

Item-Based Collaborative Filtering

Find Carol's estimated rating for "Touch"

	Steve Jobs	The Cat's Table	The Help	Touch	Sarah's Key
Anne	2	3	5	4	-
Bob	3	4	-	5	5
Carol	5	4	2	-	3
Dave	4	2	2	4	5

ITEM-BASED

(and Touch)

- Find similarity scores between each item. (Pearson Correlation Coef)

$$\text{Fmla, "Pearson" Sim}(x, y) = \frac{\sum (r_x - \bar{r}_x)(r_y - \bar{r}_y)}{\sqrt{\sum (r_x - \bar{r}_x)^2 \cdot \sum (r_y - \bar{r}_y)^2}}$$

$$\left(\begin{array}{l} \text{Touch: avg}(5, 4) = 4.5 \\ + \\ \text{Sarah's Key: avg}(5, 5) = 5 \end{array} \right.$$

$$\text{Sim}(\text{Touch, Sarah's Key}) = \frac{(5-4.5)(0) + (4-4.5)(0)}{\sqrt{0+0}} \rightarrow 0$$

$$\left(\begin{array}{l} \text{Touch: avg}(4, 4) = 4 \\ + \\ \text{The Help: avg}(5, 2) = 3.5 \end{array} \right.$$

$$\text{Sim}(T, TH) = \frac{(4-4)(5-3.5) + (4-4)(2-3.5)}{\sqrt{0+0}} \rightarrow 0$$

$$\left(\begin{array}{l} \text{Touch: avg}(4, 5, 4) = 4.3 \\ + \\ \text{Cat's Table: avg}(3, 4, 2) = 3 \end{array} \right.$$

$$\begin{aligned} \text{Sim}(T, CT) &= \frac{(4-4.3)(0) + (5-4.3)(1) + (4-4.3)(-1)}{\sqrt{[(0.3)^2 + (0.7)^2 + (0.3)^2][1+1]}} \\ &= \frac{1}{1.158} = 0.86 \end{aligned}$$

$$\left(\begin{array}{l} \text{Touch: avg}(4, 5, 4) = 4.3 \\ + \\ \text{Steve Jobs: avg}(2, 3, 4) = 3 \end{array} \right.$$

$$\begin{aligned} \text{Sim}(T, SJ) &= \frac{(4-4.3)(4) + (5-4.3)(0) + (4-4.3)(4-3)}{\sqrt{[(0.3)^2 + 0.7^2 + 0.3^2][2]}} \\ &= 0 \end{aligned}$$

Now use a Weighting Fmla to estimate Carol's rating for Touch:

$$\hat{r}_x = \frac{\sum \text{Sim}(x, y) \cdot \text{rate}_y}{\sum \text{Sim}(x, y)} = \frac{(0.86)(4)}{0.86} = 4$$

This is Carol's estimated rating for Touch.