

MODUL 8

Recommender System (Sistem Rekomendasi)

Recommender System (RS) adalah perangkat lunak dan teknik yang memberikan saran untuk item yang paling mungkin menarik bagi pengguna tertentu.

Tujuan utama dari sistem rekomendasi adalah untuk membantu pengguna menemukan item yang relevan, seperti barang apa yang akan dibeli, musik apa yang harus didengarkan, atau berita online apa yang harus dibaca, sesuai riwayat perilaku pengguna di masa lampau.

Riwayat perilaku pengguna menggambarkan preferensi pengguna terhadap suatu hal, dimana perilaku pengguna bisa terekam dari umpan balik yang diberikan secara langsung (*explicit feedback*) dan secara tidak langsung (*implicit feedback*) ketika pengguna berinteraksi dengan sebuah aplikasi atau situs.

Contoh *explicit feedback*, yaitu pemberian rating, *like*, *dislike* atau *subscribe* akun, sedangkan *implicit feedback* contohnya seperti riwayat pembelian, riwayat kunjungan, atau klik pada sebuah link.

Metode-metode RS yang dapat digunakan pada HRS diantaranya *Content-Based Filtering*, *Collaborative Filtering*, dan *Hybrid Filtering*.

Content Based Filtering

1	<pre>#import library import pandas as pd from scipy import sparse</pre>
2	<pre>#dataset ratings = pd.read_csv('ratings.csv') movies = pd.read_csv('movies.csv') ratings = pd.merge(movies, ratings).drop(['genres', 'timestamp'], axis=1) print(ratings.shape) ratings.head()</pre>
3	<pre>userRatings = ratings.pivot_table(index=['userId'], columns=['title'], values='rating') userRatings.head() print("Before: ", userRatings.shape) userRatings = userRatings.dropna(thresh=10, axis=1).fillna(0, axis=1) print("After: ", userRatings.shape)</pre>
4	<pre>#matriks korelasi corrMatrix = userRatings.corr(method='pearson')</pre>

	<code>corrMatrix.head(100)</code>
5	<pre>#Hitung similarity def get_similar(movie_name,rating): similar_ratings = corrMatrix[movie_name]*(rating-2.5) similar_ratings = similar_ratings.sort_values(ascending=False) #print(type(similar_ratings)) return similar_ratings</pre>
6	<pre>romantic_lover = [(" (500) Days of Summer (2009)",5),("Alice in Wonderland (2010)",3),("Aliens (1986)",1),("2001: A Space Odyssey (1968)",2)] similar_movies = pd.DataFrame() for movie,rating in romantic_lover: similar_movies similar_movies.append(get_similar(movie,rating),ignore_index = True) similar_movies.head(10)</pre>
7	<code>similar_movies.sum().sort_values(ascending=False).head(20)</code>
8	<pre>action_lover = [("Amazing Spider-Man, The (2012)",5),("Mission: Impossible III (2006)",4),("Toy Story 3 (2010)",2),("2 Fast 2 Furious (Fast and the Furious 2, The) (2003)",4)] similar_movies = pd.DataFrame() for movie,rating in action_lover: similar_movies = similar_movies.append(get_similar(movie,rating),ignore_index = True) similar_movies.head(10) similar_movies.sum().sort_values(ascending=False).head(20)</pre>

Latihan

1. Cobalah semua source code diatas
2. Lihat hasil outputnya dan berikan penjelasan
3. Buatlah laporan yang berisi penjelasan dari source code tersebut, kumpulkan di Hebat dengan nama PrakClustering_NIM

Tugas

1. Kerjakan tugas secara kelompok
2. Buatlah program Sistem Rekomendasi dengan data yang lain (data silahkan dicari sendiri)
3. Buatlah laporan dari program yang dibuat
4. File yang dikumpulkan data asli, program phyton, dan laporan
5. Penamaan Tugas TugasRS_Kelompok XXX.Zip
6. Dikumpulkan paling lambat Senin / 15 Mei 2023 Pukul 23.59 Wib