Marta Pereira Neves

Descrição da dor através de aprendizagem não supervisionada

Unsupervised learning for pain description

DOCUMENTO PROVISÓRIO

Universidade de Aveiro
2025

Marta Pereira Neves

Descrição da dor através de aprendizagem não supervisionada

Unsupervised learning for pain description

Projeto apresentado à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Licenciatura em Engenharia Biomédica, realizado sob a orientação científica da Doutora Susana Brás, Professora Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e investigadora do Instituto de Engenharia Eletrónica e Informática de Aveiro (IEETA).

o júri / the jury

presidente / president

Doutora Ana Luísa Monteiro da Silva
Professora Auxiliar da Universidade de Aveiro

vogais / examiners committee

Doutora Raquel Sebastião
Investigadora do Instituto de Engenharia Eletrónica e Informática de Aveiro (IEETA) (Arguente)

Doutora Susana Manuela Martinho dos Santos Baía Brás
Professora Auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e investigadora do Instituto de Engenharia Eletrónica e Informática de Aveiro (IEETA) (Orientadora)

agradecimentos / **acknowledgements**

Agradeço toda a ajuda a todos os meus colegas e companheiros.

palavras-chave texto livro, arquitetura, história, construção, materiais de construção, saber tradicional.

resumo Um resumo é um pequeno apanhado de um trabalho mais longo (como uma tese, dissertação ou trabalho de pesquisa). O resumo relata de forma concisa os objetivos e resultados da sua pesquisa, para que os leitores saibam exatamente o que se aborda no seu documento.
Embora a estrutura possa variar um pouco dependendo da sua área de estudo, o seu resumo deve descrever o propósito do seu trabalho, os métodos que você usou e as conclusões a que chegou.
Uma maneira comum de estruturar um resumo é usar a estrutura IMRaD. Isso significa:

- Introdução
- Métodos
- Resultados
- Discussão

Veja mais pormenores aqui:
`https://www.scribbr.com/dissertation/abstract/`

keywords textbook, architecture, history, construction, construction materials, traditional knowledge.

abstract

An abstract is a short summary of a longer work (such as a thesis, dissertation or research paper).
The abstract concisely reports the aims and outcomes of your research, so that readers know exactly what your paper is about.
Although the structure may vary slightly depending on your discipline, your abstract should describe the purpose of your work, the methods you've used, and the conclusions you've drawn.
One common way to structure your abstract is to use the IMRaD structure. This stands for:

- Introduction
- Methods
- Results
- Discussion

Check for more details here:
`https://www.scribbr.com/dissertation/abstract/`

acknowledgement of use of AI tools

Recognition of the use of generative Artificial Intelligence technologies and tools, software and other support tools.

I acknowledge the use of [insert AI system(s) and link] to [specific use of generative artificial intelligence or other tasks]. I acknowledge the use of [software, codes or platforms] to [specific use software, codes or platforms or to other tasks].

Example 1: I acknowledge the use of ChatGPT 3.5 (Open AI, https://chat.openai.com) to summarise the initial notes and to proofread the final draft and the use of Office365 (Microsoft, https://www.office.com) for text writing and productivity.

Example 2: No content generated by AI technologies has been used in this Thesis.

Contents

List of Figures

List of Tables

Glossary

ANOVA	analysis of variance
ANS	Autonomous Nervous System
BL	baseline
BP	blood pressure
CNS	Central Nervous System
CPT	Cold Pressor Test
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EDA	Electrodermal Activity
EMG	Electromyogram
FN	False Negatives
FP	False Positives
GSR	Galvanic Skin Response
HR	heart rate
HRV	Heart Rate Variability
HSD	Honest Significant Difference
IQR	interquartile range
KNN	K-Nearest Neighbours
LOSO	leave-one-subject-out
LSTM	Long-Short-Term-Memory
NPS	Numerical Pain Scale
RF	Random Forest
SVM	Support Vector Machine
TDE	time-dependent entropy
TN	True Negatives
TP	True Positives
VAS	Visual Analog Scale

CHAPTER 1

Introduction

Face your life, its pain, its pleasure, leave no path untaken. - Neil Gaiman

1.1 Motivation

Pain is unpleasant. No one enjoys experiencing it, whether it's the muscle soreness from working out or the more severe pain from a broken bone. In fact, it might even represent a significant burden on a person's quality of life. The unpleasant feeling it provokes may result in functional disability or even a change of behaviour, making an individual feel anger and frustration [1]. However, pain serves a valuable purpose: it alerts us to potential dangers and helps us adapt to situations that could harm us, helping us avoid further injuries [2]. Nevertheless, when it interferes with the way people typically live their lives, it's important to find ways to attenuate it. This burden is particularly profound for patients suffering from chronic pain conditions, for whom discomfort becomes a persistent and debilitating element of life.

Attempts of measuring pain surged with the intent of improving quality of pain management [3]. For this reason, scales like the Numerical Pain Scale (NPS) and the Visual Analog Scale (VAS) have been used since 1950 [4]. Although these are the most used methods, they're very subjective, depending on a patient's own perspective of pain. Moreover, some patients might undermine the severity of their pain or be unwilling to share their experience. For those with cognitive difficulties, these methods of verbal report may not even be appropriate, especially for non-verbal individuals [5].

The challenge of measuring pain makes prescribing pain medication a difficult job for healthcare providers since the only comparison they have is a patient's previous results. This means that the prescription can be excessive and the organism will grow accustomed to it quicker or, if the patient minimizes the severity, they will continue to experience pain as their tolerance increases. This highlights the importance of finding an objective way of describing pain.

With this aim, researchers have analysed physiological signals with the intent of selecting features that might allow for an objective pain description, due to those signals' direct correlation with the Autonomous Nervous System (ANS). Various techniques have been explored, most employing machine learning models such as Random Forest (RF) and Support Vector Machine (SVM) to classify pain. Although these have proven to be successful, this project suggests the use of unsupervised learning for feature selection, through the use of clustering. By doing this, the goal is to find natural partitions within the data, attempting to distinguish pain from the lack of it without using class labels.

1.2 Work Objectives

1.3 Project Structure

CHAPTER 2

Theoretical Introduction

2.1 Pain

Pain is often defined as "an unpleasant experience which we primarily associate with tissue damage, or describe in terms of such damage, or both" [6]. Although pain affects the Central Nervous System (CNS) directly, it also impacts the ANS, since it connects the CNS to the internal organs [7].

According to its duration, pain can be classified as acute or chronic. Acute pain is induced by the activation of nociceptor sensory neurons, which occurs in the presence of actual or potential damaging stimuli, such as intense heat or cold and excessive mechanical force, or due to inflammation [8]. On the other hand, chronic pain is defined as lasting more than three months [9] and can be classified into nociceptive, neuropathic or nociplastic pain. Nociceptive pain results from continuous stimuli associated with tissue injury, while neuropathic pain results from damage to the peripheral or central nervous system. Lastly, nociplastic pain is a broader term, that is applied to chronic pain when it can't be described by the other two terms [10].

The Numerical Pain Scale is one of the most widely used traditional methods for assessing pain. It typically involves asking patients to rate their pain on a scale from 0 to 10, where 0 represents no pain and 10 signifies the worst pain imaginable [3]. While simple and easy to administer, this method relies entirely on the individual's subjective interpretation of their pain, which can be influenced by factors such as mood, stress and gender, not allowing for viable continuous monitoring. Besides this, the difference between pain scores may not be comparable in scaling the intensity of pain [11]. Finally, its verbal component is a limitation for non-verbal patients who may not be able to describe their pain using this method.

Another traditional method for assessing pain is the Visual Analogue Scale. This scale represents a continuous range of values, and it mainly uses a horizontal line measuring exactly 100 mm. The patient is asked to make a mark on the line according to their level of pain, then, the distance of the mark is measured and recorded in millimetres or centimetres [4]. However, even though it's non-verbal, a minimum level of motor abilities is necessary to correctly use

the VAS, which makes it inadequate for scaling pain in some patients with motor impairment, for example. This method, although it seems to be more detailed in values than the NPS, is still very subjective, not allowing for an accurate pain description and comparison. This method can also be used for assessing other symptoms, however. For example, in the dataset used in this project, it was used to assess the level of anxiety, happiness, fear and stress that the participants felt, on a scale of 0-100%, and the arousal and valence states in a -5 to 5 scale [12].

Other similar scales, that also depend on a patient's perception of their own pain, end up also being subjective [11][13]. Due to this, researchers have attempted to find objective ways to describe pain using physiological signals.

2.2 Pain Description

The premise that pain induces changes in the ANS has led to the idea that physiological measures can be used to describe it [14]. For this reason, multiple physiological signals have been analysed in an attempt to select features that distinguish pain from the lack of it, or even classify it as more or less severe.

