Assignment 3 (ML for TS) - MVA

Jordi Baroja Fernàndez baroja.jordi@gmail.com
Janis Aiad janisaiad.JA@gmail.com

December 20, 2025

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_ValerioGuerrini.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 apply changepoint detection to the time-frequency representation of the signal to identify symbol boundaries and classify each segment. The complete procedure consists of the following steps:

Step 1: Time-Frequency Representation. We compute the Short-Time Fourier Transform (STFT) with window length 256 samples and overlap 128 samples (sampling frequency $f_s = 22050$ Hz). For each frequency bin, we compute the energy by integrating the squared magnitude over the time axis: $E(f) = \sum_t |Z(f,t)|^2$.

Step 2: Signal-Noise Separation. Since the signal contains significant white noise, we perform k-means clustering ($k = 2$) on the energy values to separate frequency bins with high energy (signal) from those with low energy (noise). We retain only the high-energy bins for further processing.

Step 3: Frequency Clustering. Within the high-energy frequency bins, multiple bins may correspond to the same DTMF frequency due to spectral leakage. We perform a second k-means clustering ($k = 8$) on the 2D feature space $(f, E(f))$ combining frequency and energy to group bins into the 8 major DTMF frequency bands (4 low frequencies: 697, 770, 852, 941 Hz; 4 high frequencies: 1209, 1336, 1477, 1633 Hz). Here, $k = 8$ is the *only* piece of prior knowledge we inject from the DTMF standard; all remaining steps are fully data-driven. We filter out clusters containing only a single frequency bin or frequencies outside the DTMF range $[500, 1800]$ Hz.

Step 4: Changepoint Detection. For each frequency cluster, we compute the energy signal $E(t)$ for each frequency bin. Since DTMF signals are dual-tone (simultaneous low and high frequencies), we apply changepoint detection (PELT algorithm) independently on each frequency bin and select, for each cluster, the bin with the minimum number of changepoints as the most stable representative. The penalty is chosen according to the BIC rule $\beta = 2\sigma^2 \log(T)$, where σ^2 is estimated from the variance of $E(t)$ and T is the signal length.

Step 5: Period Extraction and Filtering. For each representative frequency bin, we consider the segments between consecutive changepoints and compute, for each segment, its partial energy and normalize it by the segment duration to obtain an energy density (energy per second). We gather all such segments across all low and high frequencies into a single list and rank them by energy density. If a signal contains N ground-truth symbols, we retain the $2N$ segments with the highest energy density (one low and one high segment per symbol), which correspond to the most energetic tone intervals. Very short, low-energy segments are implicitly discarded by this ranking.

Step 6: Period-Symbol Matching. Given N ground truth symbols, we match them to the $2N$ selected segments: for each symbol, we pair one low-frequency and one high-frequency segment that overlap (or are closest in time), and define the symbol period as the union of the paired segments. These symbol periods are then ordered chronologically and matched one-to-one with the ground truth symbols in their original order.

Step 7: Frequency Pair Detection and Symbol Verification. For each matched period, we retain the dominant low and high frequency (the centers of the paired segments) and map this frequency pair to a DTMF symbol using the standard DTMF frequency table. We finally compare the resulting symbol with the assigned ground truth symbol to validate the matching.

Hyperparameters: STFT window length 256, overlap 128, energy clustering $k = 2$, frequency clustering $k = 8$ (the only DTMF prior), BIC penalty coefficient $2\sigma^2 \log(T)$, and selection of the top $2N$ energy-density segments for N symbols.

This approach is more robust to noise than applying changepoint detection directly on the raw signal, as it leverages the known frequency structure of DTMF signals and the dual-tone property to improve changepoint detection accuracy.

