5731095821 Wasuwat Satittamajittra

2110430 Time Series Mining and Knowledge Discovery

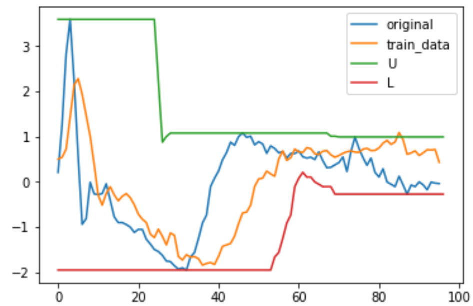
Final Project

1. **Dynamic Time Warping (DTW) modifications**

   In this project, I implemented the DTW and tested for the 1-nearest-neighbor (1NN) classification accuracies.

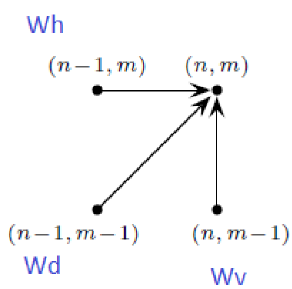
   The data are normalized before classified.

   I used the Sakoe-Chiba band and set the window length to (length of samples / 10)

   Keogh's lower bound was also used.

   Weights(Wv, Wd, Wh) = (vertical,diag,horizontal)

   L2 norm distances were used, unless specified.

1-A. **Asymmetric Weights**

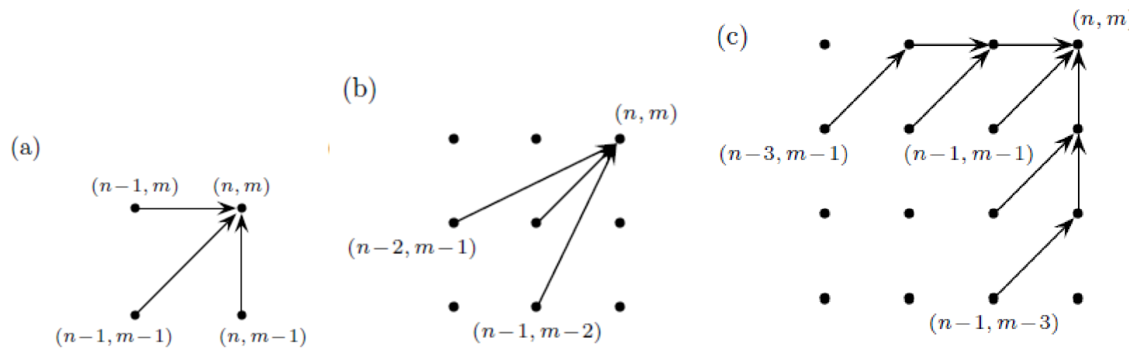
(Wv, Wd, Wh)/Instances	ECG200	Coffee	Beef	Beef(L1)	Gun_Point	OliveOil	Lightning7
(1,1,1)	1.0	1.0	0.5666	0.6666	1.0	1.0	0.808219
(1,1,2)	0.989	1.0	0.5666	0.6666	1.0	1.0	0.8356
(1,2,1)	1.0	0.96428	0.5666	0.6666	1.0	1.0	0.83561
(2,1,1)	0.979	1.0	0.533	0.7000	1.0	1.0	0.8630

Table1 – The classification accuracies.

There is no optimal weights for the classification accuracies in general. It depends on the instances.

1-B. **Another step patterns for DTW**

I tested on the following 3 patterns, the weights are symmetric (1,1,1).



Patterns/Instances	ECG200	Coffee	Beef	Gun_Point	OliveOil	Lightning7
Pattern (a)	1.0	1.0	0.5666	1.0	1.0	0.808219
Pattern (b)	0.89000	1.0	0.500	1.0	1.0	0.8904
Pattern (c)	0.920000	0.964285	0.5333	1.0	1.0	0.917808

Table2 – The classification accuracies.

Most instances work best on the pattern (a), except for Lightning7 that works best on Pattern (c).

