Heart rate variability

Sofia Almirante Castillo, 1527718
Rut Blanco Prieto, 1490761

January, 2023

1. Introduction

Cardiovascular diseases are one of the most common causes of death, of which the most predominant causes are coronary heart disease and cerebrovascular accidents. These cardiovascular diseases can be detected by auscultation of heart sounds, electrocardiograms (ECG) or ultrasound scans. The problems that make cardiovascular diseases one of the leading causes of death are smoking, high cholesterol, hypertension and bad habits among other causes.

The ECG is the graphical representation of the electrical activity of the heart as a function of time, this allows to know the heart rate variability (HRV) whitch is the temporal variation between sequences of consecutive heartbeats. The ECG consists of 5 peaks (PQRST) and the different intervals between the peaks:

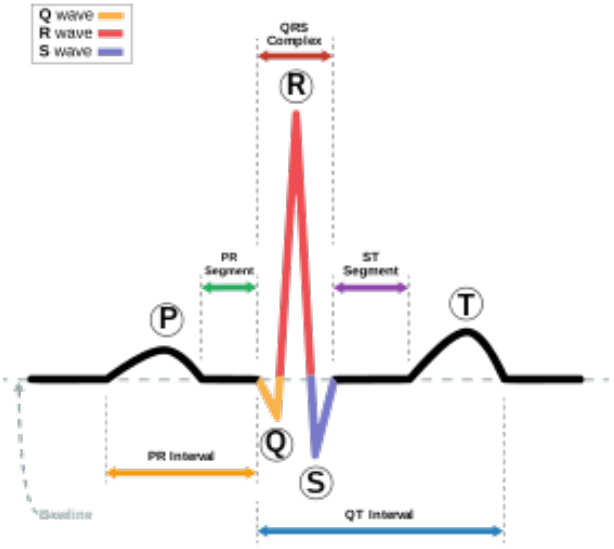


Figure 1: Picks and intervals in a normal ECG signal

- P wave: This wave in ECG results in the atrial contraction as it represents the atrial depolarization
- Q : This wave represents the first negative deflection in the ECG signal.
- R : This wave will represent the first positive deflection in the ECG signal.
- S : After the positive deflection of the R wave next negative deflection is represented by the S wave.

- T wave: This wave represents the ventricular repolarization in the ECG signal.

In this case we will focus on the detection of arrhythmias from ECG signals.

1.1. Arrhythmias

Any of the disorder or abnormality in the normal activation sequence of the myocardium is called cardiac arrhythmias. This activation of the myocardial sequence is represented in the difference of consecutive R-R' peaks. There are two type of arrhythmia on the basis of the R-R' interval:

- Bradycardia: the heart rate is less than 60 beats/min. The patients with increase in the pressure myxoedema and jaundice can rise the effect of the bradycardia.
- Tachycardia: the heart rate is greater than the 100 beats/min, the presence of the ectopic focus in the atrium can regularly cause the tachycardia.

2. Analysis

The set of ECGs analysed in this work were obtained from the MIT-BIH arrhythmia database [4]. This database consists of 23 records obtained randomly from over 4,000 long-term Holter recordings obtained at Beth Israel Hospital between 1975 and 1979, and 25 records also selected from this set but representing rare phenomena that if not selected in this way would not be well represented because of the small sample size. The subjects were 25 men aged 32 to 89 years and 22 women aged 23 to 89 years and each ECG recordings with a length of a 30 min. The signals are sampled with a frequency of 360 Hz.

In most recordings, the upper signal is a modified limb lead II (MLII), which is obtained by placing the electrodes on the chest. The lower signal is usually a modified V1 lead (occasionally V2 or V5, and in one case V4), where normal QRS complexes are usually prominent in the upper signal.

Python has been used to perform the analysis of this data. The detection of the beat in the ECG signal is divided into three stages:

1. Imput ECG signal
2. Procesing and filtering the signal
3. Feature selection and R peaks detection

First, the signals are downloaded in .csv format from the above-mentioned database at random. Then, they are read by our program with the *pandas* library, and stored as a *dataframe*. To make it easier to work with the data, the names of the columns are changed.

In a first approach to the data, histograms of the MLII and V data in the files are generated. In addition to this, the visualization of the MLII series over time in a short period time window is also generated, in order to understand the behaveour of the heartbeat.

The code used to generate these first representations is provided below, and the resulting figures are shown in Fig. 2 and Fig. 3.

```
import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv("#ECG.csv")
```

