Item-Based Collaborative Filtering: Step-by-Step Example

This presentation walks through a complete example of item–based collaborative filtering, demonstrating how to calculate similarities between items and generate personalized recommendations.

MUKESH KUMAR

What is Item-Based Collaborative Filtering?

Definition

A recommendation technique that finds similarities between items based on user ratings patterns.

Core Concept

Recommends items similar to those the user already liked. Focuses on item relationships rather than user behaviors.

Advantages

More stable than user-based filtering. Handles the cold start problem better. Works well with sparse datasets.

Our Example Setup

Small Dataset

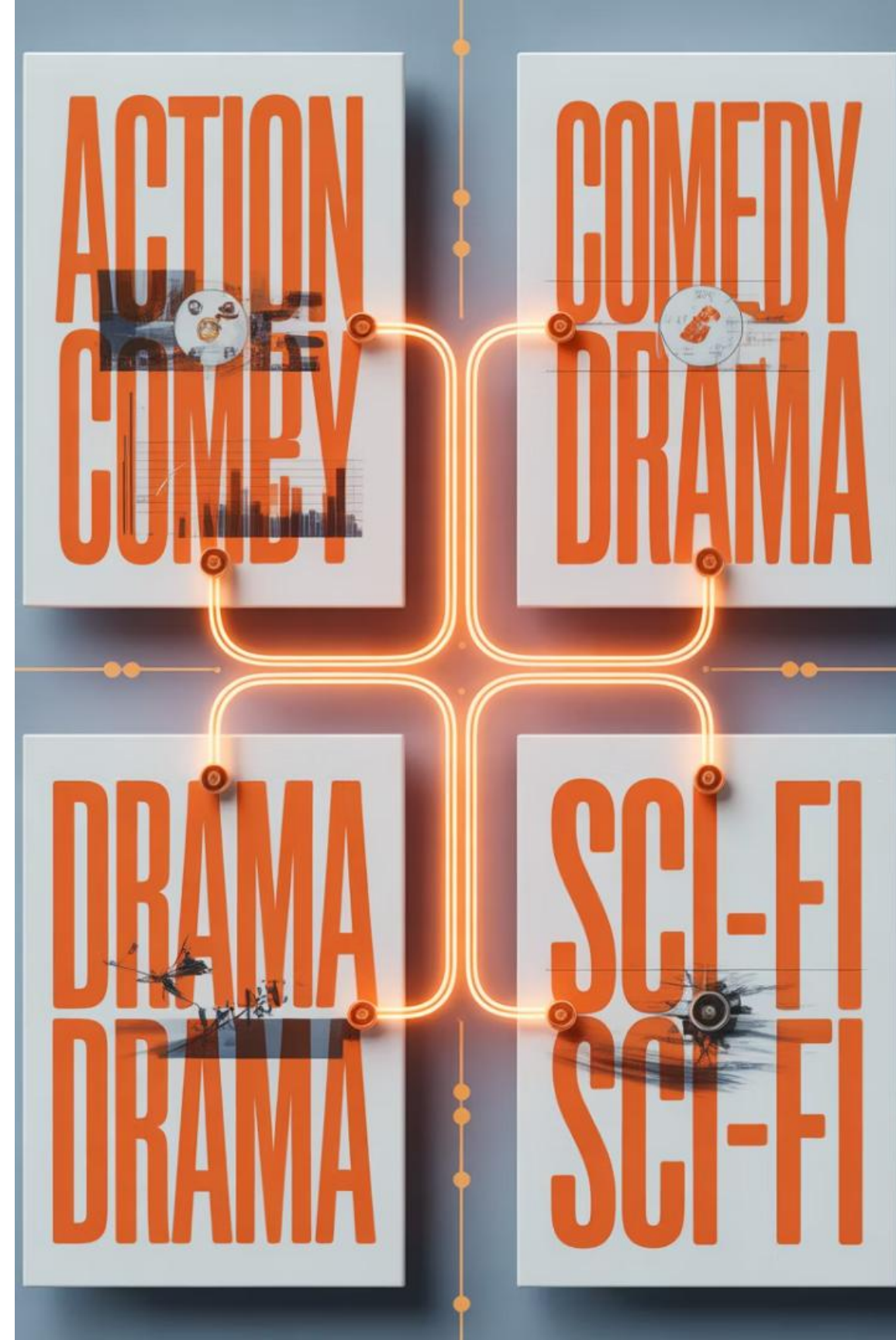We'll use 4 movies and 2 users to demonstrate the complete process.
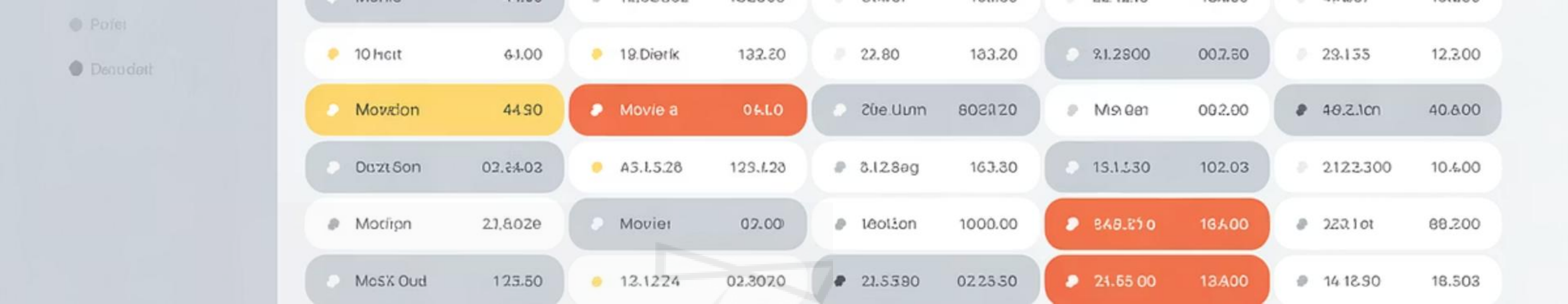
Rating Scale