Question 2

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

Answer 2

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

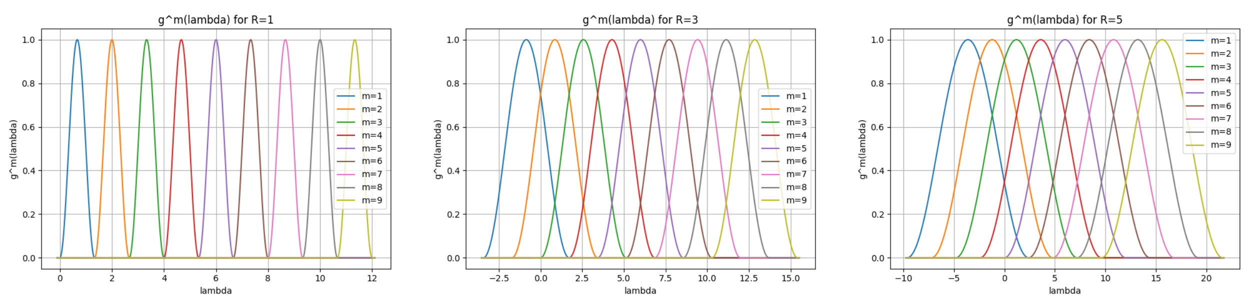
$$\hat{g}^U(\lambda) := \frac{1}{2}\left[1 + \cos\left(2\pi\left(\frac{\lambda}{aR} + \frac{1}{2}\right)\right)\right] \mathbb{1}(-Ra \leq \lambda < 0) \tag{3}$$

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



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

Figure 1: The SAGW kernels functions

We can see that as R increases, the support of the kernels grows (because $R < M + 1$), and the oscillation frequency decreases. Each kernel spans exactly π radians, forming a single "mountain". Also, the kernels get more and more overlapped as R increases.

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 0.8291457286432161. We computed the maximum threshold and not the minimum since for low values of epsilon the condition will be surely met.

The signal is the least smooth at 2014-01-21 06:00:00.

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

Question 5

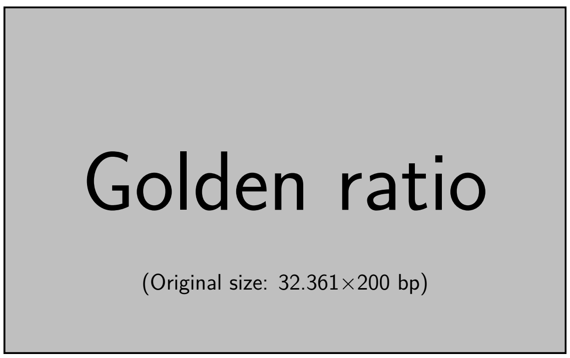
(For the remainder, set $R = 3$ for all wavelet transforms.)

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

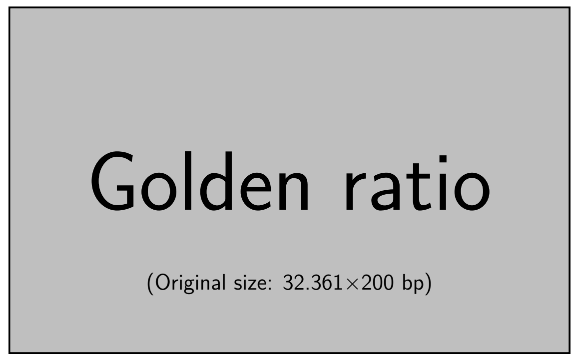
- a node is considered low frequency if the scales $m \in \{1, 2, 3\}$ contain most of the energy,
- a node is considered medium frequency if the scales $m \in \{4, 5, 6\}$ contain most of the energy,
- a node is considered high frequency if the scales $m \in \{6, 7, 9\}$ contain most of the energy.

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).

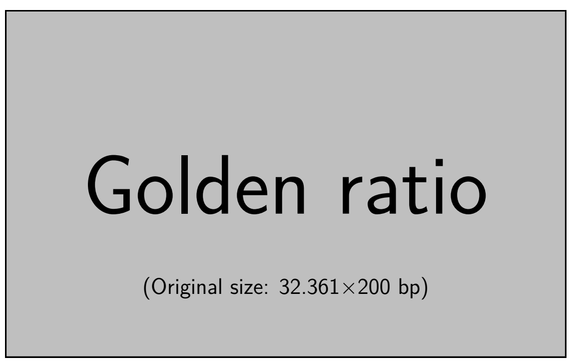
Answer 5



(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.

