Overview

In the first part of this exercise session, you will build a recommender system for predicting movies using different techniques. In the second part, you will explore an important concept in time series analysis: dynamic time warping.
Have fun!

1 Recommending movies using collaborative filtering

1.1 Similarity-based

A). First, you will apply the similarity-based collaborative filtering algorithm presented in class by hand to make a prediction for an unrated movie. The following table illustrates the ratings.

	User A	User B	User C
Movie 1	1		3
Movie 2	2		3
Movie 3	3		3
Movie 4	3	4	
Movie 5	5	4	
Movie 6	2	4	**?**
Movie 7		2	
Movie 8		4	2
Movie 9		4	5
Movie 10			4

Use the collaborative filtering technique with Pearson correlation to predict the rating of user C for movie 6.

Let's compute first the similarity between User C and, respectively, User A and User B:

$$w(C,A) = \frac{(3-3.33)\cdot(1-2.67)+(3-3.33)\cdot(2-2.67)+(3-3.33)\cdot(3-2.67)}{\sqrt{((3-3.33)^2+(3-3.33)^2+(3-3.33)^2)\cdot((1-2.67)^2+(2-2.67)^2+(3-2.67)^2)}} = 0.63;$$
$$w(C,B) = \frac{(2-3.33)\cdot(4-3.67)+(5-3.33)\cdot(4-3.67)}{\sqrt{((2-3.33)^2+(5-3.33)^2)\cdot((4-3.67)^2+(4-3.67)^2)}} = 0.11.$$

Then, we can make the prediction using Equation 1:

$$\hat{R}_{C,6} = 3.33 + \frac{1}{|0.63+0.11|}\cdot[0.63\cdot(2-2.67)+0.11\cdot(4-3.67)] = 2.81.$$

B). Now download from Toledo the `collaborative-filtering.ipynb` notebook and the data `u.data` which contains 100,000 movie ratings by different users. The goal is to build a recommendation engine that implements the functions above.

1.2 Model-based

Another approach to collaborative filtering is to use a model-based method such as singular value decomposition or non-negative matrix factorization. In the `collaborative-filtering.ipynb` notebook, you will use an implementation of non-negative matrix factorization to predict ratings.

2 Dynamic Time Warping

For the time series $s_1 = [1,1,1,3,2,0]$ and $s_2 = [2,4,3,2,1,1]$, do the following:

1. Compute the Euclidean distance.

$$\sqrt{(1-2)^2+(1-4)^2+(1-3)^2+(3-2)^2+(2-1)^2+(0-1)^2} = 4.12$$

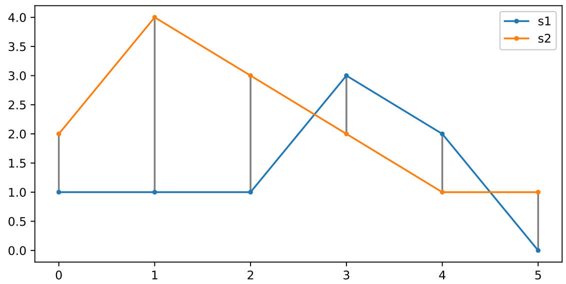
2. Compute the DTW distance.

	1	**1**	**1**	**3**	**2**	**0**
2	1	2	3	4	4	6
4	4	4	5	4	6	8
3	6	6	6	4	5	8
2	7	7	7	5	4	6
1	7	7	7	7	5	5
1	7	7	7	9	6	6

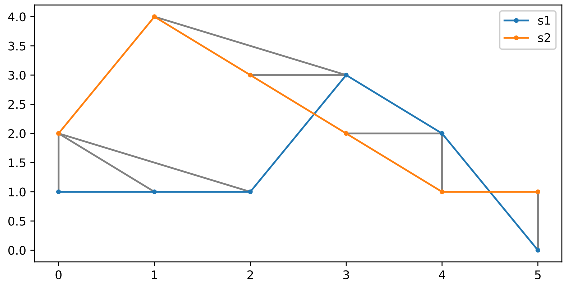
DTW distance = lower right corner = 6

3. Draw a plot of s_1 and s_2 and show visually which points of s_1 are compared to which points of s_2, for both the Euclidean distance and the DTW distance.

For Euclidean distance:



For DTW distance:



4. Compute the DTW distance, but now use a warping constraint of 1.

	1	**1**	**1**	**3**	**2**	**0**
2	1	2				
4	4	4	5			
3		6	6	5		
2			7	6	5	
1				8	6	6
1					7	7

DTW distance = lower right corner = 7