Users rate movies on a scale from 1–5. Unrated items are marked as 0.

Goal

Predict which movie to recommend next for User 1 based on similarities.

User-Item Rating Matrix

User / Movie	Movie A	Movie B	Movie C	Movie D
User 1	5	3	0	0
User 2	4	0	4	2

The matrix shows existing ratings. Zero values indicate unwatched movies. These are the ratings we'll try to predict.

Step 1: Compute Item Vectors

Extract Columns

We represent each movie as a vector of ratings from all users.
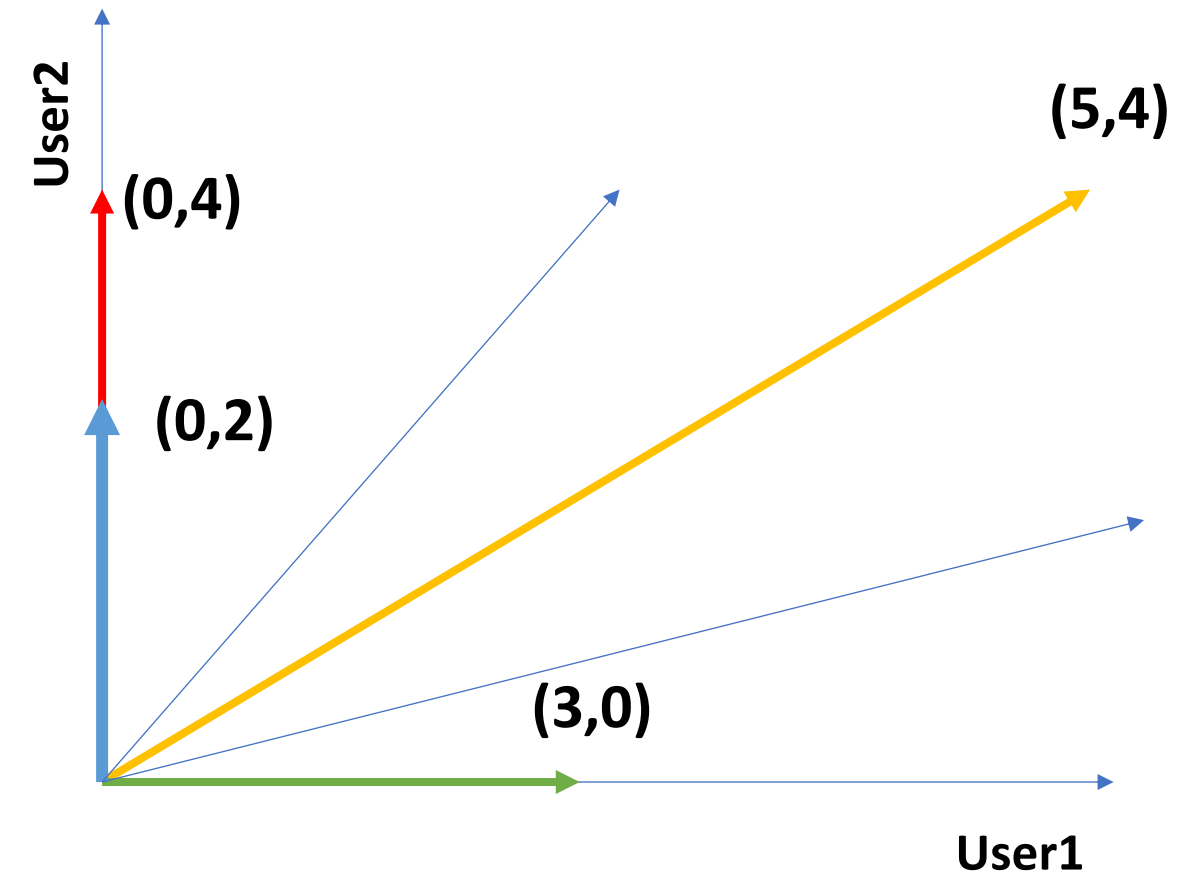
Create Vectors

Movie A → [5, 4]

Movie B → [3, O]

Complete Set

Movie C → [O, 4]

Movie D → [O, 2]

