Assignment 3 (ML for TS) - MVA 2023/2024

Rania Bennani rania.bennani@ens-paris-saclay.fr
Raphael Razafindralambo raphael.razafin@gmail.com

January 14, 2024

1 Introduction

Objective. The goal is to implement (i) a signal processing pipeline with a change-point detection method and (ii) wavelets for graph signals.

Warning and advice.

- Use code from the tutorials as well as from other sources. Do not code yourself well-known procedures (e.g. cross validation or k-means), use an existing implementation.
- The associated notebook contains some hints and several helper functions.
- Be concise. Answers are not expected to be longer than a few sentences (omitting calculations).

Instructions.

- Fill in your names and emails at the top of the document.
- Hand in your report (one per pair of students) by Sunday 31st December 11:59 PM.
- Rename your report and notebook as follows:
  `FirstnameLastname1_FirstnameLastname1.pdf` and
  `FirstnameLastname2_FirstnameLastname2.ipynb`.
  For instance, `LaurentOudre_CharlesTruong.pdf`.
- Upload your report (PDF file) and notebook (IPYNB file) using this link: docs.google.com/forms/d/e/1FAIpQLScqLsYuKeQbsDEOie5OqpOH7YwCnWmudzApMC005HvxOaOv

2 Dual-tone multi-frequency signaling (DTMF)

Dual-tone multi-frequency signaling is a procedure to encode symbols using an audio signal. The possible symbols are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, *, #, A, B, C, and D. A symbol is represented by a sum of cosine waves: for $t = 0, 1, \ldots, T-1$,

$$y_t = \cos(2\pi f_1 t / f_s) + \cos(2\pi f_2 t / f_s)$$

where each combination of (f_1, f_2) represents a symbols. The first frequency has four different levels (low frequencies), and the second frequency has four other levels (high frequencies); there are 16 possible combinations. In the notebook, you can find an example symbol sequence encoded with sound and corrupted by noise (white noise and a distorted sound).

Question 1

Design a procedure that takes a sound signal as input and outputs the sequence of symbols. To that end, you can use the provided training set. The signals have a varying number of symbols with a varying duration. There is a brief silence between each symbol.

Describe in 5 to 10 lines your methodology and the calibration procedure (give the hyperparameter values). Hint: use the time-frequency representation of the signals, apply a change-point detection algorithm to find the starts and ends of the symbols and silences, and then classify each segment.

Answer 1

In our approach to decoding Dual-tone multi-frequency signaling (DTMF) from sound signals, we undertake the following steps:

- **Preprocessing with Short-Time Fourier Transform (STFT):** We initiate the process by converting the time-domain signal into the frequency domain using the Short-Time Fourier Transform (STFT). This step allows us to analyze the signal in terms of its frequency components over short time intervals. The hyperparameter 'nperseg' is dynamically chosen in the code, representing the number of data points used in each segment for STFT computation. It is set to a minimum of 256 or adjusted based on the signal's characteristics for effective detection (i.e fft)
- **Change-Point Detection Algorithm:** By applying a dynamic programming-based change-point detection algorithm, we identify the starting and ending points of each symbol and the intervening silences.
- **Segmentation:** Built on the detected change points, we segment the spectrogram to obtain individual segments corresponding to symbols and silences. This step is crucial for subsequent classification.
- **Classification:** Each segment is directly used as input of the Convolutional Neural Network (CNN) classifier. Before this classification, we preprocess the data by padding sequences to a maximum length, ensuring a consistent input size for the neural network.
- **Decoding:** By applying the trained classifier to each segment of the test dataset, we reconstruct the sequence of symbols based on the classification results.

Question 2

What are the two symbolic sequences encoded in the test set?

Answer 2

- Sequence 1: 721C99
- Sequence 2: 1#2#

Edit: These sequences are the ones that had to be found, but our code did not predict the exact same sequences. Despite the fact that our model achieved high accuracies, there might be an overfitting issue or an error in the choice of our hyperparameters.

3 Wavelet transform for graph signals

Let G be a graph defined a set of n nodes V and a set of edges E. A specific node is denoted by v and a specific edge, by e. The eigenvalues and eigenvectors of the graph Laplacian L are $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $u_1, u_2, \ldots, u_n$ respectively.

