Assignment 3 (ML for TS) - MVA

Pierre Aguié pierre.aguie@gmail.com
Paco Goze paco.goze@polytechnique.edu

December 30, 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 one report per pair of students.
- 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 the link given in the email.

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

We analyzed the frequency-domain representation of the signals using their spectrograms and implemented the following pipeline:

1. **STFT Calculation:**
   For each signal, we computed its STFT using $n_{\text{fft}} = 2048$ and hop_length $= 256$. Intensities below a manually fine-tuned threshold (threshold = 44) were set to zero to refine the spectrogram.

2. **Identifying Frequencies of Interest:**
   We aimed to identify the eight key frequencies associated with DTMF signaling (F_1 and F_2 sets). To achieve this, we computed the overall energy for each frequency across the training set and identified the eight most intense frequencies, presumed to correspond to the DTMF keypad. The frequencies in F_1 and F_2 were closely aligned with the known DTMF values (see the notebook for comparisons). Ultimately, we used the standard DTMF frequencies for further analysis.

3. **Detecting Symbol Boundaries:**
   For each signal, we determined the start and end times of symbols. This was done by calculating E_F, the energy of the STFT at the selected frequencies ($F_1 \cup F_2$) over time. A manually tuned threshold was applied to E_F; symbols were detected at time t if $E_F(t) >$ threshold (threshold = 95). This threshold was optimized using the training set to maximize prediction accuracy.

4. **Symbol Identification:**
   Using the detected symbol boundaries, we determined the symbol at each time t. This in-

volved identifying the two most intense frequencies, (f_1, f_2), at time t and finding the closest pair in $F_1 \times F_2$ using Euclidean distance.

This approach delivered impressive performance, achieving a precision of 1.0 on the training set.

Question 2

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

Answer 2

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

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

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

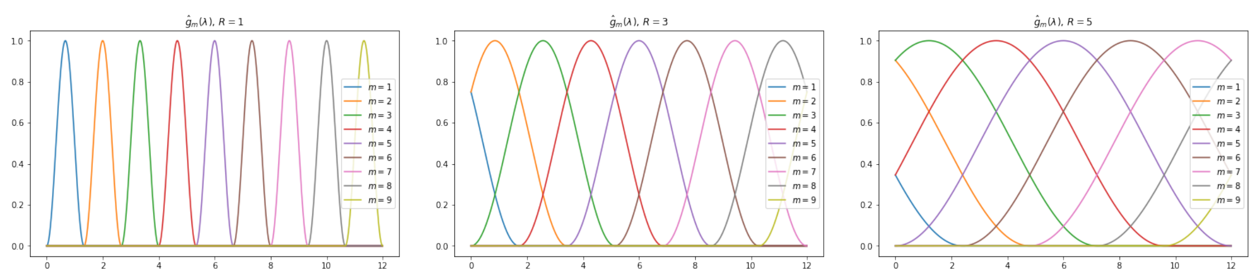
and $R > 0$ is defined by the user.

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

R changes the silhouette of the kernels. At fixed m, the higher R, the wider its "bump" becomes, and the more its maximum is shifted to the left. Thus, for higher values of R, the kernel function $\hat{g}_m$ covers a wider band of the frequency spectrum (the eigenvalues of L).



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

Figure 1: The SAGW kernels functions

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', 'VANNES-MEUCON'.

The threshold is equal to 0.82.

The signal is the least smooth at 2014-01-21 02:00:00.

The signal is the smoothest at 2014-01-22 03:00:00.

