→ Recommendation Systems:-

Nearest - Neighbor:-

1) Consider the table given below, which shows the rating given by 4 users for 3 different movies.

user	Movie	Rating
U1	M1	2
U1	M3	3
U2	M1	5
U2	M2	3
U3	M1	3
U3	M2	3
U3	M3	1
U4	M2	2
U4	M3	2

Predict the rating for the movies by users who haven't watched it yet. Recommend to watch them keeping threshold rating as 2.5.

A)

STEP-1 :- Create a pivot table (or matrix) showing the ratings of each user.

Movie / user	M1	M2	M3
U1	2	?	3
U2	5	3	?
U3	3	3	1
U4	?	2	2

STEP-2 :- Compute item similarities

i) Compute item similarity between M_1 & M_2 (rated by U_2 & U_3) can be given by

$$V_1 = 5U_2 + 3U_3$$
$$V_2 = 3U_2 + 3U_3$$

Cosine similarity between V_1 & V_2

$$\cos(V_1, V_2) = \frac{5 \times 3 + 3 \times 3}{\sqrt{(5^2+3^2)(3^2+3^2)}}$$

$$= 0.9701$$

ii) Compute item similarity between M_2 & M_3 (rated by U_3 & U_4) can be given as

$$V_1 = 3U_3 + 2U_4$$
$$V_2 = 1U_3 + 2U_4$$

Cosine similarity between V_1 & V_2

$$\cos(V_1, V_2) = \frac{(3 \times 1) + (2 \times 2)}{\sqrt{(3^2+2^2)(1^2+2^2)}}$$

$$= 0.8682$$

iii) Compute item similarity between M_1 & M_3 (rated by U_1, U_3) can be given as

$$V_1 = 2U_1 + 3U_3$$
$$V_2 = 3U_1 + 1U_3$$

$$\cos(V_1, V_2) = \frac{(2 \times 3) + (3 \times 1)}{\sqrt{(2^2+3^2)(3^2+1^2)}}$$

$$= 0.7894$$

STEP-3:- predict the rating of each user who haven't watched the movie yet.

$$U_1 \rightarrow M_2 \Rightarrow \text{Rating} = \frac{(2\times0.9701)+(3\times0.8682)}{0.9701+0.8682}$$

$$= 2.4723$$

$$U_2 \rightarrow M_3 \Rightarrow \text{Rating} = \frac{(5\times0.7894)+(3\times0.8682)}{0.7894+0.8682}$$

$$= 3.9525$$

$$U_4 \rightarrow M_1 \Rightarrow \text{Rating} = \frac{(2\times0.9701)+(2\times0.7894)}{0.9701+0.7894}$$

$$= 2.0$$

STEP-4:- Recommendations

- Keeping threshold on 2.5, We can recommend user-2 to watch Movie M_3.
- But, We may not recommend user-1 to watch Movie M_2 and user-4 to watch Movie-1.

→ Formulas for Association Rules:-

1) $$\text{support}(\{x\}\rightarrow\{y\}) = \frac{\text{Transactions containing both } x \text{ \& } y}{\text{Total numbers of Transactions}}$$

2) $$\text{Confidence}(\{x\}\rightarrow\{y\}) = \frac{\text{Transactions containing both } x \text{ \& } y}{\text{Transactions Containing } x}$$

3) $$\text{Lift}(\{x\}\rightarrow\{y\}) = \frac{\text{Confidence}(\{x\}\rightarrow\{y\})}{\text{support}(\{y\})}$$

→ Association Rules:-

1) Consider the transaction given in the table. Keeping min-sup = 0.6 and min-conf = 0.8, find the association rules.

Transactions	Itemlist
T_1	{F, A, D, B}
T_2	{D, A, C, E, B}
T_3	{C, A, B, E}
T_4	{B, A, D}

A)

STEP-1 :- Create 1-itemset and compute the supports

Itemset	SupCount	Support
A	4	4/4 = 1 ✓
B	4	4/4 = 1 ✓
C	2	2/4 = 0.5 ✗
D	3	3/4 = 0.75 ✓
E	2	2/4 = 0.5 ✗
F	1	1/4 = 0.25 ✗

Min-sup > 0.6

As only the item A, B & D satisfy the min-sup requirement, We use only these items for computation of 2-items.

STEP-2 :- 2-itemsets generation:-

Itemset	Support	
AB	4/4=1	✓
AD	3/4=0.75	✓
BD	3/4=0.75	✓

min-sup >0.6

As all the three 2-itemsets supports the min-sup rule, we use all items A, B & D

STEP-3 :- 3-itemsets generation

Itemset	Support	
{A, B, D}	3/4=0.75	✓

The itemset to be considered is {A, B, D}

STEP-4 :- Create the rules & check their confidence.

Possible subsets	Rules	Confidence	
{A}	{A}→{D, B}	3/4 = 0.75	
{B}	{B}→{A, D}	3/4 = 0.75	
{D}	{D}→{A, B}	3/3 = 1	✓
{A, B}	{A, B}→{D}	3/4 = 0.75	
{A, D}	{A, D}→{B}	3/3 = 1	✓
{B, D}	{B, D}→{A}	3/3 = 1	✓

min-con >0.8

The Association Rules Satisfy min-conf are highlighted.

STEP-5 :- Lift [one can check whether these rules are really making sense using lift values]

i) $\text{Lift}(D \rightarrow A,B) = \dfrac{\text{conf}(D \rightarrow A,B)}{\text{sup}(A,B)} = \dfrac{1}{1} = 1$

ii) $\text{Lift}(A,D \rightarrow B) = \dfrac{\text{conf}(A,D \rightarrow B)}{\text{sup}(B)} = \dfrac{1}{1} = 1$

iii) $\text{Lift}(B,D \rightarrow A) = \dfrac{\text{conf}(B,D \rightarrow A)}{\text{sup}(A)} = \dfrac{1}{1} = 1$