For a signal $f \in \mathbb{R}^n$, the Graph Wavelet Transform (GWT) of f is $W_f : \{1, \ldots, M\} \times V \longrightarrow \mathbb{R}$:

$$W_f(m, v) := \sum_{l=1}^{n} \hat{g}_m(\lambda_l) \hat{f}_l u_l(v) \tag{1}$$

where $\hat{f} = [\hat{f}_1, \ldots, \hat{f}_n]$ is the Fourier transform of f and $\hat{g}_m$ are M kernel functions. The number M of scales is a user-defined parameter and is set to $M := 9$ in the following. Several designs are available for the $\hat{g}_m$; here, we use the Spectrum Adapted Graph Wavelets (SAGW). Formally, each kernel $\hat{g}_m$ is such that

$$\hat{g}_m(\lambda) := \hat{g}^U(\lambda - am) \quad (0 \leq \lambda \leq \lambda_n) \tag{2}$$

where $a := \lambda_n / (M + 1 - R)$,

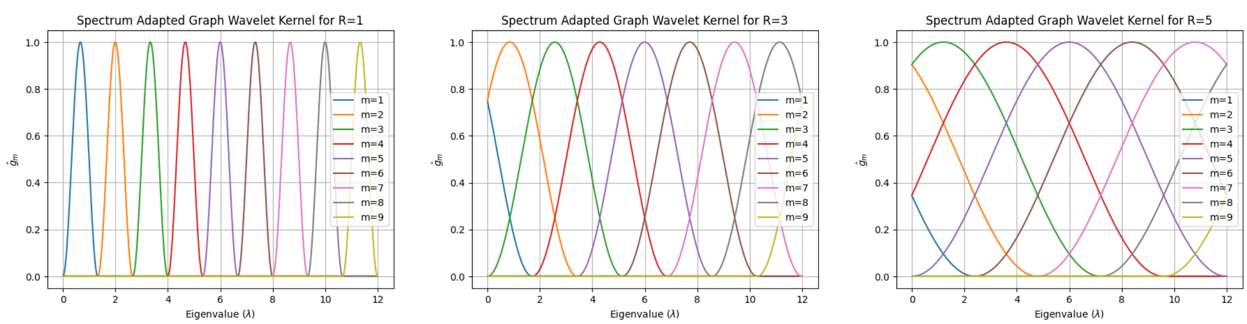
$$\hat{g}^U(\lambda) := \frac{1}{2} \left[1 + \cos\left(2\pi \left(\frac{\lambda}{aR} + \frac{1}{2} \right) \right) \right] \mathbb{1}(-Ra \leq \lambda < 0) \tag{3}$$

and $R > 0$ is defined by the user.

Question 3

Plot the kernel functions $\hat{g}_m$ for $R = 1$, $R = 3$ and $R = 5$ (take $\lambda_n = 12$) on Figure 1. What is the influence of R?

Answer 3



(a) $R = 1$ (b) $R = 3$ (c) $R = 5$

Figure 1: The SAGW kernels functions

Regarding the figures, we can see that: when R increases, the Spectrum Adapted Graph Wavelet (SAGW) kernels exhibit a more oscillatory behavior with a narrower frequency concentration. Specifically, as the parameter R becomes larger, the support of the cosine term in the kernel function $\hat{g}^U$ decreases, resulting in a higher frequency of oscillations within the kernel.

We will study the Molene data set (the one we used in the last tutorial). The signal is the temperature.

Question 4

Construct the graph using the distance matrix and exponential smoothing (use the median heuristics for the bandwidth parameter).

- Remove all stations with missing values in the temperature.
- Choose the minimum threshold so that the network is connected and the average degree is at least 3.
- What is the time where the signal is the least smooth?
- What is the time where the signal is the smoothest?

Answer 4

The stations with missing values are ARZAL, BATZ, BEG MEIL, BREST-GUIPAVAS, BRIGNOGAN, CAMARET, LANDIVISIAU, LANNAERO, LANVEOC, OUESSANT-STIFF, PLOUAY-SA, PLOUDALMEZEAU, PLOUGONVELIN, QUIMPER, RIEC SUR BELON, SIZUN, ST NAZAIRE-MONTOIR, and VANNES-MEUCON.

The threshold is equal to 0.8319.

The signal is the least smooth at 2014-01-10 09:00:00.

The signal is the smoothest at 2014-01-24 19:00:00.

Question 5

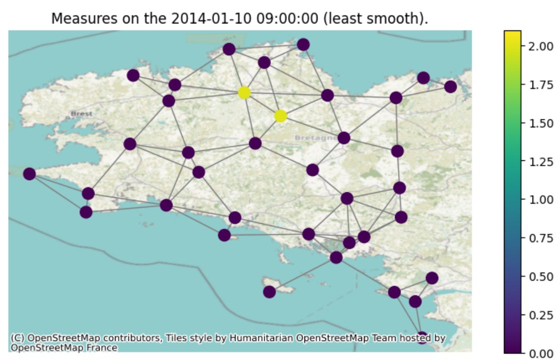
(For the remainder, set $R = 3$ for all wavelet transforms.)