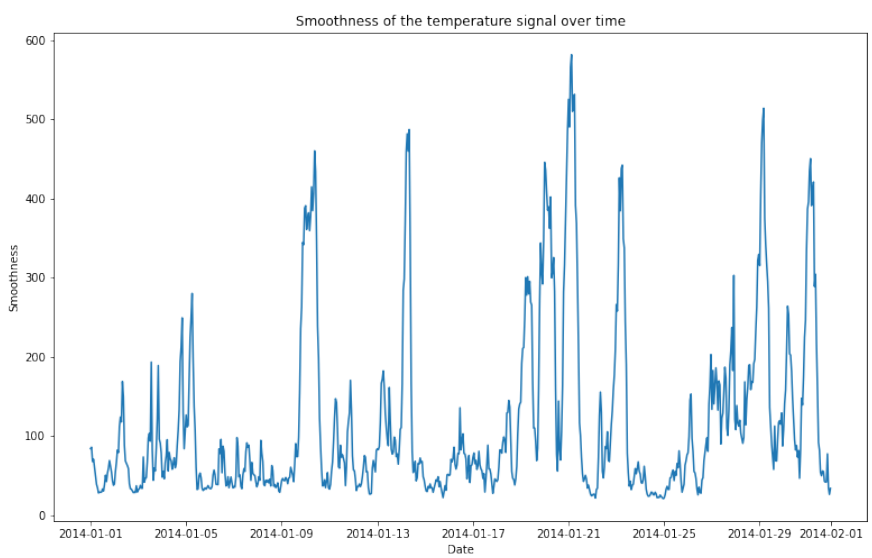


Figure 2: Smoothness of the temperature signal at each timestep

Question 5

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

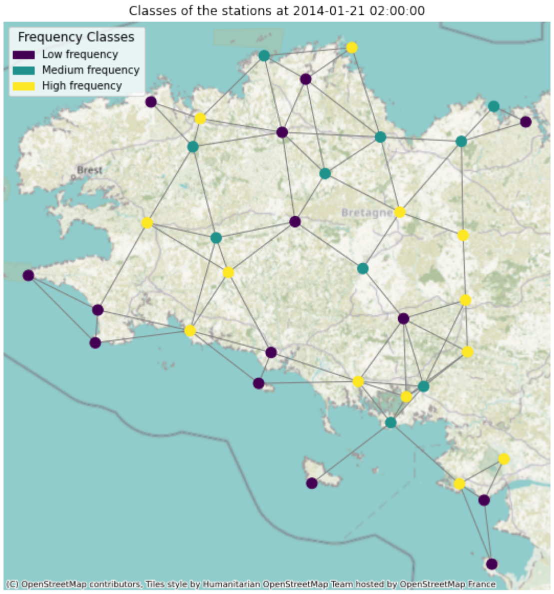
For each node v, the vector $[W_f(1, v), W_f(2, v), \dots, W_f(M, v)]$ can be used as a vector of features. We can for instance classify nodes into low/medium/high frequency:

- 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.

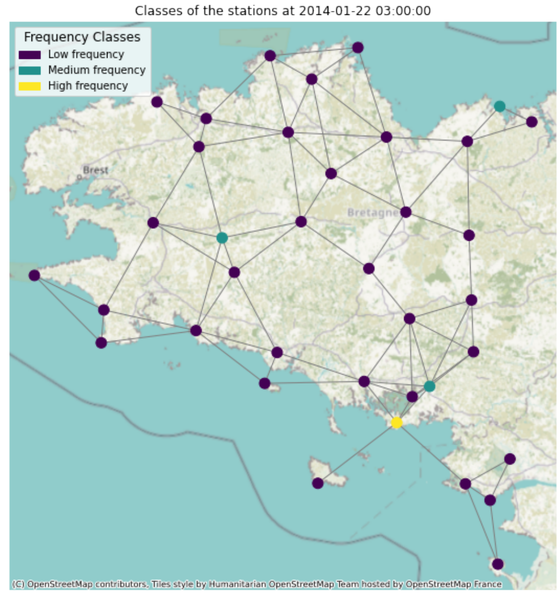
For both signals from the previous question (smoothest and least smooth) as well as the first available timestamp, apply this procedure and display on the map the result (one colour per class).

