%

Nuevos puntos más visitados

POI	Top	Descripción
737	1	Aeropuerto
24820	2	Estadio baseball
2527	3	Puente

POI	Top	Descripción
1043	3	Tienda electrónica
57364	4	Gym
19101	6	Plaza



Conclusiones y trabajo futuro



Conclusiones

1. **Misma relevancia**
2. **Mejor novedad y diversidad**
3. **Eficacia**





Trabajo futuro

1. Distintas técnicas de recomendación
2. Distintas plataformas
3. Estudio de rutas turísticas

Muchas gracias por su atención



¿Preguntas?

	Usuarios	POIs	Check-ins	Rating
Tokio	849	7,309	29,784	16,290
Nueva York	6,111	13,755	97,154	71,106
San Francisco	7,113	14,485	185,941	120,406

	Rating Gowalla	Aumento	% de aumento
Tokio	13,032	2,795	21.45%
Nueva York	56,884	16,272	28.61%
San Francisco	96,324	9,366	9.72%

Tokio	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	1.6333938e-03	4.1474538e-03	0.9973428	6,283	551
Skyline	0.3290581	0.7996931	0.9832436	1,373	499
Popularity	0.03411978	0.0740604	0.9233216	95	551
Knn	0.0373016	0.0889116	0.9535372	2,163	504
Hybrid	0.0333938	0.0858767	0.9433859	4,165	551
Random	3.6297641e-04	3.7269899e-04	0.9993243	6,439	551
Skyline	0.338189	0.8166799	0.9956476	1,538	508
Popularity	0.007441	0.0147034	0.9281864	67	551
Knn	0.032874	0.0806071	0.981821	2,281	508
Hybrid	0.0174229	0.0523354	0.954302	4,148	551

Nueva York	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	3.0496257e-03	7.9266318e-04	0.9992413	12,764	3,607
Skyline	0.2821307	0.940221	0.9810545	5,332	3,520
Popularity	0.0380926	0.1467304	0.9256439	77	3,607
Knn	0.0428059	0.13615	0.959412	7,490	3,065
Hybrid	0.0372054	0.1430849	0.9448733	11,794	3,607
Random	2.772387e-04	9.5812299-04	0.9995265	12,853	3,607
Skyline	0.2834043	0.9438676	0.9907201	5,421	3,525
Popularity	0.031023	0.1238669	0.9624801	93	3,607
Knn	0.0416368	0.1318707	0.981303	7,803	3,067
Hybrid	0.0370945	0.142112	0.9735638	11,579	3,607

San Francisco	Precision	Recall	EPC	Aggregate diversity	User coverage
Random	4.4062568e-04	1.0065721-03	0.9989523	13,678	4,539
Skyline	0.348474	0.9363187	0.9740478	7,211	4,489
Popularity	0.0320115	0.1124272	0.9294165	94	4,539
Knn	0.0403656	0.117811	0.9613333	7,036	3,879
Hybrid	0.0320776	0.1106193	0.9562269	12,445	4,539
Random	2.4234413e-04	2.9015108e-04	0.9992082	13,716	4,539
Skyline	0.348553	0.9368733	0.9810598	7,249	4,492
Popularity	0.0292575	0.1039796	0.9422459	96	4,539
Knn	0.0393658	0.1135853	0.9711517	7,552	3,939
Hybrid	0.0307997	0.1086237	0.9637322	12,517	4,539

Análisis de los resultados

Vecinos

Ciudad	Cambio
Tokio	12.54%
Nueva York	14.79%
San Francisco	28.98%

Ciudad	Vecinos
Tokio	36
Nueva York	100
San Francisco	100