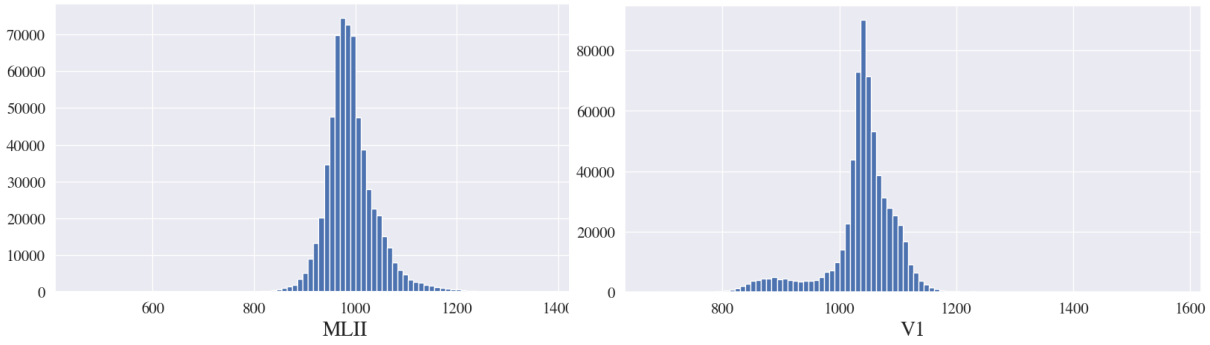
```
df = data.rename(columns={"'sample #'": 'sample', "'MLII'":'ML', "'V1'":'V'})

plt.hist(df['ML'], bins = 100)
plt.xlabel('MLII')
plt.show()

plt.hist(df['V'], bins = 100)
plt.xlabel('V1')
plt.show()

plt.plot(df['sample'], df['ML'])
plt.xlim([0, 700])
plt.ylim([850, 1200])
plt.xlabel('sample')
plt.ylabel('MLII')
plt.show()
```

As explained above, priority shall be given to the use of the MLII signal related to the upper recorded signal due to a better clarity when reading them. This difference in clarity can be seen with histograms obtained from the same ECG but with the upper and lower signals shown in Fig. 2, where it can be clearly observed that the upper signals are more symmetrical.



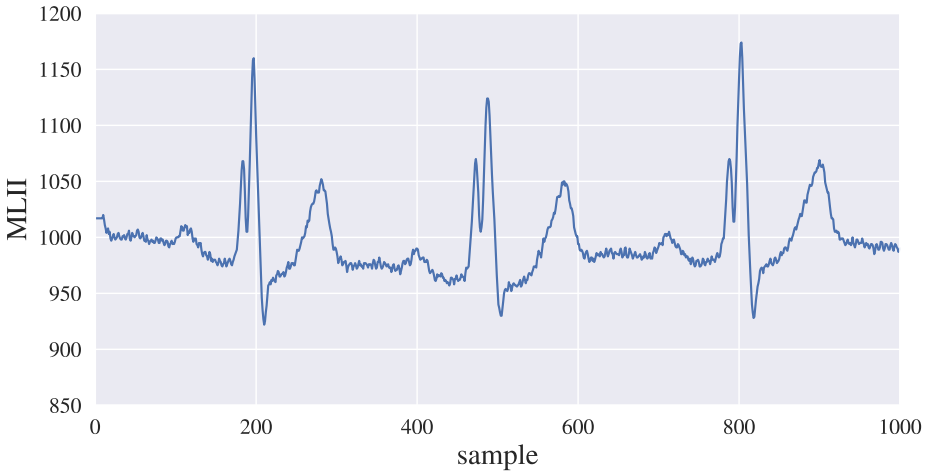
(a) MLII histogram representation

(b) VI histogram representation

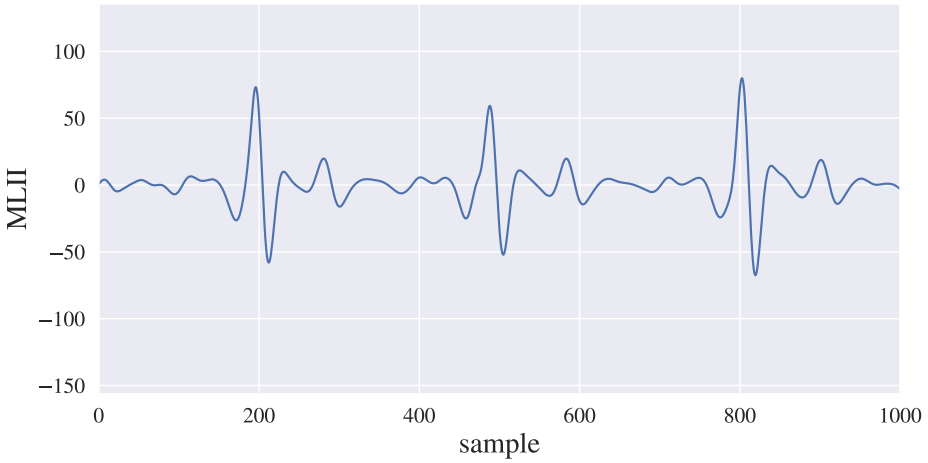
Figure 2: Comparison of histograms obtained from the MLII and V1 series of the sample 111.

If we compare the histogram obtained by plotting these data with a dirac function, which corresponds to a vertical impulse, it can be assumed that the more scatter there is in the histogram the more variation in heart rate variability there is.

In Fig. 3.a, the MLII signal is represented in function of the sample (which is equivalent to time) in order to visualize the heartbeats. There, the different theoretical parts shown in the theoretical ECG signal in Fig. 1 can be identified, being the R peak the most obvious one.