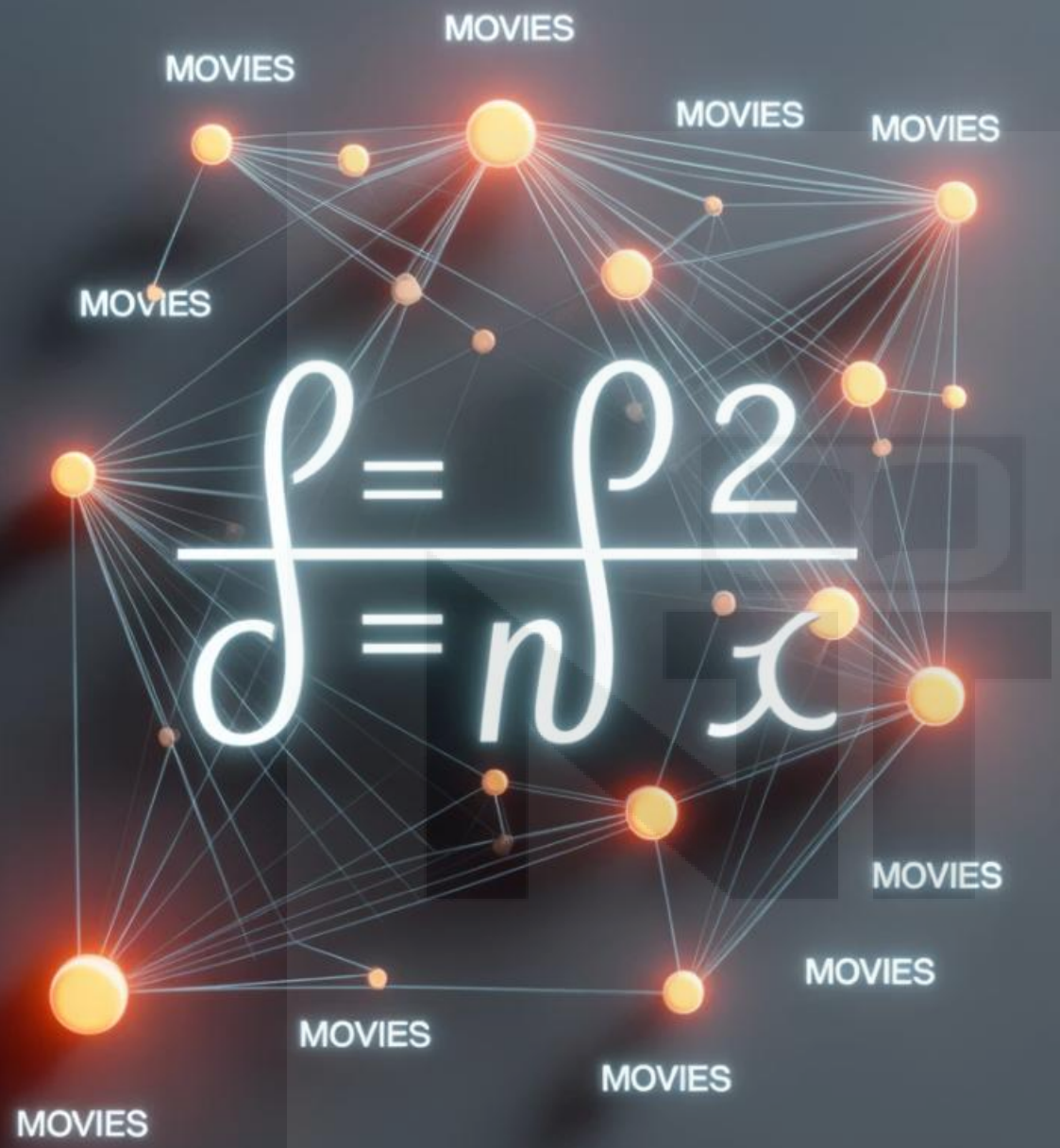# Step 2: Calculate Item-to-Item Similarities

Formula:

$$\text{cosine}(A, B) = \frac{A \cdot B}{||A|| \times ||B||}$$

Example: Sim(Movie A, Movie B):

$$\frac{5 \times 3 + 4 \times 0}{\sqrt{5^2 + 4^2} \times \sqrt{3^2 + 0^2}} \approx 0.78$$

Sample similarities:

- Sim(C, A) ≈ 0.62

- Sim(C, B) ≈ 0.1

- Sim(D, A) ≈ 0.624

- Sim(D, B) = 0

Step 3: Predict Rating for User 1 on Movie C

- User 1 rated Movie A = 5, Movie B = 3

- Similarities to Movie C: Sim(C,A)=0.62, Sim(C,B)=0.1

- Prediction:

$$\frac{(0.62 \times 5) + (0.1 \times 3)}{0.62 + 0.1} = \frac{3.1 + 0.3}{0.72} \approx 4.72$$

Step 4: Predict Rating for User 1 on Movie D

- User 1 rated Movie A = 5, Movie B = 3

- Similarities to Movie D: Sim(D,A)=0.624, Sim(D,B)=0

- Prediction:

$$\frac{(0.624 \times 5)}{0.624} = 5.0$$

Final Recommendation

4.72

Movie C

High predicted rating based on item similarities

5.0

Movie D

Highest predicted rating

Recommendation for User 1: Movie D

Movie D receives our highest prediction score of 5.0, making it the optimal recommendation for User 1.