Homework 4

Hichem Belhocine 18-206-722

Task1:

Task 1

Predict the missing rating using the item based collaborative
filtering with $K = 2$ and cosine similarity.

	Michael	Paul	Ann	Julie	Pierre	Sophie
Forest Gump	3	5	5	5	4	4
Intouchables	4	5	?	4	3	5
Fight club	5	3	2	3	1	4
Lion King	3	5	5	4	3	5

the average rating:

$$R_{forest\,Gump} = \frac{3+5+5+5+4+4}{6} = \boxed{4,33}$$

$$R_{Intouchables} = \boxed{4,2}$$

$$R_{Fight\,Club} = \boxed{3}$$

$$R_{lion\,King} = \boxed{4,17}$$

Cosine Similarity:

$$S_{ij} = \frac{\sum_{k \in I}(R_{ik} \cdot R_{jk})}{\sqrt{\sum_{k \in I} R_{ik}^2} \sqrt{\sum_k R_{jk}^2}}$$

- the missing rating is in intouchables movie so we will compute
the cosine similarity for intouchables with others.

$$S(Intouchables, Forest\,Gump) = \frac{4.3+5.5+4.5+3.4+5.4}{\sqrt{4^2+5^2+4^2+3^2+5^2} \cdot \sqrt{3^2+5^2+5^2+4^2+4^2}} = \frac{89}{\sqrt{91}\sqrt{91}}$$

$$= \boxed{0,978}$$

$$S(Intouchables, Fight\,club) = \frac{4.5+5.3+4.3+3.1+5.4}{\sqrt{4^2+5^2+4^2+3^2+5^2} \cdot \sqrt{5^2+3^2+3^2+1^2+4^2}} =$$

$$= \boxed{0,947}$$

$$S(\text{Untouchables, Lion King}) = \frac{4.3 + 5.5 + 4.4 + 3.3 + 5.5}{\sqrt{91} \cdot \sqrt{3^2 + 5^2 + 4^2 + 3^2 + 5^2}} = \boxed{0.995}$$

- Using $K = 2$, Forest Gump and Lion King will the similar neighbors.

then:

$$R_{\text{Untouchables}} = \bar{R}_i + \frac{\sum_{u_j \in N_i} S_{ij}(R_{jk} - \bar{R}_j)}{\sum_{u_j \in N_i} S_{ij}}$$

$$= 4.2 + \frac{0.978(5 - 4.33) + 0.995(5 - 4.17)}{0.978 + 0.995} = \boxed{4.952}$$