(a) MLII electrocardiogram



(b) MLII electrocardiogram filtered

Figure 3: Comparison between an electrocardiogram with unfiltered and filtered frequencies of the sample 111

As seen in the paper by F. Yaghouby et al. [6], in order to be able to analyze the electrocardiogram properly, the ECG signal seen in Fig. 3.a is filtered, thus simplifying the selection of the maxima (R peaks). Following the of the mentioned paper,

The frequencies are filtered, as seen in F. Yaghouby et al., using a 5-15 Hz bandpass filter. This results in Fig. 3.b, in which one can observe a decrease in the intensity of the smaller peaks while preserving the position of the R peaks. The code used to filter the frequencies is provided below and it is based on *heartpy* [2, 3], a *python* library created to study ECG series.

```
import heartpy as hp
filtered = hp.filtering.filter_signal(df['ML'], [5,15], 360, order=2,
           filtertype='bandpass', return_top=False)
```

In order to know if a patient has an arrhythmia, either tachycardia or bradycardia, it is necessary to know the beats per minute of the provided signal. To obtain this, the maximums obtained in the previous diagram are represented as a function of time, where it is necessary to ensure that all the maximums are uniform in their temporal distribution and the speed at which they occur.

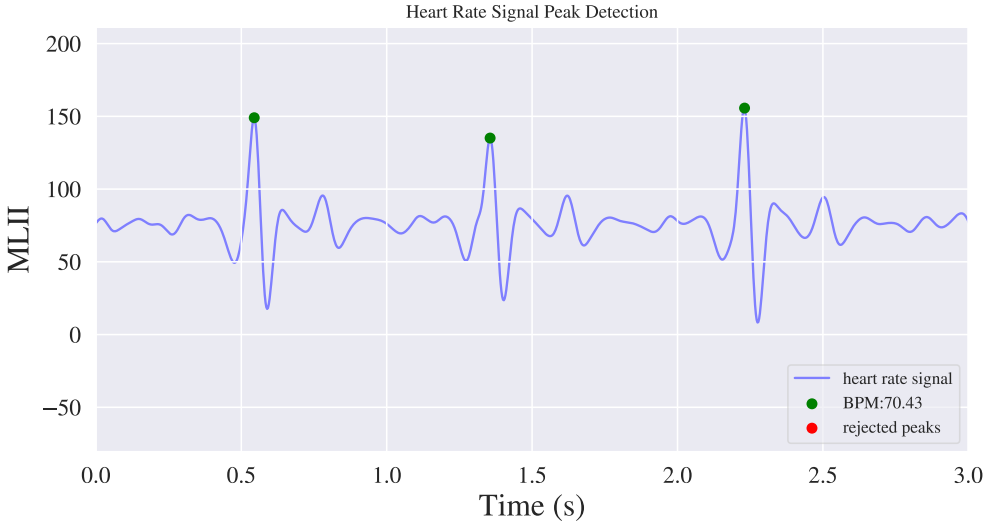


Figure 4: In blue, filtered ECG of the sample 111 in function of time. In green dots, the R peaks detected.

A graphic representation of the R peak detection can be observed in Fig. 4, where it is represented the filtered ECG of the sample number 111 by a blue line, and by green dots the R peaks detected in a small window of time.

The code used to create this figure is shown below, and it also provides the beats per minute of the signal, based on the time difference between the detected peaks.

```
wd, m = hp.process(filtered, 360)
plt.figure(figsize=(12,4))
hp.plotter(wd, m)
plt.xlim([0,10])
plt.ylabel('MLII')
```

Once the maxima that relate to the R-peaks are well identified, a graph is made showing all the maxima obtained in the complete electrocardiogram. As can be seen in Fig. 5, each ascending vertical line corresponds to an R-peak.

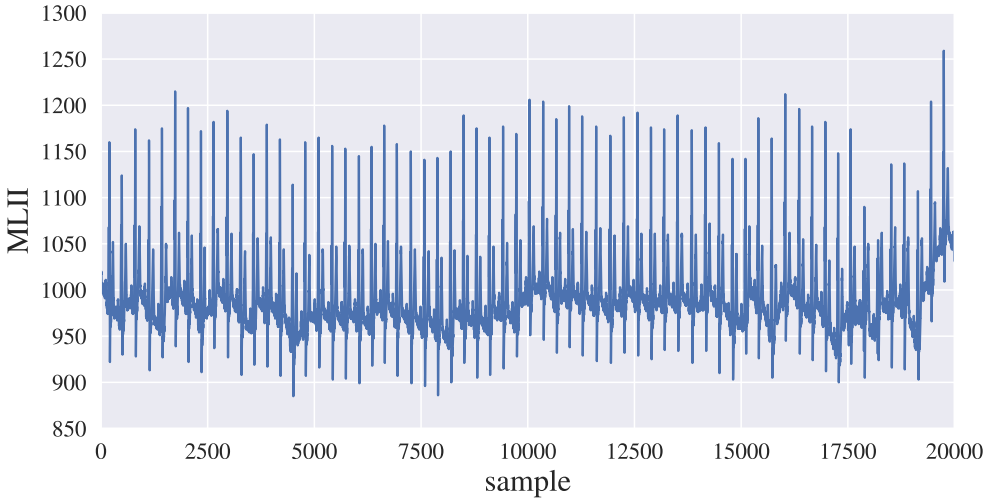


Figure 5: ECG of the MLII signal of the sample 111.

As far as we can deduce, the patient in electrocardiogram 111 is a patient without cardiac problems since his pulse is between 60 and 100 BPM, and throughout the whole electrocardiogram his peaks remain constant throughout the sample, so it can be deduced that he does not have irregularities in his heartbeat either.

In order to compare this electrocardiogram with that of another patient with a cardiovascular disease and to be able to analyse the differences in the results obtained, the electrocardiogram of patient 203, belonging to the 25 samples selected to represent the arrhythmia as explained at the beginning of section 2, was chosen.

