

Data Mining 2

CAT 3 - 2019/2020

Name _____ Surname _____ ID: _____ Test id. AUTO

Q1. Given $w = 4$ and the time series $X = \langle 20, 15, 7, 7, 22, 29, 19 \rangle$, apply the Moving Average Smoothing. What type of distortion the smoothing reduce?

A1. _____

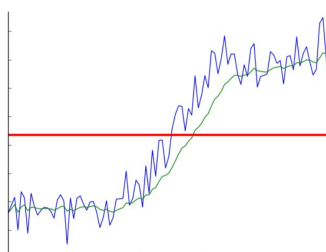
Q2. Given the time series $X_1 = \langle 4, 3, 2, 4, 1 \rangle$ and $X_2 = \langle 2, 3, 4, 6, 3 \rangle$, compute their distance using the DTW with distance between points computed as $d(x, y) = |x - y|$

Point-to-point costs					Cumulative costs				

Q3. Given the time series $X = \langle 2, 1, 7, 1, 2, 1, 2 \rangle$, build the Matrix Profile with $m = 3$ using the Manhattan distance. Which is/are correct value/s for m that would have retrieved more motifs with distance equals to 0?

A3. _____

Q4. Mark the correct sentence/s about the **time series**, the **mean** and the **exponentially weighted mean** below, knowing that a Dickey-Fuller Test gives a Statistics of -1.42 and a Critical Value of -9.7



- 1) A weakly stationary behaviour is highlighted by the **mean**
- 2) The Dickey-Fuller Test reports not-stationarity
- 3) The Dickey-Fuller Test reports seasonality
- 4) Oldest points are more important in the calculus of the **weighted mean**
- 5) The **mean** is only representative of the data around the middle

A4. _____

N.B.: this question can have more than one correct answer

Q5. In the Holt forecast equation $\hat{y}_{t+h|T} = l_t + hb_t$ the trend equation is equal to ...

- 1) $l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$
- 2) h
- 3) $l_t = \alpha y_t + (1 - \alpha)l_{t-1}$
- 4) $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$
- 5) $b_t = \beta(y_t - y_{t-1}) + (1 - \beta)b_{t-1}$

A5. _____

Q6. Identify the CNN/s which do/es not preserve space and write down how many zero-pad pixels are needed to adjust its size (e.g. "5, 2 pixels").

- 1) input 9x9 filter 5x5 stride 3
- 2) input 7x7 filter 3x3 stride 1
- 3) input 7x7 filter 5x5 stride 1
- 4) input 5x5 filter 3x3 stride 2
- 5) input 5x5 filter 3x3 stride 3

A6. _____

N.B.: this question can have more than one correct answer

Q7. LSTMs ...

- 1) ... are a particular type of CNN
- 2) ... encode memory as convolutional layers
- 3) ... maintain information in memory for short periods of time
- 4) ... need the independence of data
- 5) ... improve backpropagation flow

A7. _____