Answer 6

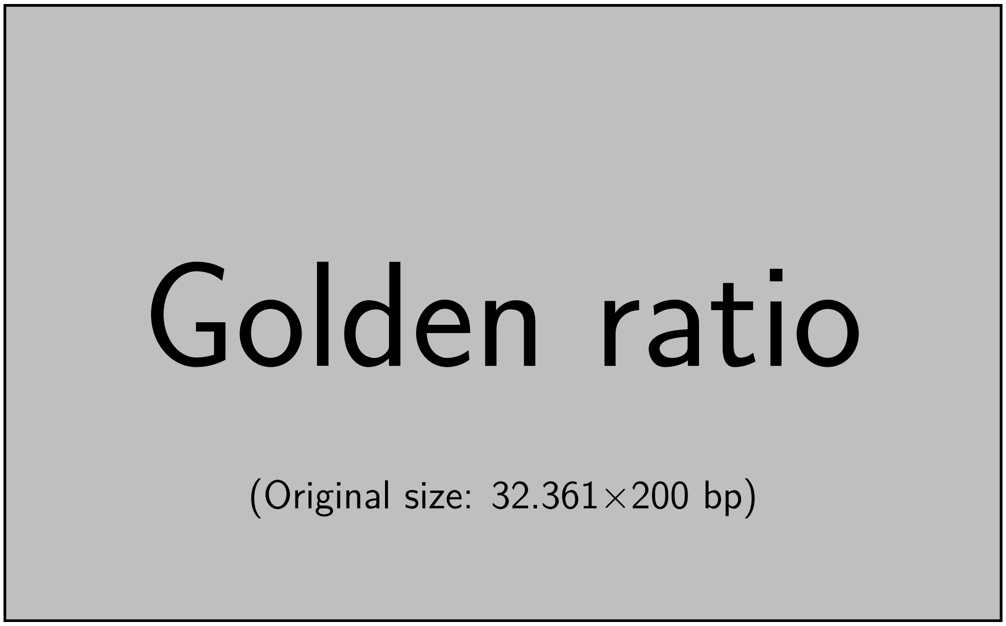


Figure 3: Average temperature. Markers' colours depend on the majority class.

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

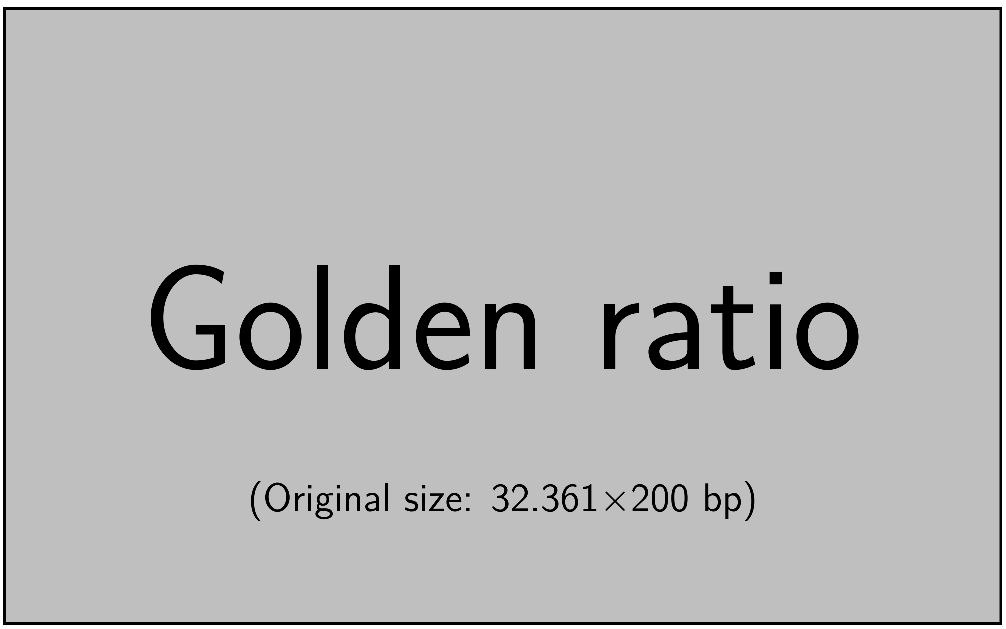


Figure 4: Average temperature. Markers' colours depend on the majority class.