Looking at Fig. 6, we can see that the Q, R and S peaks, as well as the PR and QT intervals, and the respective P wave and T wave, can be quite clearly differentiated on the 111 electrocardiogram. However, the differentiation of these parts in the electrocardiogram 6 is more difficult.

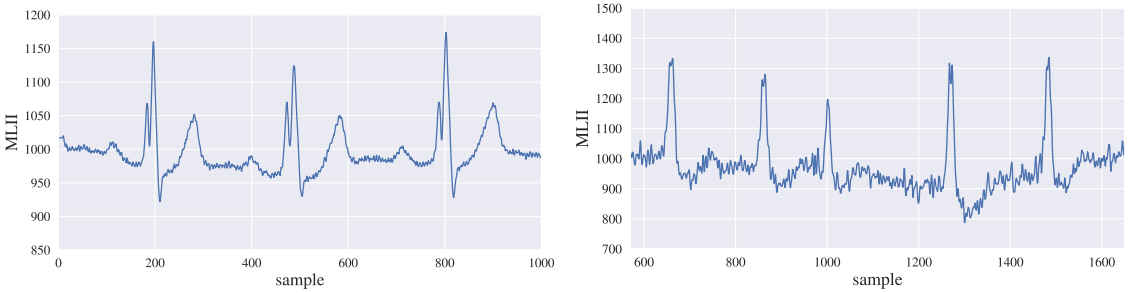


Figure 6: ECG of the MLII signal of the samples 111 (left) and 203 (right).

Once the signal has been filtered to be able to analyse the maximum points as can be seen in Fig. 7, the pacient on sample 111 has a pulse per minut of 70.43 BPM while in the patient in the sample 203 is 111.77 BPM. This high BPM in the sample 203 confirm that the patient has arrhytmia, specifically tachicardia.

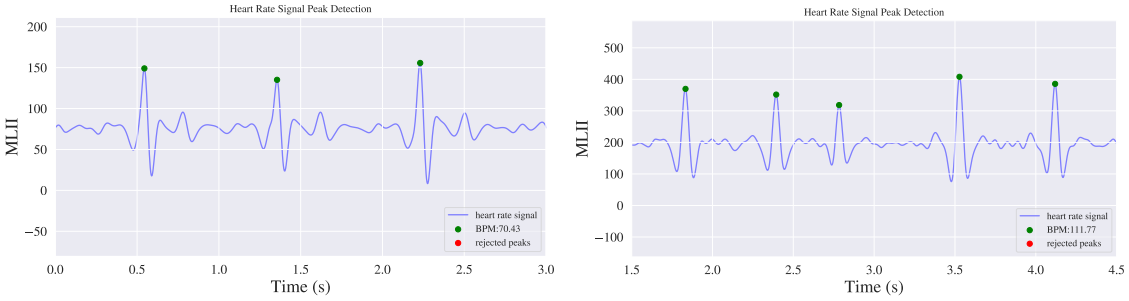


Figure 7: Filtered ECG of the sample 111 (left) and 203 (right) in function of time. In green dots, the R peaks detected.

Finally, as it can be seen in Fig. 8 where the sample maxima of the complete electrocardiogram are represented, it can be seen that those referring to the sample 111 are more uniform the thoses referring to the sample 203.

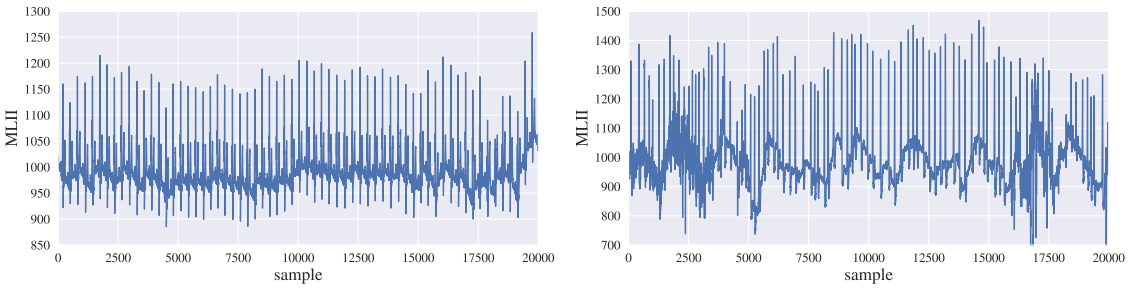


Figure 8: Filtered ECG of the MLII signal with the maximum of the samples 111 (left) and 203 (right).

With this code, it is efficient to obtain the BPM of a sample series, which combined with the heartbeat plot has been used to study the samples between 200 and 234, where according to the database providers are expected to find arrhythmia patients.

The results of the study are shown in Table 1.

Table 1: BPM obtained of the analysis of the samples 200-234.

sample	**BPM**	**comments**
200	87.61	
201	79.54	
202	65.94	
203	111.77	Tachycardia
205	87.68	
207	140.44	Tachycardia episode, Ischemic heart disease[a]
208	99.6	
209	99.91	
210	88.46	
212	91.29	
213	108.03	Tachycardia
214	76.58	
215	112.11	Tachycardia
217	72.71	Duplicated peaks if not filtered
219	76.75	Ischemic heart disease
220	67.07	
221	85.69	Ischemic heart disease
222	94.03	
223	89.78	Ischemic heart disease
228	70.57	Duplicated peaks if not filtered
230	74.96	
231	62.70	
233	103.63	Tachycardia
234	91.53	

[a]The S-T segment observed in the heartbeats of these samples is not isoelectric, which is a pattern that can be associated with an ischemic heart disease.