2. Shape Averaging

I found an interesting research paper [1] about this topic. They proposed an algorithm called DTW Barycenter Averaging (DBA) method. The idea is that they want to minimize the sum of squared DTW distances from the average sequence to the set of sequences. DTW between each individual sequence and the temporary average sequence will be computed to find the associated coordinates between them. Then, each coordinate of the average sequence will be updated and used as the barycenter of coordinates associated to it. Here is the pseudocode of the algorithm.

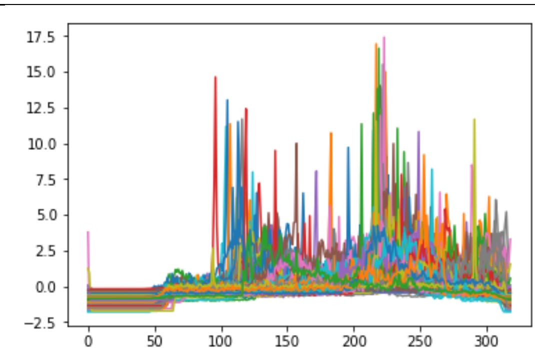
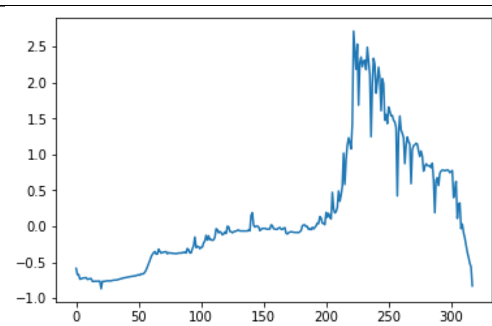
Algorithm 5. DBA