2.2.1 Electrocardiogram

In this project, the chosen physiological signal was the ECG.

The heart functions as a pump, its activity determined by the sinus node, that transmits excitation to the atrium and ventricule successively. An ECG provides a recording of this electrical activity. Alterations in the typical conduction of electrical impulses through the heart can be detected on an ECG and may reflect underlying physiological changes in the individual [15].

The ECG signal can be segmented into successive waveforms that exhibit a consistent pattern and, therefore, it's possible to split those into specific components. The P wave surges as a result from the depolarization of the atrium. This electrical impulse spreads to the ventricule during the P-R segment and, as the ventricular muscle contracts, the QRS complex appears. Then, the S-T segment corresponds to the period where the ventricles are completely activated. Lastly, the T wave matches the repolarization of the ventricule [16]. These waves are shown in Figure 2.1, as well as other divisions, such as the P-R and S-T intervals. The onset of the P and T waves match the beginning of the P-R interval and the end of the S-T segment, while their offset corresponds to the start of the P-R segment and the end of the S-T interval, respectively. Similarly, the onset and offset of the R wave coincide with the start and end of the QRS complex [17].

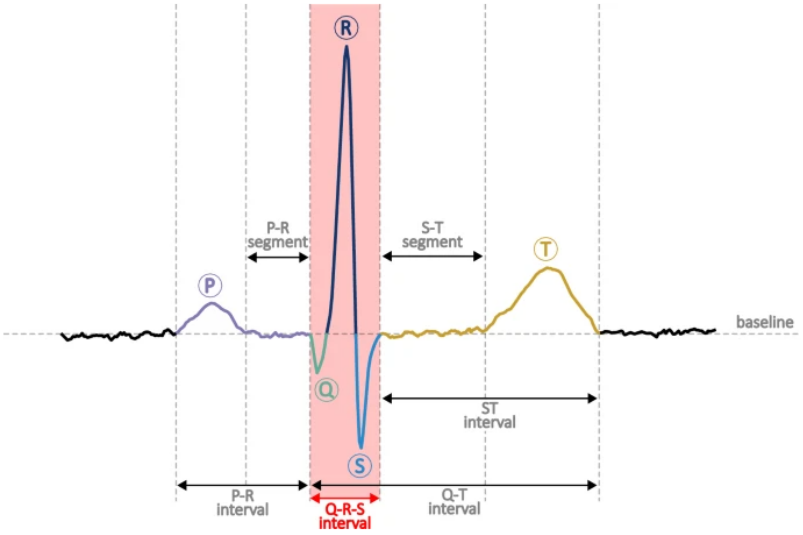

Figure 2.1: ECG Schematic (Adapted from [18])

2.2.2 Related Works

In literature, the most common feature extracted from the ECG is Heart Rate Variability (HRV), which is related to time variations in R-R intervals. Nevertheless, other ECG features have been analysed and selected as successful pain descriptors, along with features extracted from different physiological signals, like the Electromyogram (EMG) and Electrodermal Activity (EDA). For this signals' analysis, various machine learning models have been used, having achieved successful results throughout the years [7][19].

For example, Pavlidou and Tsiknakis [20] used the BioVid Heat Pain dataset, in which heat pain was induced through a thermode, with the intent of obtaining significant features for pain classification. On that account, they extracted features from the ECG, EMG and Galvanic Skin Response (GSR) from 87 participants, where the first are HRV related. To detect pain, SVMs and Long-Short-Term-Memory (LSTM) models were created while a leave-one-subject-out (LOSO) cross validation method was used to evaluate its intensity, distinguishing between a baseline (BL) and four levels of pain intensity (PA1, PA2, PA3 and PA4). The highest accuracy was obtained for multimodal approaches, with a maximum accuracy of 76.69% for SVM and 77.21% for LSTM. However, in unimodal approaches, GSR features led to the best results, when compared with the other signals. Finally, the LOSO method achieved a maximum accuracy of 82.83% in BL versus PA4, in accordance with the remaining results, where the best classification matched the models that compared the highest and lowest level of pain or the BL.

Silva and Sebastião [21] used a Cold Pressor Test (CPT) to induce pain in 36 participants, while eliciting distinct emotions through fear-inducing and neutral videos across two different sessions. This was done with the aim of measuring the effect of emotional contexts on physiological responses to pain. During the study, EMG, ECG, EDA, and blood pressure (BP) data was collected. However, only the ECG signals were analysed, using sliding windows for feature extraction. Subsequently, a variety of machine learning algorithms, including

AdaBoost, Decision Trees, K-Nearest Neighbours (KNN), Linear Discriminant Analysis, Logistic Regression, RF, and SVMs, were employed across three classification strategies: dependent, session-independent, and participant-independent approaches. In the first, a maximum balanced accuracy of 97.4% was obtained, while the second reached 97.7% of balanced accuracy, which supports the hypothesis that the physiological response to pain remained consistent despite differing emotional contexts. Across the three approaches, the RF and Adaboost models attained the best results. Meanwhile, the most relevant features were the amplitude of the S peaks, the amplitude of the T peaks, the amplitude of the T offset, and the amplitude of the R peaks. Although emotion elicitation was intended, self-reported pain levels did not reveal statistically significant differences between the fear and neutral sessions. This result was attributed to patients being too focused on the pain to pay attention to the video.

Choosing an alternative approach, Balasubramani et al [22] analysed ECG signals from 142 patients who underwent surgical procedures. Their methodology involved using two dimensional Continuous Wavelet Transform on the signals, and processing the results through a framework that combines a quantum neural network architecture and classical deep learning models. This new model achieved an accuracy of 94.8% in pain assessment, implying that quantum transfer learning might acquire a fundamental role in future medical diagnostic systems.

CHAPTER 3

Methods and Procedure

In Figure 3.1, the processing pipeline for this project is shown.

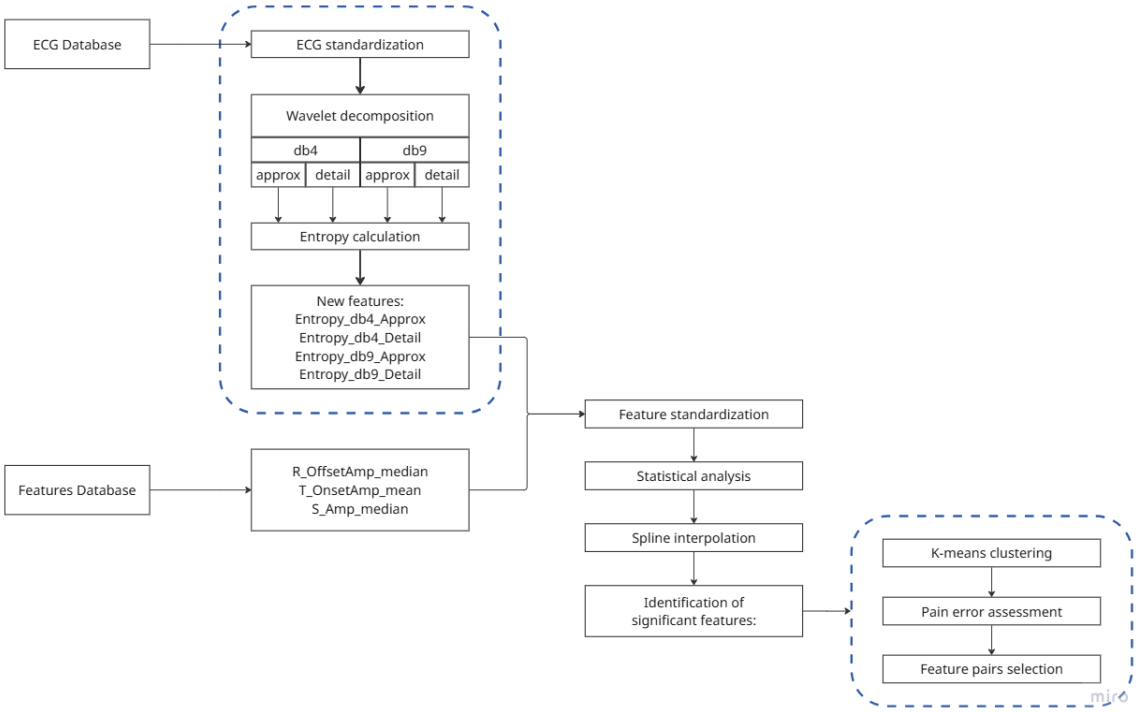

Figure 3.1: Processing pipeline.

3.1 Database Acquisition

In this project, two databases were used: one has data from an ECG signal while the other contains features extracted from that ECG, as described in the article by Alves et al [12].