In order to obtain a better understanding of the carried analysis and some of the situations found, in the Annex can be found some Figures explaining the comments.

Another useful tool when analyzing HRV and revealing rhythms disorders is the Poincare graph [1], which consists of the plot of every RR interval against the consecutive RR interval. Each couple of intervals is represented by a dot in the graph, resulting in clusters. The shape, size and positions of the clusters can be

used to analyze the rhythm of the studied signal [1].

- **Position:** When the cluster is in the lower part of the line, it represents tachycardia. Usually, it will have a narrow shape that indicates the dominance of the nonrespiratory components regulating the heart rate. When it is in the upper part, it represents bradycardia. It usually presents a wider shape, indicating the dominance of the respiratory components regulating the heart rate.
- **Other clusters:** The appearance of other clusters indicates an arrhythmical behavior. They will typically represent premature supraventricular or ventricular beats or pauses.
- **Shattered image:** Might be a proof of Atrial Fibrillation.
- **Symmetry:** If the cluster presents assymetry, it is possible that ryhtm disorders are occuring.
- **Pauses:** The pauses manifest themselves as dots above 2000ms.

The code used to create the Poincaré graph is the following, and it is also based in the *heartpy* library.

```
wd, m = hp.process(df['ML'], 360)
hp.plot_poincare(wd, m)
```

In Fig. 9, the Poincaré graph for the sample 111 and 207 are shown, as they represent clearly different behaviors of the graph. The first one presents a normal behavior, and it is positioned in the upper part of the graph, which indicates a lower BPM. This is compatible with the results previously obtained for this sample. In the other hand, the second graph presents different clusters, one situated in the tachycardia zone and one in the center. This is also compatible with the previous analysis of the sample, which presented a normal BPM followed by a tachycardic episode (as shown in the Appendix C.

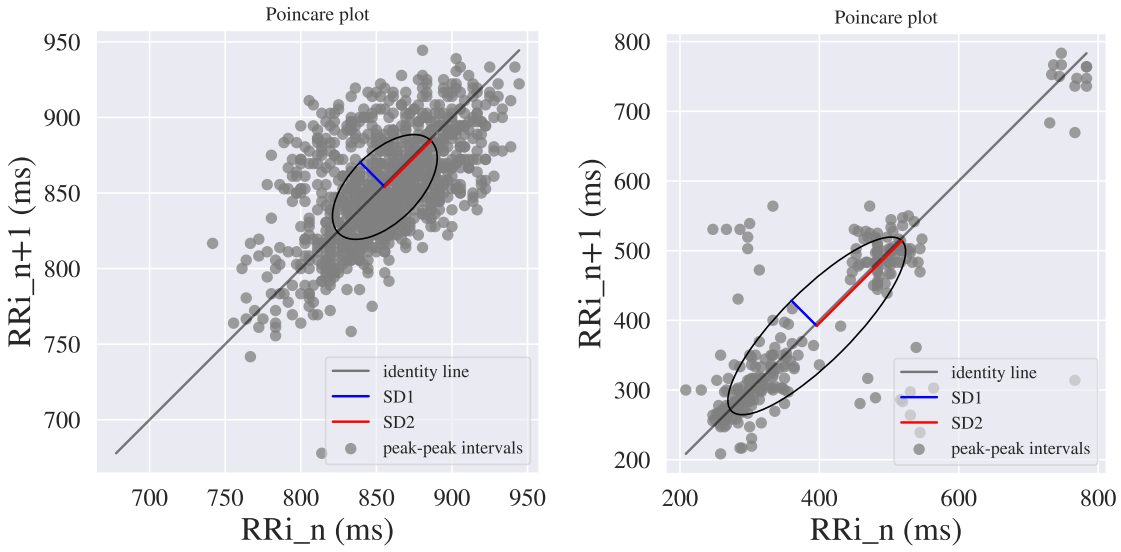


Figure 9: Poincaré graph of the samples 111 and 207.

In Fig. 10, other examples of the Poincaré graph are shown. The graph for the sample 215 is displayed in the left and is clearly not symmetric, which is compatible with the arrhythmia found in the previous analysis. The graph for the sample 203 is displayed in the right, and it is an example of a shattered image, which could be a proof of Atrial Fibrillation (arrhythmia characterized by rapid and irregular beating of the atrial chambers of the heart). This is compatible with the high BPM obtained for this sample.