For each node v, the vector $[W_f(1, v), W_f(2, v), \dots, W_f(M, v)]$ can be used as a vector of features. We can for instance classify nodes into low/medium/high frequency:

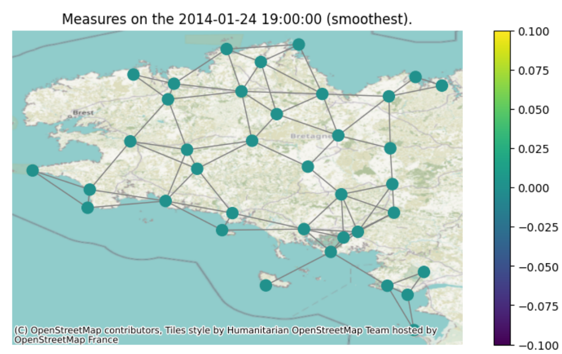
- a node is considered low frequency if the scales $m \in \{1, 2, 3\}$ contain most of the energy,
- a node is considered medium frequency if the scales $m \in \{4, 5, 6\}$ contain most of the energy,
- a node is considered high frequency if the scales $m \in \{6, 7, 9\}$ contain most of the energy.

For both signals from the previous question (smoothest and least smooth) as well as the first available timestamp, apply this procedure and display on the map the result (one colour per class).

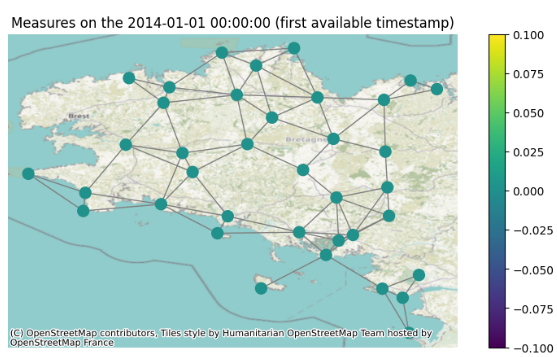
Answer 5



(a) Least smooth signal



(b) Smoothest signal



(c) First available timestamp

Figure 2: Classification of nodes into low/medium/high frequency

Question 6

Display the average temperature and for each timestamp, adapt the marker colour to the majority class present in the graph (see notebook for more details).

Answer 6

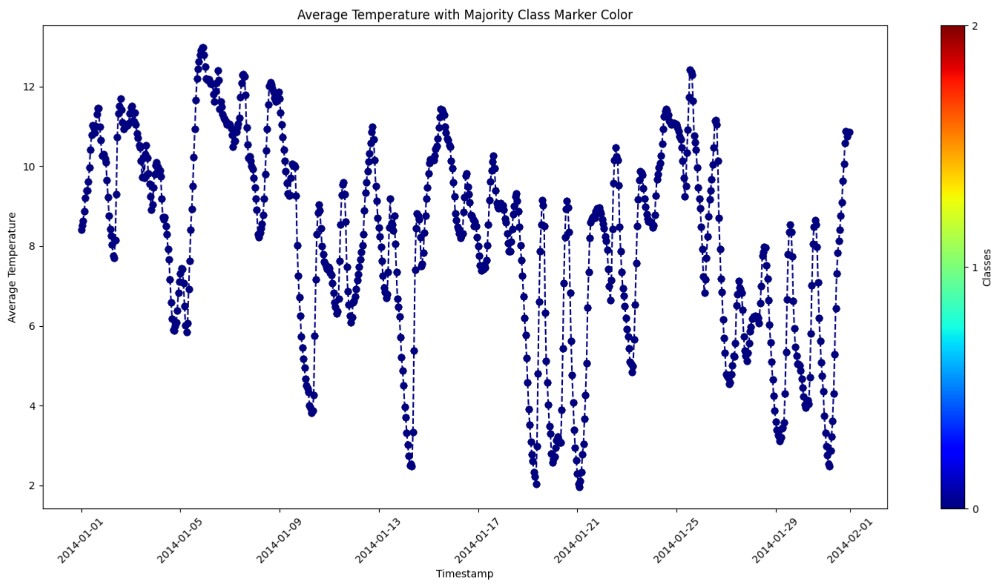


Figure 3: Average temperature. Markers' colours depend on the majority class.