The data collection protocol followed a structured sequence designed to elicit and record both emotional and pain-related responses. It began with a 5-minute baseline period, during

which participants remained seated in a relaxed position without any external stimuli. During this time, only physiological signals were recorded to establish a baseline. Subsequently, participants watched a 10-minute video, composed of excerpts from comedy, horror, or documentary films, aimed at inducing positive, negative, or neutral emotional states, respectively. This was immediately followed by a second 5-minute stimulus-free period, allowing for recovery and stabilization of physiological responses. Next, participants underwent a CPT, in which they immersed their non-dominant hand in a tank of cold water maintained at 7 ± 1°C. This procedure aimed to induce pain, and participants reported their experience using the NPS at four key time points: (1) before immersing the hand, (2) at the moment pain was first perceived (Pain Threshold), (3) when the pain became unbearable (Pain Tolerance), and (4) three minutes after hand removal. The CPT phase concluded either when the participant reached their pain tolerance or after a maximum of 2 minutes, whichever occurred first. Finally, the protocol concluded with a third 5-minute period without stimuli, serving as a rest phase to observe post-task physiological responses. This process is depicted in Figure 3.2.

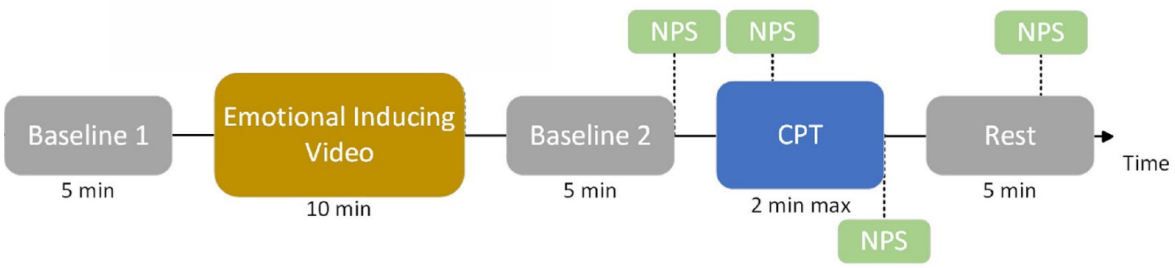

Figure 3.2: Scheme of the protocol applied to obtain the database. (Adapted from [12])

During the study, three physiological signals were recorded: ECG, EDA, and EMG, with EMG signals collected specifically from the trapezius and triceps muscles. Among these, the ECG data were stored in a dedicated database, which was subsequently used in the present project. From the ECG signals, several time series were extracted, including heart rate (HR), amplitude of wave peaks, amplitude of onsets and offsets, and the intervals between consecutive onsets, offsets, and peaks. These features were computed using 10-second sliding windows with 50% overlap. For each window, a set of statistical metrics—namely the mean, median, and variance—was calculated, resulting in a total of 237 features. These were compiled into a features database, which was also employed in the current project.

3.2 Feature Extraction

Four of all the extracted features revealed to be more related to pain assessment, by the use of a gradient booster regressor with 200 estimators and learning rate of 0.1 [12]. So, under the scope of this project, the median of the R wave offset amplitude, the median of the T wave onset amplitude, and the mean and median of the S wave peak amplitude were used as a base for subsequent analysis. Considering that the previously described features are examples of the state of the art features in ECG analysis, and knowing that the pain quantification or identification is still an open area, this work wants to understand and find out other possible

features that may describe the pain process. So, to accomplish such goal, an exploratory analysis to the collected ECG was performed.

3.2.1 ECG Data Exploration

When looking at the time series of the signal, in Figure 3.3, it's noticeable that the range of amplitudes of the QRS complex is way bigger than that of the P and T waves. To tackle this, the ECG signal was standardised according to equation 3.1, where X is the original ECG, μ is the mean, σ is the standard deviation and Z is the standardised signal, which has mean zero and standard deviation one. This means the range of values is the same for all participants, reducing the intervariability that is induced by the initial state of the participant, since it is an uncontrolled variable.

$$Z = \frac{X - \mu}{\sigma} \tag{3.1}$$

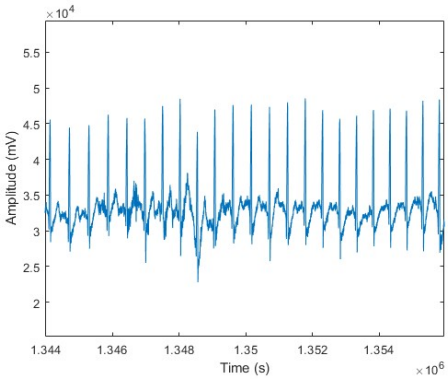

Figure 3.3: Section of the ECG signal of a participant.

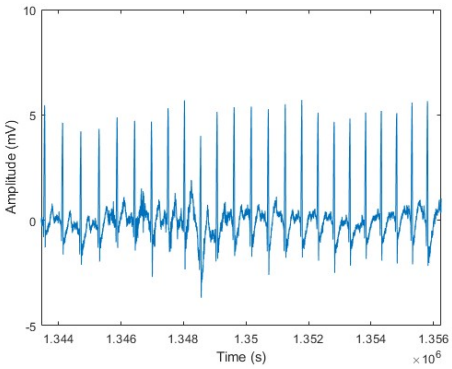

Figure 3.4: Section of the standardised ECG signal of a participant.

Once the signal was standardised, fluctuations were noticed in the time series. By removing high frequencies from the signal, by filtering for example, these would be alleviated. However, there could be information related to pain associated to those frequencies. A solution to these problems is to decompose the signal in wavelets, using Discrete Wavelet Transform (DWT).

Wavelets are built upon basic functions that preserve properties of localization in both frequency and time domains, resembling short wave-like signals with specific formats. The computation of a DWT is primarily based on the employment of low-pass and high-pass filters with different cutoff frequencies to process signals in various frequency bands. This process leads to the decomposition of the signal into approximation and detail components. The approximation component represents the low-frequency content (signal energy), while the detail component captures the high-frequency content associated with short pulses, sudden changes, and noise. The amount of information is changed by filtering, and the scale is changed by decimation and interpolation. Decimation corresponds to lowering the sampling rate, but this can be fixed when the the wavelets are reconstructed [23].

There are many wavelet families, that differ from each other in the format of the approximation wave, but the selection of one should consider the characteristics of the signal and

the decomposition level must be chosen accordingly. For this project, two types of wavelet were chosen, according to two specific criteria: one's format should be similar to that of the ECG wave, complementing the remaining features in gathering whether the QRS complex is relevant to describe pain, and the other should have a higher frequency, to understand if there's important information associated with the fluctuations of the signal, while still maintaining its smoothness. Accordingly, Daubechies-4 ('db4') and Daubechies-9 ('db9') were chosen, which waveforms can be seen in Figure 3.5.

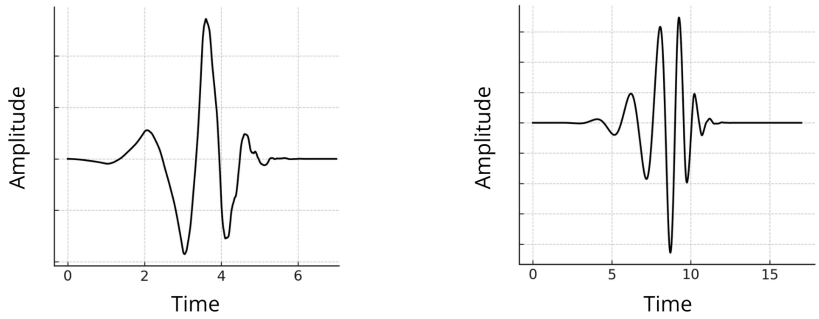

Figure 3.5: Daubechies-4 and Daubechies-9 Waveform.

Regarding the number of levels of decomposition, as can be seen in Figure 3.6, when there are two levels the R peak is noticeable in the first detail, which is not something this project means to highlight in the detail. Therefore, only one level of decomposition was applied.