Answer 5

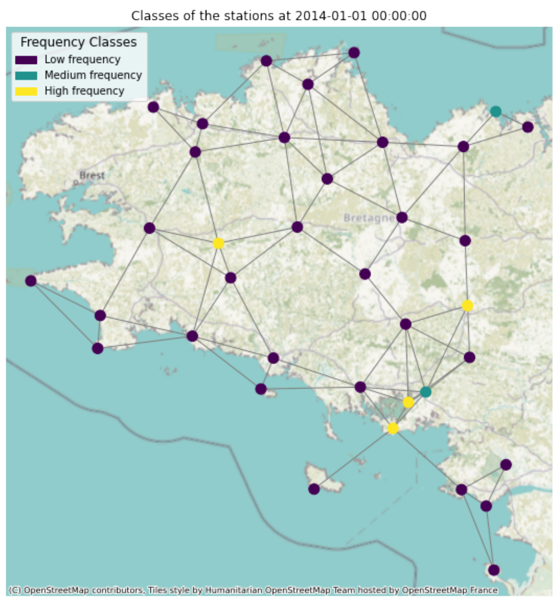
As shown in Figure 3, the least smooth signal has many nodes whose energy is mainly contained in the high and medium frequencies, while the smoothest signal has almost only nodes whose energy is contained in the low frequencies. The signal at the first timestep also has mostly low-frequency nodes but has more high-frequency nodes than the smoothest signal. This is the result that we expected.



(a) Least smooth signal



(b) Smoothest signal



(c) First available timestamp

Figure 3: 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

It seems that at each timestep, the majority class is the low frequency nodes, even for the higher values of smoothness (i.e. the least smooth signals). Note however that for some graphs, the low frequency nodes and high frequency nodes were very close in representation (e.g. 14 vs. 13 for the least smooth signal).

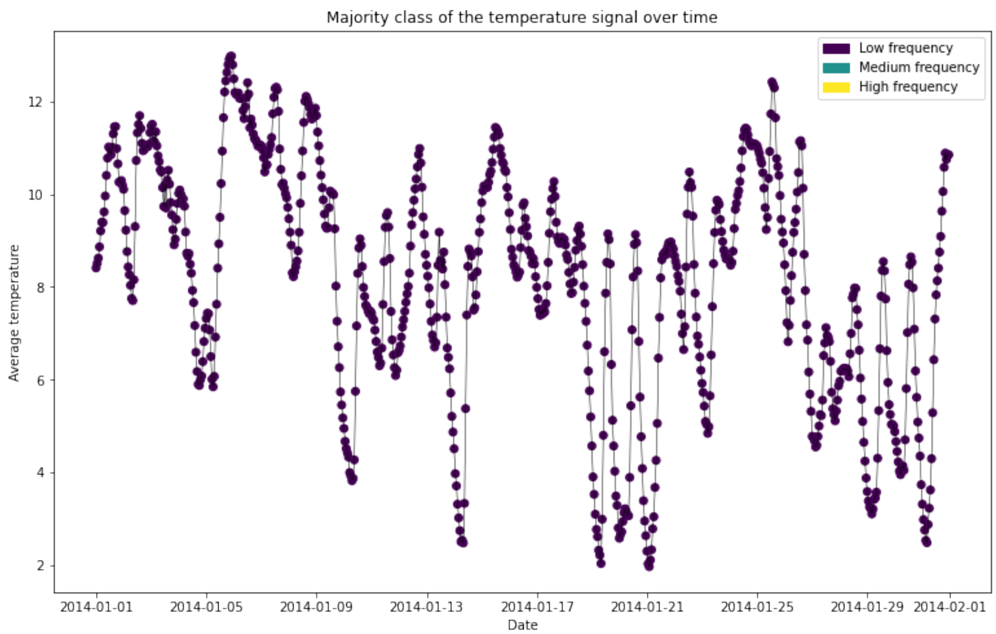


Figure 4: Average temperature. Markers' colours depend on the majority class.