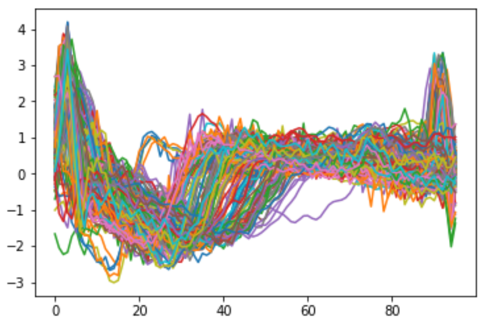
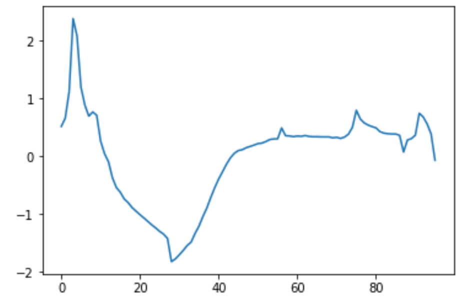
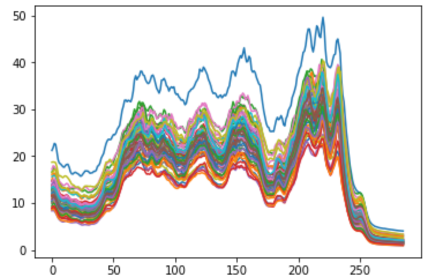
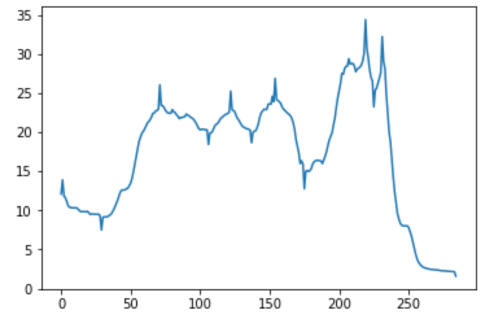
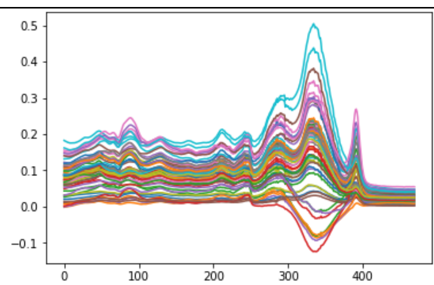
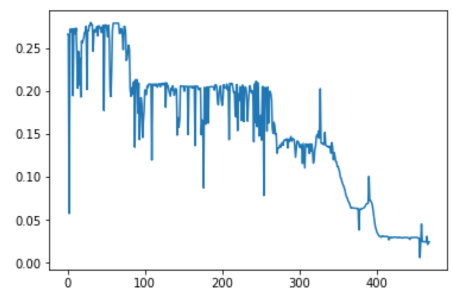
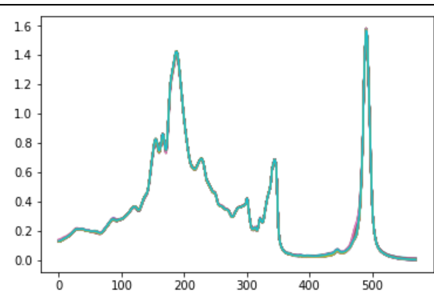
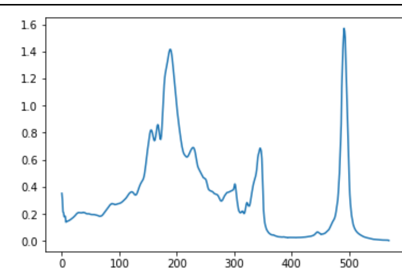
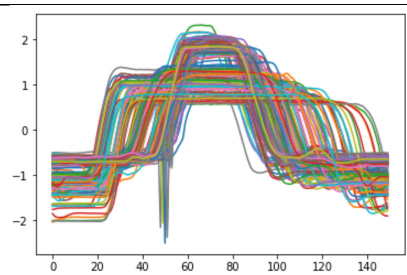
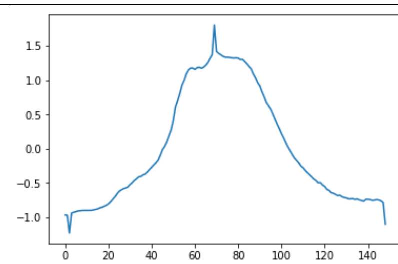
Require: $\mathcal{C} = \langle \mathcal{C}_1, \ldots, \mathcal{C}_{T'} \rangle$ the initial average sequence
Require: $\mathbb{S}_1 = \langle s_{1_1}, \ldots, s_{1_T} \rangle$ the 1st sequence to average
$\vdots$
Require: $\mathbb{S}_n = \langle s_{n_1}, \ldots, s_{n_T} \rangle$ the nth sequence to average
Let T be the length of sequences
Let $assocTab$ be a table of size T' containing in each cell a set of coordinates associated to each coordinate of $\mathcal{C}$
Let $m[T,T]$ be a temporary DTW (cost,path) matrix
  $assocTab \leftarrow [\emptyset, \ldots, \emptyset]$
  for seq in $\mathbb{S}$ **do**
    $m \leftarrow DTW(\mathcal{C}, seq)$
    $i \leftarrow T'$
    $j \leftarrow T$
    while $i \geq 1$ and $j \geq 1$ **do**
      $assocTab[i] \leftarrow assocTab[i] \cup seq_j$
      $(i,j) \leftarrow second(m[i,j])$
    end while
  end for
  for $i = 1$ to T **do**
    $\mathcal{C}_i' = barycenter(assocTab[i])$ {see Eq. (6)}
  end for
  return $\mathcal{C}'$

$$barycenter\{X_1, \ldots, X_\alpha\} = \frac{X_1 + \cdots + X_\alpha}{\alpha}$$

I implemented the DBA and tested on the datasets. I also shuffled the data and went through the algorithm again to see if they will give the same results regardless of the orders, and they did.

The results are shown below.

Instances Test+Train	**Original data**	**Averaged Data**
Lightning7		

ECG200		
Coffee		
Beef		
OliveOil		
Gun_Point		

Reference(s):

[1]Francois Petitjean, AlainKetterlin, PierreGanc-arski, A global averaging method for dynamic time warping, with applications to clustering, Pattern Recognition44, 678–693, 2001.