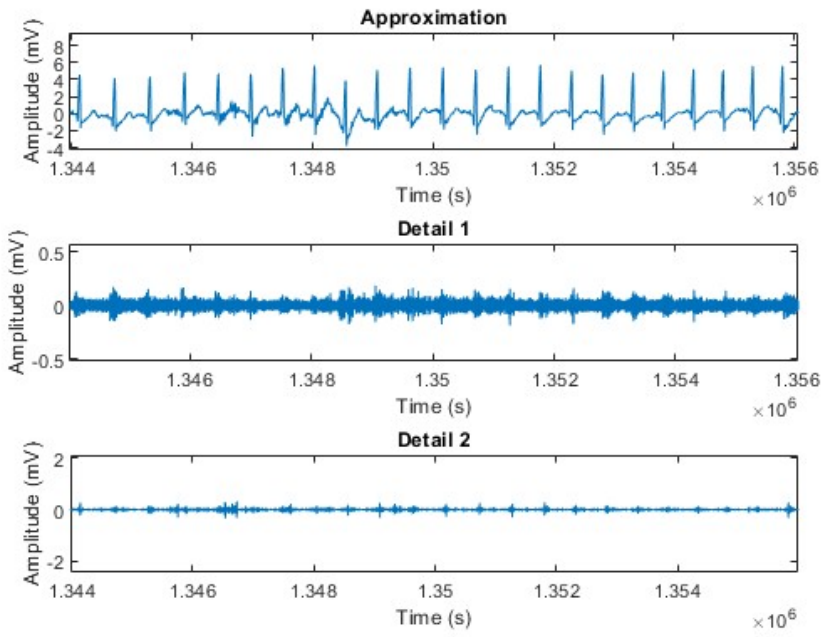

Figure 3.6: Daubechies-4 Wavelet Decomposition - level 2.

3.2.2 Entropy Calculation

Entropy quantifies the disorder in time series data. Due to this, features associated with it have been used in the analysis of physiological signals [24]. Since this quantity might be affected in the ECG when someone experiences pain, it was calculated for the approximation and detail components of the resulting wavelets. In order to keep consistency with the other features, a same sized window was used, that is, 10 seconds with 50% overlap.

There are various types of entropy. To select one, a research was executed on the tool Scopus using the key words entropy, ecg, signal, pain, and classification. Only four articles were found, but two of them weren't focused on pain description. Due to this, only the remaining two were analysed. Maity and Jana [25] analysed four types of entropy applied to wavelets using a RF model, while Adjei et al [26] calculated permutation entropy using Shannon entropy. These articles show that there isn't a lot of research that uses entropy for pain description, so the type of entropy was chosen based on the aim of the selected features, which is to evaluate if there are signal transitions in the presence of pain. Due to this, a time-dependent entropy (TDE) was implemented. The amplitude range of the signal was segmented into 100 non-overlapping and successive partitions. For each of these, the probability P_i was determined by computing the ratio between the number of signal samples in that interval, and the total number of samples in the window. The entropy was then calculated using Equation 3.2, where M equals the number of partitions [27].

$$\text{TDE} = -\sum_{i=1}^{M} P_i log(P_i) \tag{3.2}$$

Although this method of computing entropy uses fixed partitioning, an adaptative partitioning alternative was also calculated, where the amplitude range that was segmented was the window's, instead of the signal's. The results for both of these entropies is displayed in Figure 3.7. However, these were very similar, so only the fixed partitioning method was used, since it displays broader changes in amplitude, as can be seen in Figure 3.7 (right).

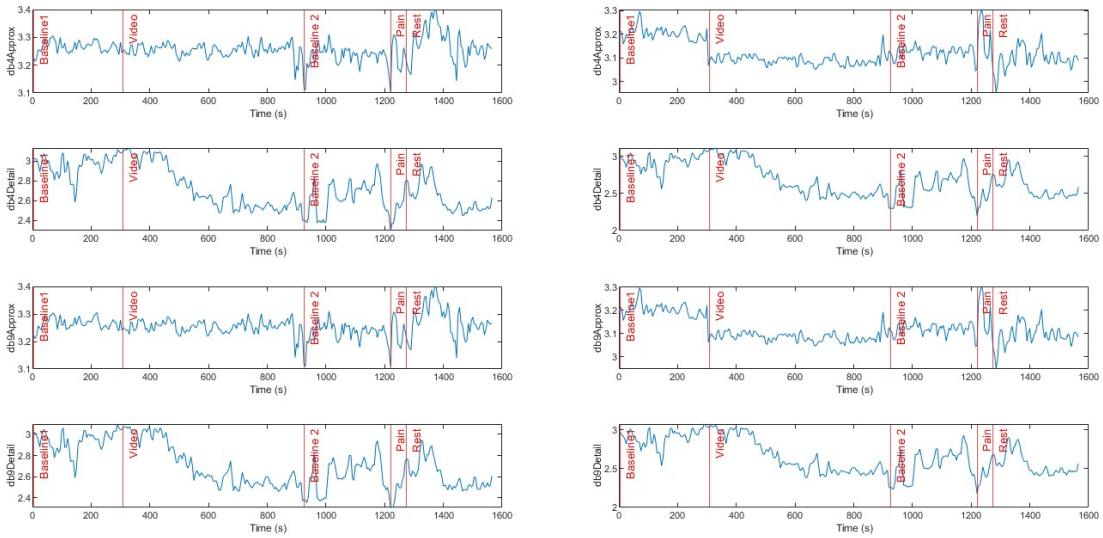

Figure 3.7: Graph of the entropy computed for the approximation and detail components of the wavelets, using adaptative (left) and fixed (right) partitioning.

3.3 Data Analysis

To do the optimal processing of the features, fifteen participants were selected at random. After standardising the features, so that they're comparable with each other, graphics of the time series of the features were plotted for each participant. An example can be seen in Figure 3.8, in which a clear change can be seen when the participant dips their hand in water, feeling pain. Since the timeseries of the mean and median of the S wave amplitude were so similar, it was decided that only the median would be analysed.

Subsequently, a statistical analysis of the participants' features was conducted using histograms, as illustrated in Figure 3.9. The histograms revealed that the distributions of the features deviate from normality, justifying the use of the median and interquartile range (IQR) as measures of central tendency and dispersion, respectively. These values were computed for each feature and considered as derived features that will integrate the analysis.

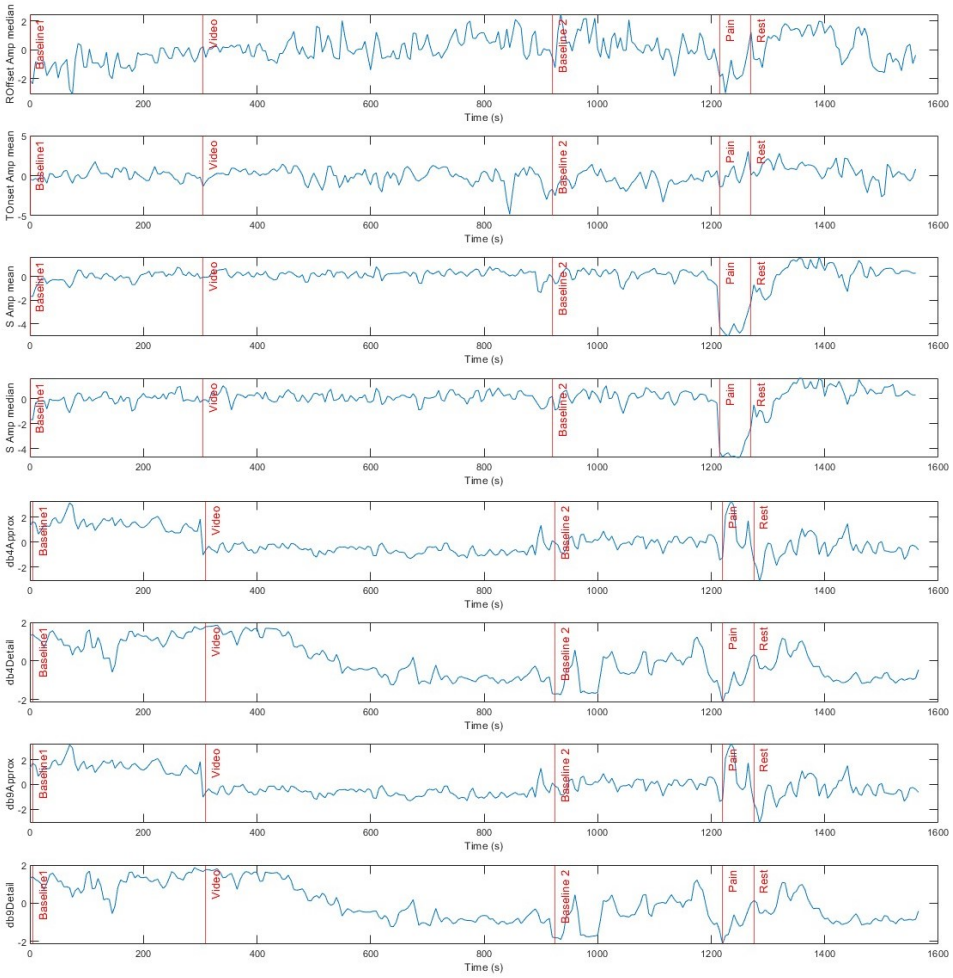

Figure 3.8: Standardised features' time series.