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

In the following, let T be the number of time steps in our time series (i.e. the number of nodes in G') and n the number of stations (i.e. the number of nodes in G). We assume that a node representing station n at time step t is connected with a weight of 1 to the nodes representing stations n at time steps $t-1$ and $t+1$ (if these time steps exist). This gives us the following expressions for the degree matrix, the adjacency matrix and the Laplacian of G':

$$
\begin{aligned}
D_{G'} &= \mathrm{Diag}(1, 2, \ldots, 2, 1) \in \mathbb{R}^{T \times T}, \\
A_{G'} &= J_0 + J_0^\top \in \mathbb{R}^{T \times T}, \\
L_{G'} &= D_{G'} - A_{G'},
\end{aligned}
$$

where J_0 is the Jordan block of size T, filled with 0s except on its superdiagonal, which is filled with 1s.

Using $D_G, D_{G'}, A_G, A_{G'}$, we deduce an expression of L_H, using Kronecker products:

$$
\begin{aligned}
D_H &= I_T \otimes L_G + L_{G'} \otimes I_n, \\
A_H &= I_T \otimes A_G + A_{G'} \otimes I_n, \\
L_H &= I_T \otimes L_G + L_{G'} \otimes I_n.
\end{aligned}
$$

Let u_G (resp. $u_{G'}$) be an eigenvector of L_G (resp. $L_{G'}$) of associated eigenvalue λ_G (resp. $\lambda_{G'}$). We find that $u_{G'} \otimes u_G$ is an eigenvector of L_H of associated eigenvalue $\lambda_G + \lambda_{G'}$, and that all eigenvectors and eigenvalues of L_H can be expressed this way.

The majority classes for this new signal at each time step are shown in Figure 5. As the smoothness metric now accounts for variations of temperature in time, we see that for some time steps at the end of the signal, the quick variations in time seem to increase the presence of energy in the high frequencies of the signal.

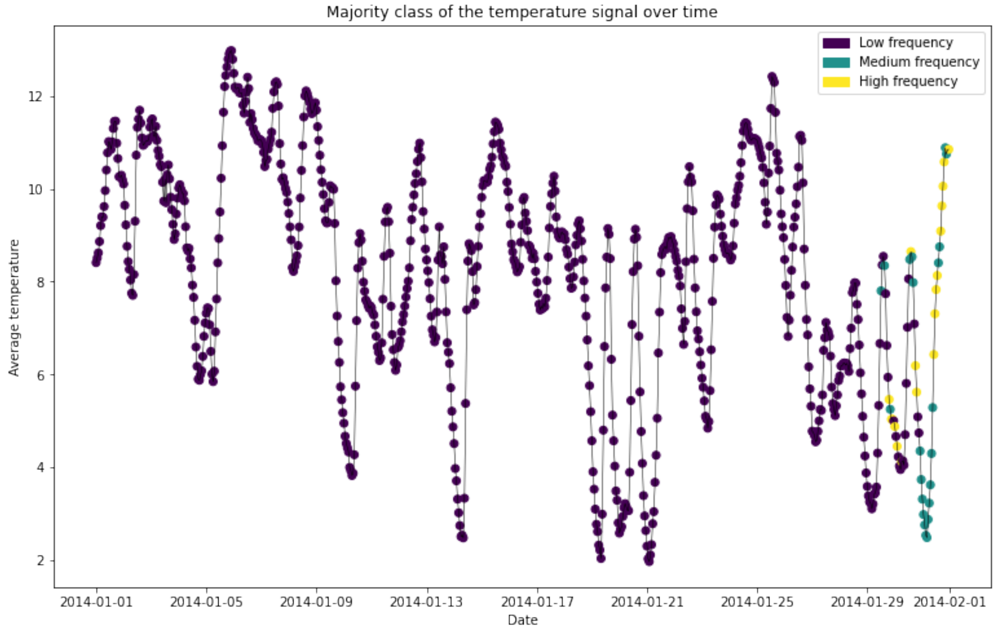


Figure 5: Average temperature. Markers' colours depend on the majority class.