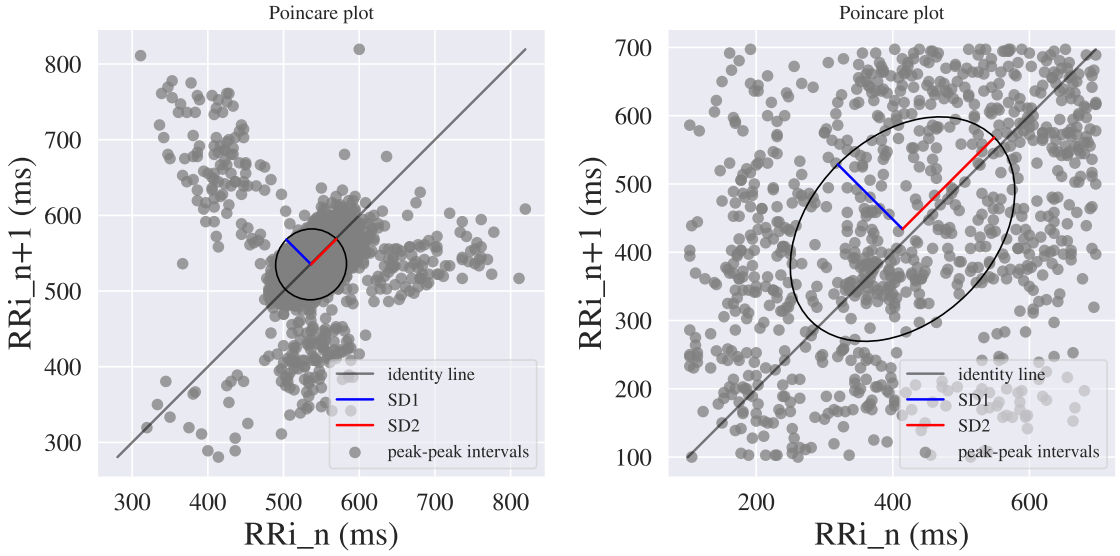


Figure 10: Poincaré graph of the samples 215 and 203.

Finally, it can be observed in the Fig. 11 the apparition of a vertical line at the left of the graph of the samples 221 and 223, marked with a green circle. This is associated with ventricular premature beats, which can be an explanation for the weird heartbeats observed in these samples shown in Appendix A.

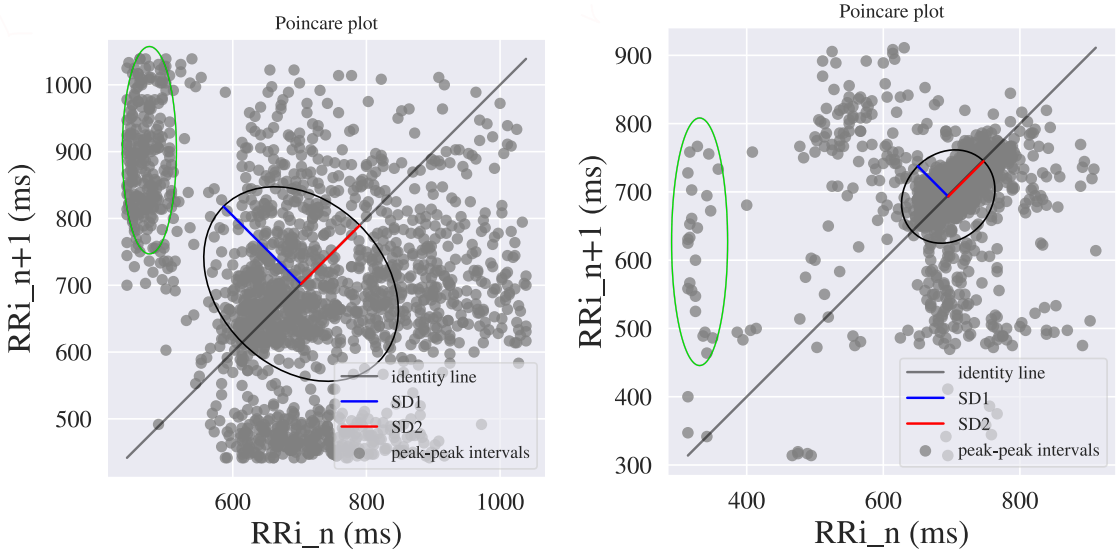


Figure 11: Poincaré graph of the samples 221 and 223.

Appendix

A Ischemic heart disease

In the samples 207, 219, 221 and 223, the S-T segment is not isoelectric or the T peak is inverted. An explanation for this anomaly could be an ischemic heart disease [5]. An example of these unusual patterns

found in the samples are provided in Fig. 12 and Fig. 13.

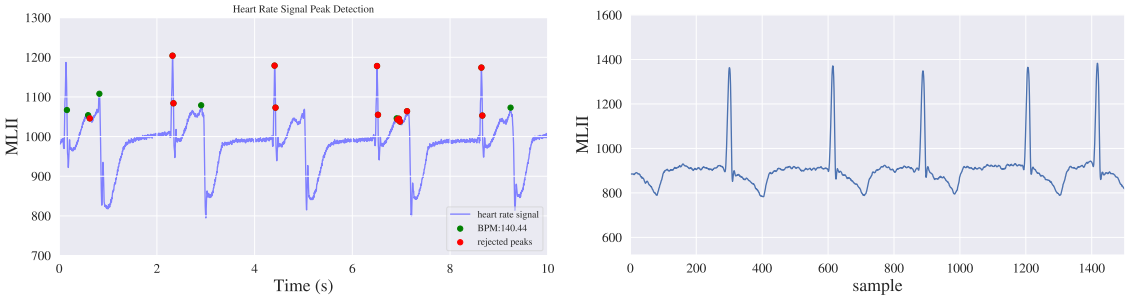


Figure 12: Heartbeats of the samples 207 (left) and 219 (right).