Question 7

The previous graph G only uses spatial information. To take into account the temporal dynamic, we construct a larger graph H as follows: a node is now *a station at a particular time* and is connected to neighbouring stations (with respect to G) and to itself at the previous timestamp and the following timestamp. Notice that the new spatio-temporal graph H is the Cartesian product of the spatial graph G and the temporal graph G' (which is simply a line graph, without loop).

- Express the Laplacian of H using the Laplacian of G and G' (use Kronecker products).
- Express the eigenvalues and eigenvectors of the Laplacian of H using the eigenvalues and eigenvectors of the Laplacian of G and G'.
- Compute the wavelet transform of the temperature signal.
- Classify nodes into low/medium/high frequency and display the same figure as in the previous question.

Answer 7

- Let $L_G \in \mathcal{M}_n$ and $L_{G'} \in \mathcal{M}_m$ be the Laplacian matrices of G (n nodes) and G' (m nodes) respectively. H is the Cartesian product of G and G' ($H = G \square G'$), thus the Laplacian L_H of H is given by
$$L_H = L_G \otimes I_m + I_n \otimes L_{G'}$$
where $\otimes$ denotes the Kroenecker product.

- Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of G and $\mu_1, \dots, \mu_m$ be the eigenvalues of G'. We denote $x_1, \dots, x_n$ and $y_1, \dots, y_m$ the corresponding eigenvectors. Then for all $i, j \in \{1, \dots, n\} \times \{1, \dots, m\}$
$$\begin{aligned} L_H(x_i \otimes y_j) &= (L_G \otimes I_m + I_n \otimes L_{G'})(x_i \otimes y_j) \\ &= (L_G \otimes I_m)(x_i \otimes y_j) + (I_n \otimes L_{G'})(x_i \otimes y_j) \end{aligned}$$

  By mixed-product property
$$\begin{aligned} &= (L_G x_i) \otimes (I_m y_j) + (I_n x_i) \otimes (L_{G'} y_j) \\ &= (\lambda_i x_i) \otimes y_j + x_i \otimes (\mu_j y_j) \\ &= \lambda_i (x_i \otimes y_j) + \mu_j (x_i \otimes y_j) \\ &= (\lambda_i + \mu_j)(x_i \otimes y_j) \end{aligned}$$

  Thus for all $i, j \in \{1, \dots, n\} \times \{1, \dots, m\}$, $(x_i \otimes y_j)$ are eigenvectors of H associated with eigenvalues $(\lambda_i + \mu_j)$.

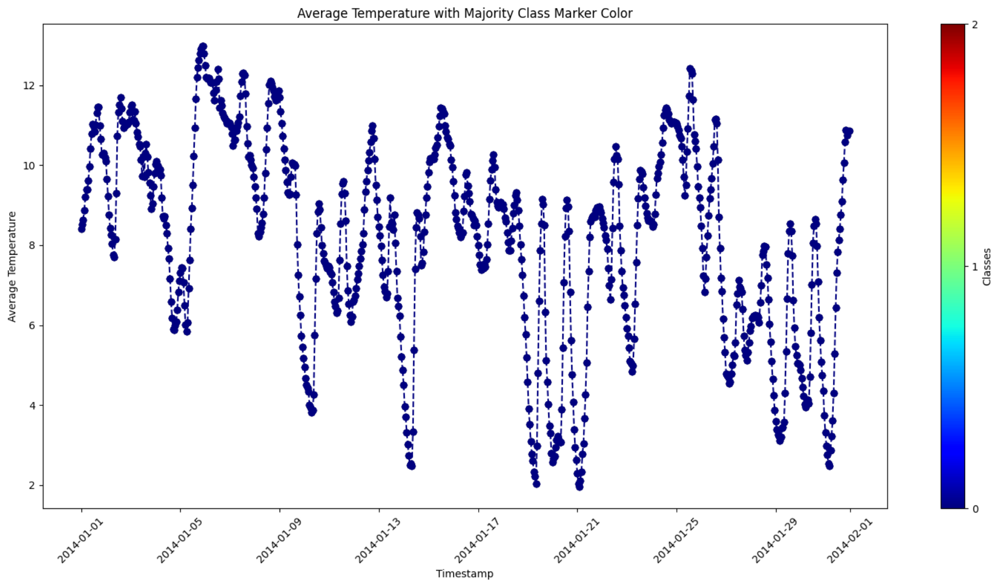


Figure 4: Average temperature. Markers' colours depend on the majority class.