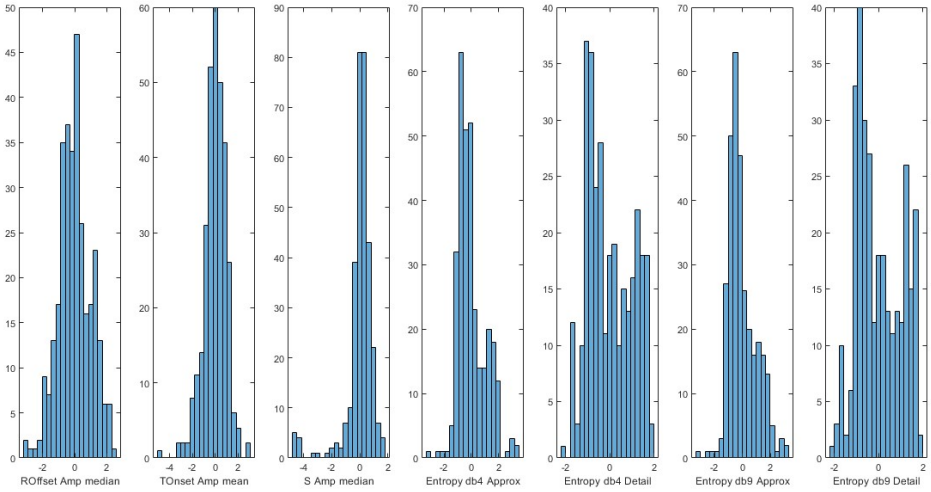

Figure 3.9: Histograms of the features.

The goal of this project is to select features that distinguish pain from no pain. To do this, it's important to guarantee the synchronization between segments, so that they're comparable. To accomplish this goal, spline interpolation was done in each step, conducing to equalize the number of points to a 60 seconds intervals. The result of this method is portrayed in Figure 3.10.

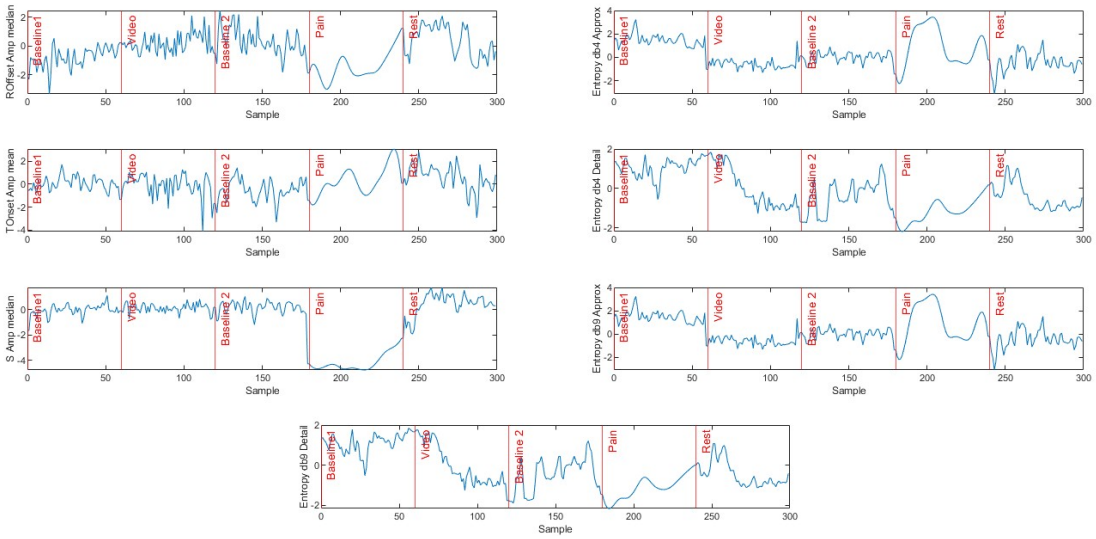

Figure 3.10: Features normalised in time after interpolation.

The features that will be analysed in this project are shown in Table 3.1. These have been computed using the median and IQR of the extracted features for each step of the protocol, as mentioned before.

Table 3.1: Features for pain classification after data preparation procedures.

Feature Source	**Features**
ECG Peaks	['Roffsetampmedian_median', 'Roffsetampmedian_iqr', 'Tonsetampmean_median', 'Tonsetampmean_iqr', 'Sampmedian_median', 'Sampmedian_iqr']
Wavelet Entropy	['Entropydb4Approx_median', 'Entropydb4Approx_iqr', 'Entropydb4Detail_median', 'Entropydb4Detail_iqr', 'Entropydb9Approx_median', 'Entropydb9Approx_iqr', 'Entropydb9Detail_median', 'Entropydb9Detail_iqr']

CHAPTER 4

Results and Discussion

4.1 Identification of Significant Features

This project employs unsupervised learning, specifically a K-means clustering algorithm, to identify the most relevant features for describing pain. This is done by generating clusters, where K is the number defined by the user for the quantity of centroids [28]. The clustering technique partitions data according to specific distance metrics, and 'cityblock' was chosen because of the way centroids are computed, that is, as the component-wise median of points in each cluster. Besides this, this metric measures the sum of absolute differences between data point coordinates, according to Equation 4.1 and assigns points to clusters based on this distance [29].

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{4.1}$$

This clustering algorithm was applied to every possible pair of features, with the number of centroids set to $K = 2$, using data from the second baseline and from the pain stimulation phase of the fifteen participants. It's important to mention that this analysis was firstly done with the first baseline but, comparing both options, the results were better using the second one, which may stem from the influence of the emotional stimuli in the pain felt by the participants.

The result of clustering applied to two features – the medians of the median of the S peak amplitude (x-axis) and of the median of the R offset amplitude (y-axis) – is displayed in Figure 4.1 (right). In it, two clusters can be seen, as well as X's in the position of the clusters' centroids. For comparison, scattering graphics were plotted for each two features, as can be seen in Figure 4.1 (left). The incorrect matches made by the clustering algorithm are circled in black.

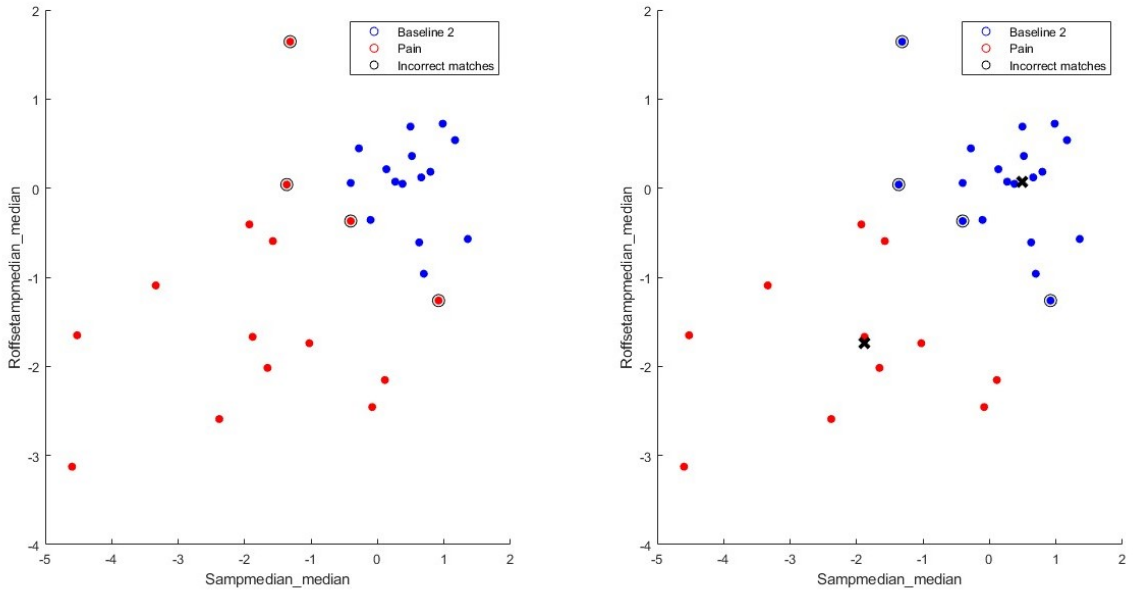

Figure 4.1: Scattering plot with data from baseline 2 and pain (left), and with the result of K-means clustering (right), when comparing the Roffsetampmedian_median and Sampmedian_median.