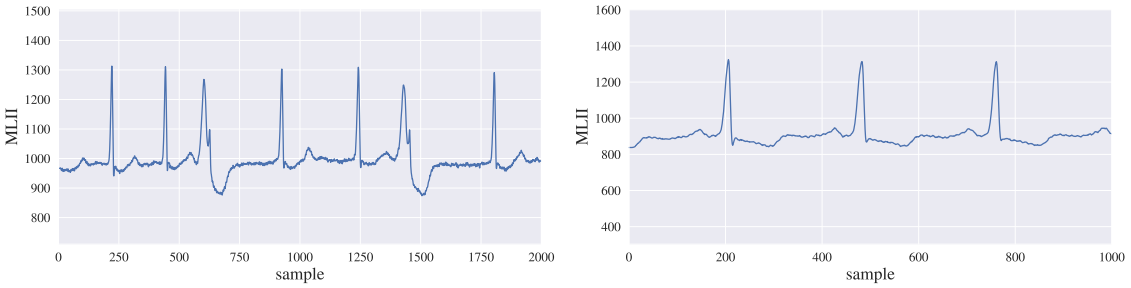


Figure 13: Heartbeats of the samples 221 (left) and 223 (right).

B Duplicated peaks

In some cases, like in the samples 217 and 228, a big difference between the BPM obtained with the filtered and not filtered signal has been observed. This is due to the peak detection algorithm of the library and the small difference in intensity between the R peak and the T peak of these samples. When filtering the signal, the R peak is more prominent, which results in a better detection of the peaks.

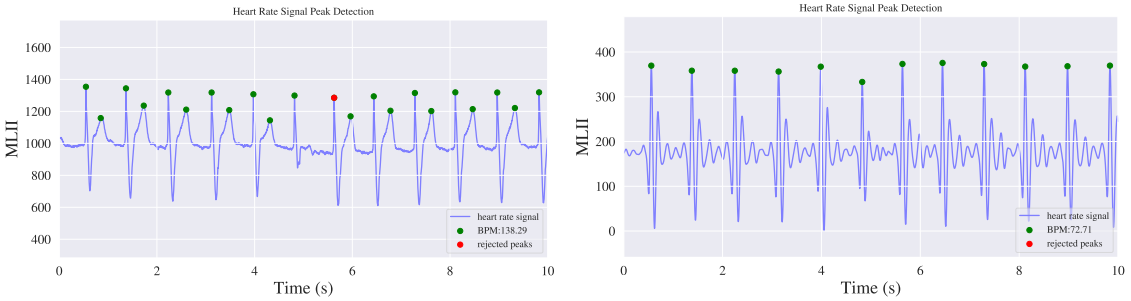


Figure 14: ECG signal of the sample 217 not filtered(left) and filtered (right).

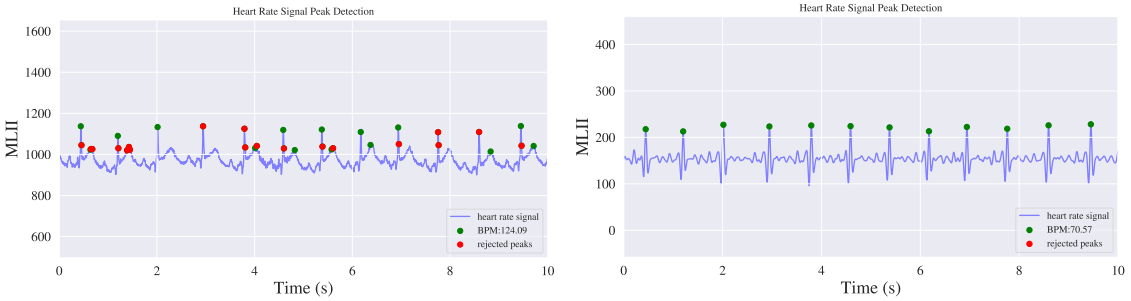


Figure 15: ECG signal of the sample 228 not filtered(left) and filtered (right).

C Ventricular tachycardia

In Fig. 16 can be observed an example of a tachycardic episode observed in the sample 207, in which the frequency of the heartbeats clearly changes for 10 minutes.

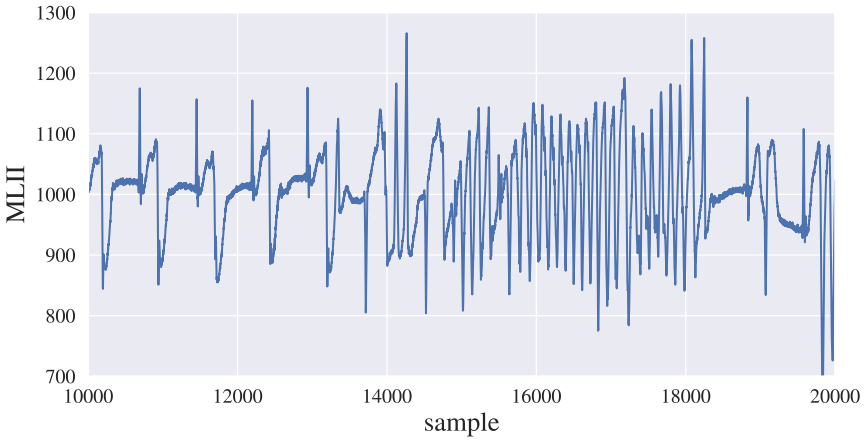


Figure 16: ECG of a tachycardic episode observed in the sample 207.

