

# Item-based collaborative filtering

Mehrdad Mohammadian

## Data

<u>Aa</u> Name	# Star wars	# Jurassic park	# Terminator 2	# Indep day
<u>sally</u>	7	6	3	7
<u>Bob</u>	7	4	4	6
<u>Chris</u>	3	7	7	2
<u>Lynn</u>	4	4	6	2
<u>Karen</u>	7	4	3	

## Pearson Similarities

<u>Aa</u> Name	# Star wars	# Jurassic park	# Terminator 2	# indep day
<u>Indep. Day</u>	0.968314	-0.105621	-0.971894	1

For **K = 2**:

Star wars: 0.96

Jurassic park: -0.10

## Prediction:

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^k (r_{ai} - \bar{r}_u) \times \text{sim}(a, u)}{\sum_{u=1}^k \text{sim}(a, u)}$$

Calculation:

$$k = 2$$

$$P_{alice} = (17/4) + \frac{(7 - 5.6) \times 0.96 + (4 - 5) \times -0.10}{0.96 - 0.10} = 5.92$$

The End