University of Molise

Department of Biosciences and Territory

Bachelor Thesis

Global and Local Prediction in Automatic detection of Atrial Fibrillation

Alex Colucci

December, 2019

Supervisor
Prof. Oliveto Rocco

Co-Supervisor
Gennaro Laudato PhD

Acknowledgement

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Abstract

Atrial fibrillation (AF) is one of the most common cardiac arrhythmia associated with an increase risk of stroke, heart failure and dementia. The estimated number of individuals with the disease is around 33 million, which is a downward forecast as the global population ages. In addition, the asymptomatic characteristic makes this arrhythmia even more severe.

The thesis describe the state of the art algorithm based on heart rate for the automatic detection of Atrial fibrillation through the use of an Electrocardiogram (ECG). The explained algorithm was used as a basis for achieving better performance by using Machine Learning techniques and algorithms globally on available medical databases. A local experimentation was then carried out, which takes into account the different physiologies of the patients and consequently groups them according to optimal discriminatory thresholds.

On the global level significant improvement were made. Instead the local level turned out to be more complicated than expected and simple improvements through a generalized least squares model couldn't be achieved. To conclude, future developments aim to create a better replica of the state of the art and carry out a comprehensive experimentation with neural networks on a global level. As for the local one, more complex regression models will have to be used to group patients together.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 What is Atrial Fibrillation?

Atrial fibrillation, also abbreviated with AF or A-Fib, is an abnormal heart rhythm that happens when electrical impulses fire off in the atria (Figure 1.1), from different spots without being organized. Characterized by rapid and irregular beating, caused by the chambers of the heart twitching [2]. This arrhythmia is associated with an increased risk of stroke, in fact the proportion of strokes associated with AF increases from 6.6%, for ages 50 to 59 years, to 36.2% for ages 80 to 89 years [3]. Other risks are heart failure and even dementia [4]. The estimated number of individuals with AF globally in 2010 was 33,5 million and as the population ages globally, the burden of AF grows [5].

The disease is classified by doctors based on how long it lasts or based on the cause. The treatment will be different for each kind [6]:

- **Paroxysmal** (holiday heart syndrome): an episode of AF, the duration of whose maybe a few minutes or a few days, but which tends to be

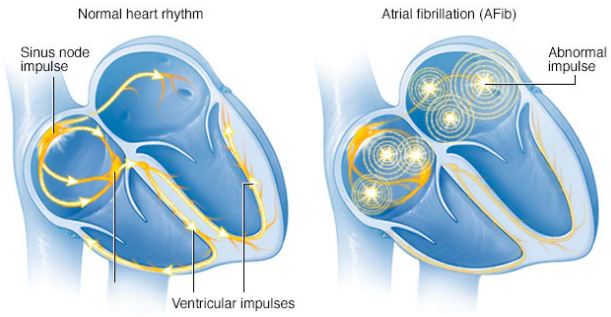

Figure 1.1: A normal heartbeat on the left, and AF heartbeat on the right. Image from mayoclinic.org

below the week. Usually, treatment is not needed;

- **Persistent**: the disease lasts longer than a week and it can stop on its own, or a specific medicine or treatment is needed. If the latter does not work, doctors opt for the electrical cardioversion, which is a low-voltage current used to reset the normal rhythm;
- **Permanent**: also called chronic, cannot be treated. The doctor decides for a long term medication to reduce the odds of associated health conditions.

1.2 Causes and Symptoms

There are many possible causes of the condition, some are controllable, others are not. Cardiovascular factors play a big role: high blood pressure, heart valve disease, congenital heart disease and even previous heart surgery. But difficulties in breathing are a key factor too, in other words, obesity and obstructive sleep apnea [7]. Alcohol consumption and tobacco smoking are associated with an increased risk of developing atrial fibrillation [8, 9]. Other factors are genetics, ageing, a sedentary lifestyle and diabetes [10, 11].

The person often feels an abnormal beating that starts to become longer and constant. There could be heart palpitations, shortness of breath, chest pain, light-headedness, or fainting [12]. But the biggest problem is that often these kind of episodes are asymptomatic [4], in fact sometimes first diagnosed when patients present a stroke [13].

1.3 Diagnosis

A doctor to diagnose AF could check your signs and symptoms, together with your medical history and conduct a different kind of tests [14]:

- **Electrocardiogram** (ECG or EKG) is the process through which a recording of the electrical activity of the patient's heart is made. To measure the electrical signals as they travel, multiple small sensors, called electrodes, are attached to the body. This test plays a key role among all the other tools used. A more in-depth explanation will be offered in Section 1.4.
- **Holter monitor** is a portable ECG device that can be carried in a pocket or even worn on a shoulder strap or a belt. The monitor will check the heart's activity for 24 hours, sometimes even longer. It is a

common practice to utilize the device when there is a strong suspect about a Paroxysmal-AF but an ECG during an office visit detects only a regular rhythm.