Observing in detail the shape of the ECG shown in Fig. 17, the pattern of a ventricular tachycardia is observed, characterized by a distorted QRS wave with no P or T component.

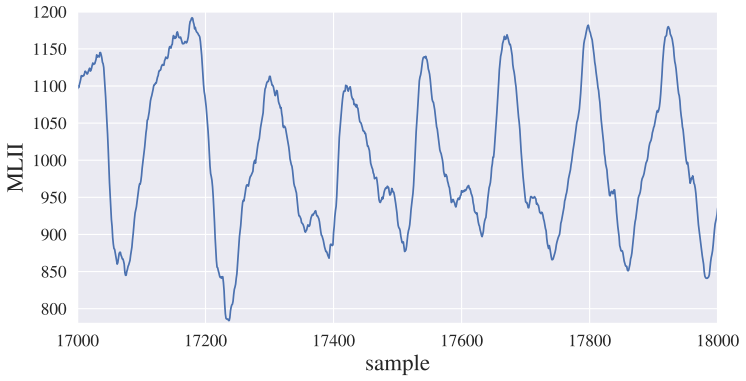


Figure 17: ECG of a shorter time window of a tachycardic episode observed in the sample 207.

D ECG with identifiable elements

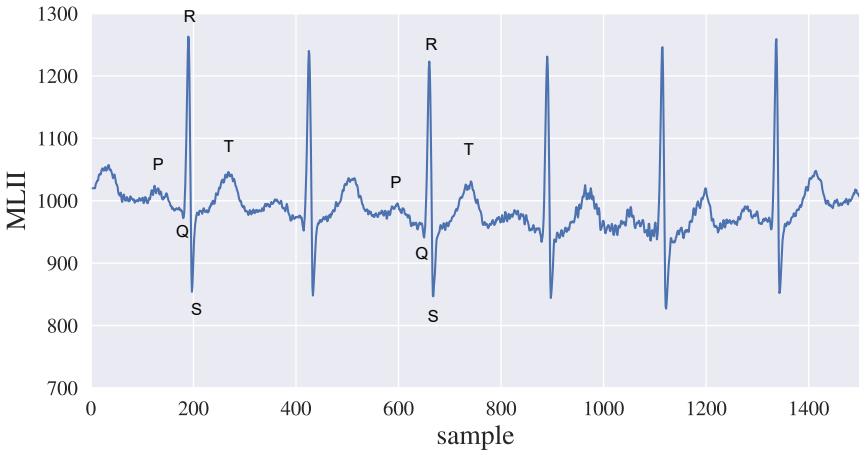


Figure 18: EGC of the sample 207 with peaks R, Q, S and waves P,T indicated.

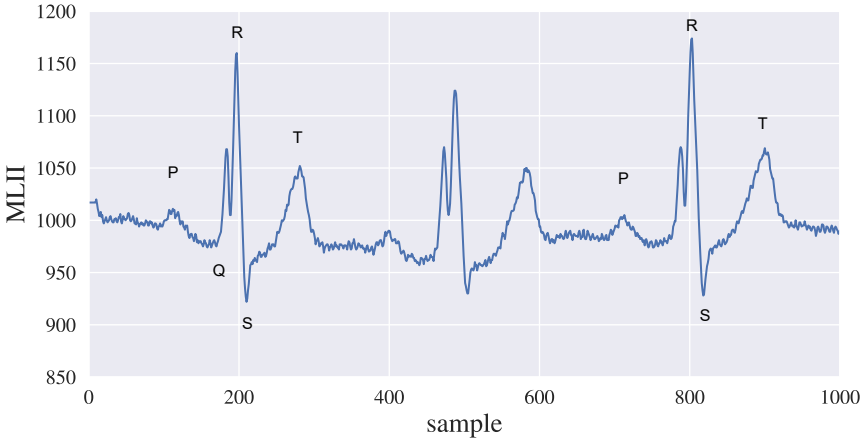


Figure 19: EGC of the sample 111 with peaks R, Q, S and waves P,T indicated. It can be seen that the R peak is duplicated and Q peak is more difficult to identify.

References

[1] Brennan, M., Palaniswami, M., Kamen, P., 2002. Poincaré plot interpretation using a physiological model of hrv based on a network of oscillators. American Journal of Physiology-Heart and Circulatory Physiology 283, H1873–H1886. doi:10.1152/ajpheart.00405.2000.

[2] van Gent, P., Farah, H., Nes, N., Arem, B., 2018a. Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. doi:10.13140/RG.2.2.24895.56485.

[3] van Gent, P., Farah, H., Nes, N., Arem, B., 2018b. Heart rate analysis for human factors: Development and validation of an open source toolkit for noisy naturalistic heart rate data .

[4] Mark, R.G., Moody, G.B., 2005. Mit-bih arrhythmia database. doi:10.13026/C2F305.

[5] Unknown, 2017. Cómo interpretar un electrocardiograma. URL: https://enfermeriacreativa.com/2017/02/04/como-interpretar-un-electrocardiograma/.

[6] Yaghouby, F., Ayatollahi, A., Soleimani, R., 2009. Classification of cardiac abnormalities using reduced features of heart rate variability signal .