Only the pairs of features where clustering was done with over 80% accuracy were considered valid. A network map illustrating these pairs was plotted in Figure 4.2. In this map, each circle corresponds to a feature, while the lines connecting them showcase the combinations that met the accuracy threshold. Additionally, the size of the circles is proportional to how many times that specific feature, paired with another, contributed to successful clustering. These results suggest that these five features, Roffsetampmedian_median, Sampmedian_median, Entropydb4Approx_median, and Entropydb9Approx_median, when paired with each other, are the most effective in distinguishing pain from no pain. A possible takeaway from this result is that pain affects the intensity of ventricular polarization, since the S wave and R offset are related to it. Regarding the entropy, pain affects it the most in the low frequencies of the ECG signal, for both the db4 and db9 wavelets. This might entail that the QRS complex, as concluded before, and the fluctuations of the signal also affect pain.

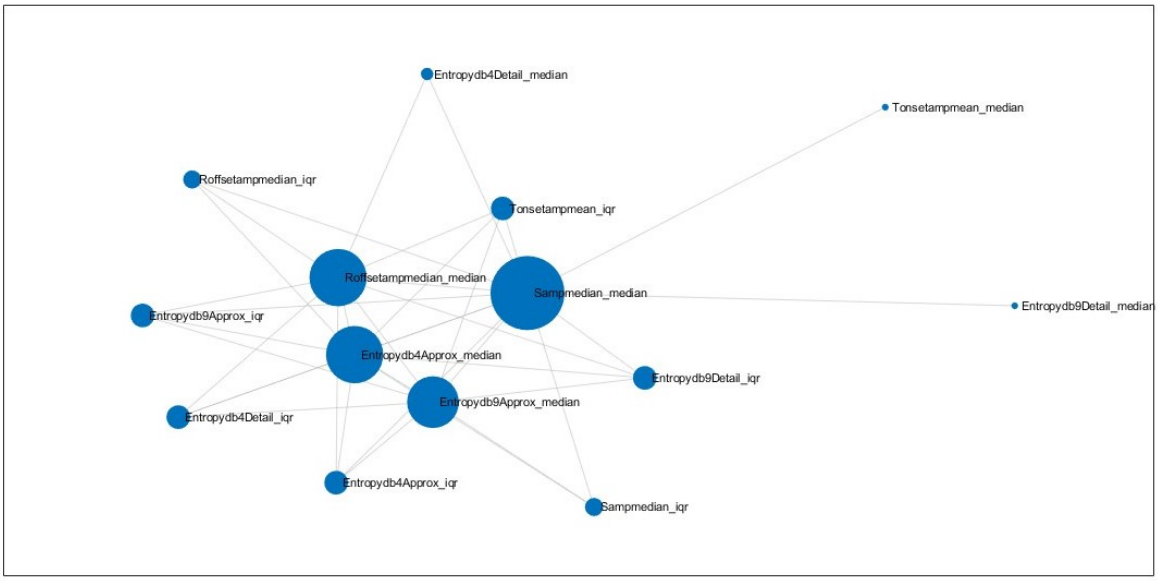

Figure 4.2: Network diagram of the pairs of features with more than 80% accuracy of clustering.

4.1.1 Multiple Comparison Test

To determine whether these features can distinguish pain from the lack of it in spite of the context, statistical analysis was performed through a multiple comparison test applied to both the baselines and to the pain phase. This was done for each participant first, using the original features that resulted in the selected ones, and then for the extracted features. Firstly, a one-way ANOVA was applied to the three groups. This test evaluates whether there are statistically significant differences between the means of each group. Afterwards, the Tukey-Kramer method, also known as the Honest Significant Difference (HSD) method, was performed as a post-hoc test to determine whether the differences detected by the ANOVA test are significant. This is done by taking the absolute value of the difference between pairs of means and dividing it by the standard error of the mean [30].

Analysing each participant's data individually, regarding the median of the R offset amplitude, the means of the pain and baseline phases were different for 100% of the participants, with 20% of the baselines' means being coincident. For the median of the S amplitude, there was 93% of successful pain and baselines' means distinction, where 43% corresponded to similar means when comparing the baselines. Lastly, the entropy of the approximation of the db4 and db9 wavelets acquired analogous results, with 80% of the participants' pain and baselines' means being significantly different. 25% of these participants' baselines 1 and 2 had similar means for both of these features. These results suggest that, for most participants, these features are successful in describing pain. Additionally, it's viable that the video watched by each participant has affected not only the entropy of the ECG and the R offset, but mostly the amplitude of the S wave. In Figures 4.3 to 4.6, examples of this multiple comparison tests are displayed for each feature.

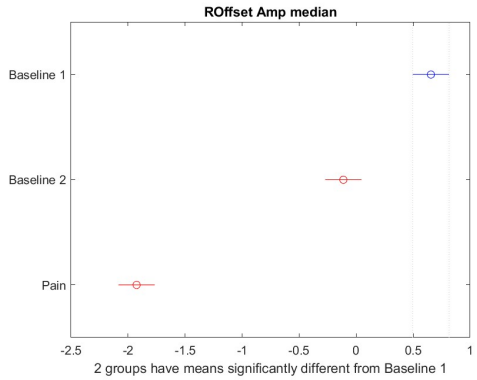

Figure 4.3: Multiple comparison test of the R offset amplitude for a participant's baselines and pain phase.

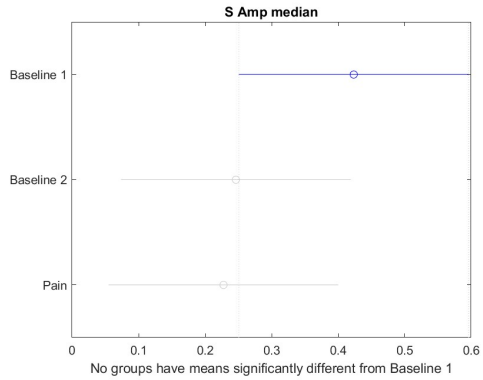

Figure 4.4: Multiple comparison test of the S wave amplitude for a participant's baselines and pain phase.

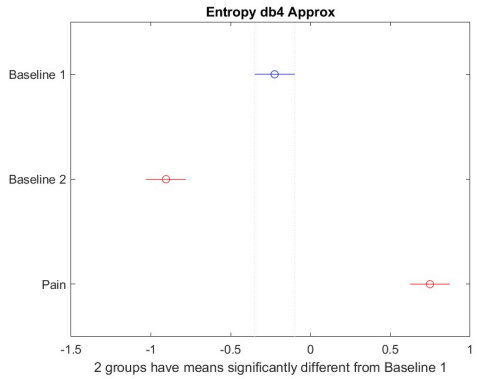

Figure 4.5: Multiple comparison test of the entropy of the db4 approximation for a participant's baselines and pain phase.

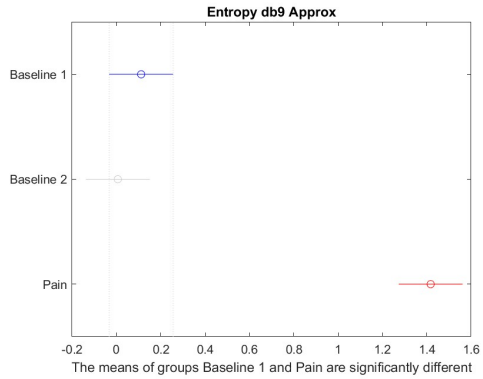

Figure 4.6: Multiple comparison test of the entropy of the db9 approximation for a participant's baselines and pain phase.

On the other hand, the result of this test for the selected extracted features is portrayed in Figure 4.7. As can be seen, the baselines' means are significantly different from the pain phase's for all the selected features, which supports the idea that pain has an impact on both the S wave and entropy. Furthermore, the means of the baselines are similar in all the graphics, which means both can be used to describe pain. However, there's a slight difference between the baselines for the amplitude of the S wave. This might be due to the visual stimulus having an impact on the ventricular polarization.

