

김종윤

2023-28318

협동과정 인공지능 전공

# 소셜컴퓨팅

## 과제 8 - 13주차

### 1. User-Movie Utility Matrix

	User 1	User 2	User 3	User 4	User 5
Item 1	4	4	4	1	1
Item 2	3	1		4	
Item 3	4	2		2	3
Item 4	???	2	3		1
Item 5			1	4	3
Item 6	1	1			2

#### 1.1. Pearson Correlation Coefficient

	User 1	User 2	User 3	User 4	User 5	(*Average)
Item 1	4	4	4	1	1	2.8
Item 2	3	1		4		2.67
Item 3	4	2		2	3	2.75
Item 4		2	3		1	2
Item 5			1	4	3	2.67
Item 6	1	1			2	1.33
(*Average)	3	2	2.67	2.75	2	

[Table 1: User-Movie Table with average]

	User 1	User 2	User 3	User 4	User 5
Item 1	1	2	1.33	-1.75	-1
Item 2	0	-1		1.25	
Item 3	1	0		-0.75	1
Item 4		0	0.33		-1
Item 5			-1.67	1.25	1
Item 6	-2	-1			0
Size	2.449	2.449	2.16	2.598	2

[Table 2: User-Movie Table with average]

Let assume  $S_{xy}$  is the Pearson correlation coefficient of both user x and y.

$$S_{12} = 0.666, S_{13} = 0.251, S_{14} = -0.392, S_{15} = 0$$

$$(S_{23} = 0.503, S_{24} = -0.538, S_{25} = -0.408, S_{34} = -0.786, S_{35} = -0.771 \text{ and } S_{45} = 0.433)$$

### 1.2. User 1

The two users having high similarity with user 1 are 2 and 3.

Therefore, the rating of item 4 by user 1 can be estimated as weighted sum (using Pearson correlation coefficient as weight),  $0.33 \cdot \frac{0.251}{0.666 + 0.251} = 0.090327$ .

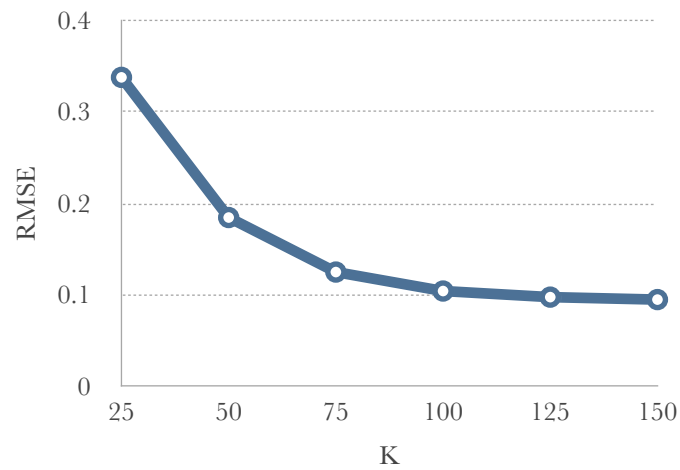
## 2. Movie Recommendation (Small Dataset)

### 2.1. UserId 10 's recommendation

The movie recommended for user 10 are given below in the descending order of the recommendation.

[ It's a Wonderful Life (1946), First Knight (1995), Postman, The (Postino, Il) (1994), Sense and Sensibility (1995), Thank You for Smoking (2006), Little Mermaid, The (1989), Sandlot, The (1993), Charlotte's Web (1973), Charlie's Angels (2000), Say Anything... (1989) ]

## 2.2. Effect of number of latent nodes ( $K$ )



**RMSE of SVD: increasing  $K$**

Increase of  $K$  led to the reduction of the error in SVD. However, after  $K=100$ , the RMSE reduction saturates around 0.1. This results that thoughtless increase of  $K$  will not reduce RMSE dramatically while it increases the computation cost.