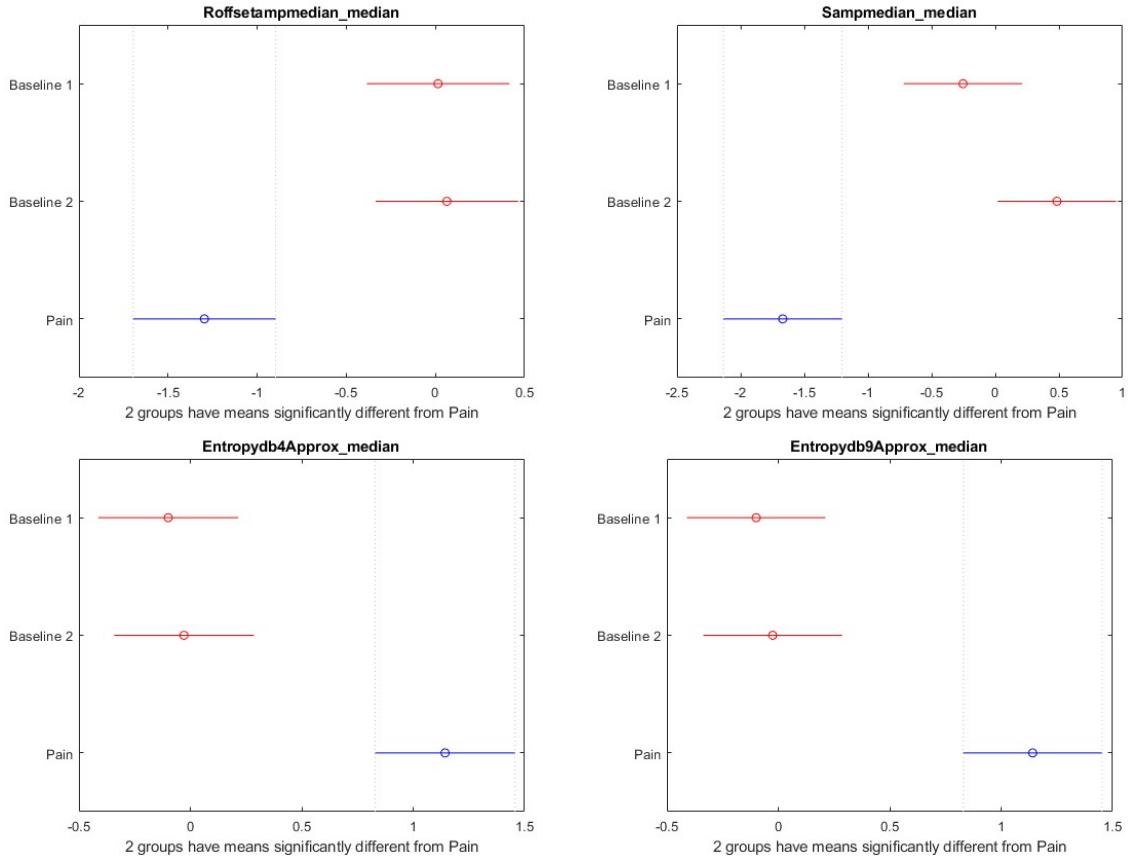

Figure 4.7: Multiple comparison test (ANOVA + HSD) applied to both the baselines and to the pain phase, for each of the selected features.

4.2 K-Nearest Neighbours

In order to validate the results of this project, a Classification Learner algorithm, particularly, a KNN classifier was employed on the selected features. This classifier is widely used due to its simplicity and effectiveness but, for this project, it was chosen because of its classification technique's similarity to the clustering algorithm. KNN classification starts with the determination of the K-nearest neighbours, which is done by computing the distance between an unknown example q, that must be classified, and x_i training samples. This distance is described by Equation 4.2, which is a summation for all the features in F, with w_f being the weight for each feature and δ a specific distance metric [31]. Then, a majority voting is applied and q is characterised as the class of the majority of its K-nearest neighbours [32].

$$d(q, x_i) = \sum_{f \epsilon F} w_f \delta(q_f, x_{if}) \tag{4.2}$$

To create this classifier, MATLAB's Classification Learner app was used, along with the processed features from the aforementioned fifteen participants. For the distance metric, Euclidian distance was chosen. This metric is computed as the square root of the sum of squared differences between the elements of both vectors, according to Equation 4.2.

$$\delta(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{4.3}$$

The remaining hyperparameters were selected using Bayesian optimisation. The goal of this optimisation is to find the ideal combination of parameters that minimizes cross-validation loss. After performing a search with 30 iterations, the optimised parameters were selected and used in the model. These are portrayed in Table 4.1.

Table 4.1: Hyperparameters optimisation search range and results.

Hyperparameters	**Search Range**	**Optimised Values**
Number of neighbours	1-15	2
Distance weight	Equal, Inverse, Squared inverse	Equal
Standardised	True, False	False

Once the hyperparameters were selected, the model was trained using the fifteen participants' features as the training set. This was done through fivefold cross-validation, that uses one fifth of the data to evaluate the model in each iteration. Once the KNN model was complete, the processing pipeline was applied to the remaining 36 participants' data to create a training dataset. Then, the model was applied to this set, predicting whether a value had derived from the baseline or pain stimulation phase. To compare these predictions with the true values, a confusion matrix was plotted in Figure 4.8. This allows for the computation of the following variables:

- **True Positives (TP):** pain observations correctly predicted.
- **False Positives (FP):** pain observations incorrectly predicted.
- **True Negatives (TN):** baseline observations correctly predicted.
- **False Negatives (FN):** baseline observations incorrectly predicted.

These variables can be used to calculate metrics for classification models' evaluation, such as sensitivity, precision, accuracy, and F-score. Sensitivity, also known as True Positive Rate, is obtained by dividing the number of TP for the sum of TP and FN. On the other hand, specificity, or True Negative Rate, corresponds to the number of TN divided by the sum of TN and FP. Meanwhile, precision is computed as the ratio between the TP and the total number of pain predictions, while accuracy is computed as the sum of TP and TN divided by the total number of observations [33].

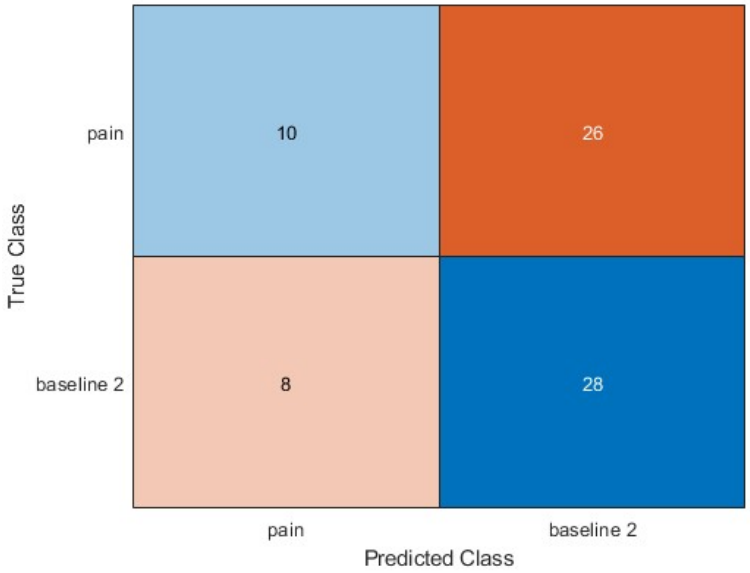

Figure 4.8: Confusion matrix for KNN model testing.

Table 4.2: Features for pain classification after data preparation procedures.

Evaluation Metric	**Percentage**
Sensitivity	27.8%
Specificity	77.8%
Precision	55.6%
Accuracy	52.8%

As shown in Table 4.2, the KNN model achieved a low sensitivity of 27.8%. This implies that the model was unsuccessful in correctly identifying true cases of pain, contesting the validity of these features for pain description. In contrast, the specificity was relatively high at 77.8%, indicating that the model was more effective in correctly identifying the absence of pain. Additionally, an accuracy of 52.8% was obtained, which indicates that approximately half of the predictions made on the test set were correct. Meanwhile, precision reached 55.6%, suggesting that slightly more than half of the observations classified as pain were true positives.

CHAPTER 5

Conclusion

Perspetivas futuras: - experimentar outros tipos de clustering
- fazer undersampling de outra maneira (SMOTE?)
- experimentar outros tipos de entropia
- aumentar numero de participantes
- usar dados de dor nao induzida a pessoas saudáveis

References

[1] D. D. Ridder, D. Adhia, and S. Vanneste, *The anatomy of pain and suffering in the brain and its clinical implications*, Nov. 2021. DOI: 10.1016/j.neubiorev.2021.08.013.

[2] S. Coninx and P. Stilwell, "Pain and the field of affordances: An enactive approach to acute and chronic pain", *Synthese*, vol. 199, pp. 7835–7863, 3-4 Dec. 2021, ISSN: 15730964. DOI: 10.1007/s11229-021-03142-3.

[3] S. M. Nugent, T. I. Lovejoy, S. Shull, S. K. Dobscha, and B. J. Morasco, "Associations of pain numeric rating scale scores collected during usual care with research administered patient reported pain outcomes", *Pain Medicine (United States)*, vol. 22, pp. 2235–2241, 10 Oct. 2021, ISSN: 15264637. DOI: 10.1093/pm/pnab110.

[4] J. Bielewicz, B. Daniluk, and P. Kamieniak, "Vas and nrs, same or different? are visual analog scale values and numerical rating scale equally viable tools for assessing patients after microdiscectomy?", *Pain Research and Management*, vol. 2022, 2022, ISSN: 12036765. DOI: 10.1155/2022/5337483.

[5] J. Qin, *Entropy-based feature extraction method from wavelet coefficient for pain intensity assessment task*, 2022.

[6] R. Fang, E. Hosseini, R. Zhang, C. Fang, S. Rafatirad, and H. Homayoun, "Survey on pain detection using machine learning models: Narrative review", *JMIR AI*, vol. 4, e53026, Feb. 2025, ISSN: 2817-1705. DOI: 10.2196/53026.

[7] S. Moscato, P. Cortelli, and L. Chiari, "Physiological responses to pain in cancer patients: A systematic review", *Computer Methods and Programs in Biomedicine*, vol. 217, p. 106 682, Apr. 2022, ISSN: 01692607. DOI: 10.1016/j.cmpb.2022.106682.

[8] S. Jayakar, J. Shim, S. Jo, B. P. Bean, I. Singeç, and C. J. Woolf, "Developing nociceptor-selective treatments for acute and chronic pain", *Science Translational Medicine*, vol. 13, 619 Nov. 2021, ISSN: 1946-6234. DOI: 10.1126/scitranslmed.abj9837.

[9] S. Raman and P. Sharma, "Self-efficacy as a mediator of the relationship between pain and disability in chronic pain patients: A narrative review", *Bulletin of Faculty of Physical Therapy*, vol. 27, 1 Dec. 2022, ISSN: 1110-6611. DOI: 10.1186/s43161-022-00101-y.

[10] M. A. Fitzcharles, S. P. Cohen, D. J. Clauw, G. Littlejohn, C. Usui, and W. Häuser, *Nociplastic pain: Towards an understanding of prevalent pain conditions*, May 2021. DOI: 10.1016/S0140-6736(21)00392-5.

[11] A. Adeboye, R. Hart, S. H. Senapathi, N. Ali, L. Holman, and H. W. Thomas, "Assessment of functional pain score by comparing to traditional pain scores", *Cureus*, Aug. 2021, ISSN: 2168-8184. DOI: 10.7759/cureus.16847.

[12] B. Alves, S. Brás, and R. Sebastião, "Decoding pain: Prediction under different emotional contexts through physiological signals", *International Journal of Data Science and Analytics*, Oct. 2024, ISSN: 2364-415X. DOI: 10.1007/s41060-024-00649-z.

[13] C. L. Robinson *et al.*, "Pain scales: What are they and what do they mean", *Current Pain and Headache Reports*, vol. 28, pp. 11–25, 1 Jan. 2024, ISSN: 1531-3433. DOI: 10.1007/s11916-023-01195-2.

[14] R. F. Rojas, N. Brown, G. Waddington, and R. Goecke, *A systematic review of neurophysiological sensing for the assessment of acute pain*, Dec. 2023. DOI: 10.1038/s41746-023-00810-1.

[15] J. Liu *et al.*, "A novel p-qrs-t wave localization method in ecg signals based on hybrid neural networks", *Computers in Biology and Medicine*, vol. 150, p. 106 110, Nov. 2022, ISSN: 00104825. DOI: 10.1016/j.compbiomed.2022.106110.

[16] J. Hampton, J. Hampton, and D. Adlam, *The ECG Made Easy E-Book: The ECG Made Easy E-Book* (Made Easy). Elsevier, 2024, ISBN: 9780323938105. [Online]. Available: https://books.google.pt/books?id=maALEQAAQBAJ.

[17] Y. Zhang *et al.*, "Abnormal recognition-assisted and onset-offset aware network for pathological wearable ecg delineation", *Artificial Intelligence in Medicine*, vol. 157, p. 102 992, Nov. 2024, ISSN: 09333657. DOI: 10.1016/j.artmed.2024.102992.

[18] H. Dogan and R. O. Dogan, "A comprehensive review of computer-based techniques for r-peaks/qrs complex detection in ecg signal", *Archives of Computational Methods in Engineering*, vol. 30, pp. 3703–3721, 6 Jul. 2023, ISSN: 1134-3060. DOI: 10.1007/s11831-023-09916-x.

[19] D. Pais, S. Brás, and R. Sebastião, "A review on the use of physiological signals for assessing postoperative pain", *ACM Computing Surveys*, vol. 57, pp. 1–39, 1 Jan. 2025, ISSN: 0360-0300. DOI: 10.1145/3685674.

[20] E. Pavlidou and M. Tsiknakis, "Multimodal pain assessment based on physiological biosignals: The impact of demographic factors on perception and sensitivity", in *Proceedings of the 11th International Conference on Information and Communication Technologies for Ageing Well and e-Health*, SCITEPRESS - Science and Technology Publications, 2025, pp. 320–329, ISBN: 978-989-758-743-6. DOI: 10.5220/0013426800003938.

[21] P. Silva and R. Sebastião, "Using the electrocardiogram for pain classification under emotional contexts", *Sensors*, vol. 23, 3 Feb. 2023, ISSN: 14248220. DOI: 10.3390/s23031443.

[22] M. Balasubramani, M. Srinivasan, W.-H. Jean, S.-Z. Fan, and J.-S. Shieh, "A novel framework for quantum-enhanced federated learning with edge computing for advanced pain assessment using ecg signals via continuous wavelet transform images.", *Sensors (Basel, Switzerland)*, vol. 25, 5 Feb. 2025, ISSN: 1424-8220. DOI: 10.3390/s25051436.

[23] A. Osadchiy, A. Kamenev, V. Saharov, and S. Chernyi, "Signal processing algorithm based on discrete wavelet transform", *Designs*, vol. 5, 3 Sep. 2021, ISSN: 24119660. DOI: 10.3390/designs5030041.

[24] J. Koh *et al.*, "Automated classification of attention deficit hyperactivity disorder and conduct disorder using entropy features with ecg signals", *Computers in Biology and Medicine*, vol. 140, p. 105 120, Jan. 2022, ISSN: 00104825. DOI: 10.1016/j.compbiomed.2021.105120.

[25] S. Maity, "Ecg heartbeat classification using wavelet packet entropy and random forest", *JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES*, vol. 18, 8 Aug. 2023, ISSN: 09738975. DOI: 10.26782/jmcms.2023.08.00002.

[26] T. Adjei, W. V. Rosenberg, V. Goverdovsky, K. Powezka, U. Jaffer, and D. P. Mandic, "Pain prediction from ecg in vascular surgery", *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 5, pp. 1–10, 2017, ISSN: 2168-2372. DOI: 10.1109/JTEHM.2017.2734647.

[27] J. Ferreira, S. Brás, C. F. Silva, and S. C. Soares, "An automatic classifier of emotions built from entropy of noise", *Psychophysiology*, vol. 54, pp. 620–627, 4 Apr. 2017, ISSN: 14698986. DOI: 10.1111/psyp.12808.

[28] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data", *Information Sciences*, vol. 622, pp. 178–210, Apr. 2023, ISSN: 00200255. DOI: 10.1016/j.ins.2022.11.139.

[29] T. B. Ricken, P. Bellmann, S. Walter, and F. Schwenker, "Pain detection in biophysiological signals: Knowledge transfer from short-term to long-term stimuli based on distance-specific segment selection", *Computers*, vol. 12, p. 71, 4 Mar. 2023, ISSN: 2073-431X. DOI: 10.3390/computers12040071.

[30] A. Nanda, D. B. B. Mohapatra, A. P. K. Mahapatra, A. P. K. Mahapatra, and A. P. K. Mahapatra, "Multiple comparison test by tukey's honestly significant difference (hsd): Do the confident level control type i error", *International Journal of Statistics and Applied Mathematics*, vol. 6, pp. 59–65, 1 Jan. 2021, ISSN: 24561452. DOI: 10.22271/maths.2021.v6.i1a.636.

[31] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers - a tutorial", *ACM Computing Surveys*, vol. 54, pp. 1–25, 6 Jul. 2022, ISSN: 0360-0300. DOI: 10.1145/3459665.

[32] M. Papanikolaou, G. Evangelidis, and S. Ougiaroglou, "Dynamic k determination in k-nn classifier: A literature review", in *2021 12th International Conference on Information, Intelligence, Systems and Applications (IISA)*, IEEE, Jul. 2021, pp. 1–8, ISBN: 978-1-6654-0032-9. DOI: 10.1109/IISA52424.2021.9555525.

[33] Ž. Đ. Vujovic, "Classification model evaluation metrics", *International Journal of Advanced Computer Science and Applications*, vol. 12, 6 2021, ISSN: 21565570. DOI: 10.14569/IJACSA.2021.0120670.