- **Event recorder** is another kind of ECG portable device that is meant to monitor the heartbeat over a few weeks to a few months. When the patient feels a symptom, then the button should be pressed to let the device memorize an ECG strip of the preceding few minutes and following few minutes.

- **Echocardiogram** is a non-invasive test that uses ultrasound waves to scan the heart and get moving pictures of the organ. The doctors aim to find problems in the valves, in the size of the left and right atrial or more general structural heart disease or blood clots.

- **Blood tests** are used to check any thyroid problems or other substances in the patient's blood that may lead to AF.

- **Stress test** can help the doctor in the task of finding AF. The reason is that some individual with the disease do well in normal activity, but not with exertion. Moreover, the nature of the symptoms can be understood.

- **Chest X-ray** help to see the condition of the lungs and heart of a specific patient. In general, it's used if a pulmonary cause of AF is suggested or if conditions like congestive heart failure are suspected.

The first type of test, the ECG, is an investigation performed routinely whenever an irregular heartbeat is suspected. And it can be done in the office and later even with a portable device, thus it's a relevant tool through which an automatic detection of atrial fibrillation can be implemented.

1.4 Diagnosis through the analysis of ECG

Electrocardiography produces an electrocardiogram (ECG), namely a recording which is a graph where the x-axis represents the time and the y-axis represents the voltage, of the electrical activity of the heart using electrodes placed on the skin [15, p.74]. In this way, small electrical changes can be detected, that are the normal consequences of cardiac muscle depolarization followed by a re-polarization during each cardiac cycle (Figure 1.2).

Normally the number of electrodes attached to the patient's limbs and on the surface of the chest is 10, this allows to form 12 ECG leads. Thus the

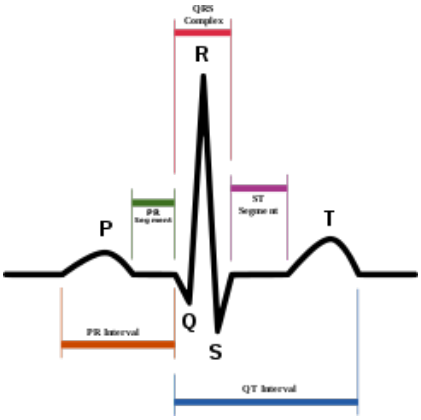

Figure 1.2: Cardiac cycle divided into different components. P, QRS and T.

Figure 1.3: ECG of a heart with atrial fibrillation on top and with normal sinus rhythm on the bottom.

overall magnitude of the electrical potential of the heart can be measured from twelve different angles (leads).

A single cardiac cycle can be divided into different components as in (Figure 1.2). The first is called P wave, which represents the depolarization of the atria. The second one is the QRS complex, that symbolizes the ventricles' depolarization. To finish with the T wave, which represents the re-polarization of the ventricles [15, p.80].

Knowing all this, to find atrial fibrillation heartbeats through the electrocardiogram is sufficient to run an investigation on the absence of P waves with disorganized electrical activity in their place and irregular $R-R$ intervals caused by irregular conduction of impulses to the ventricles [16]. Furthermore, problems over fast heart rates arise since A-Fib may look more regular, which could make it indistinguishable from other supraventricular tachycardias or ventricular tachycardia [17]. Besides QRS complexes should be quite narrow because it means that they are initiated by a normal flow of electrical activity through the intraventricular conduction system. Otherwise wide complexes are disquieting for ventricular tachycardia, albeit in cases where there is a disorder with the conductions system, wide QRS complexes may

be present in A-fib with a rapid ventricular response. A good example is shown in (Figure 1.3).

In Chapter 3 and Chapter 4, two different automatic detection of AF approaches are described based on two public datasets offered by PhysioNet [18], respectively MIT-BIH Atrial Fibrillation Database [19] and Long Term AF Database [20].

Chapter 2

State of the art

2.1 Introduction

In this Chapter, an automatic approach to detect Atrial Fibrillation is analysed. The state of the art is based on different public datasets offered by PhysioNet [18], among which MIT-BIH Atrial Fibrillation Database [19] and Long Term AF Database [20] are used. The method is based on ECG, whose explanation has been given in Section 1.4. The reason that lies behind the use of ECG, is its intrinsic simplicity, that cannot be found in methods like blood tests, chest x-ray, etc.

2.1.1 Description of the MIT-BIH AFDB

The database includes 25 long-term ECG recordings of patients with atrial fibrillation, which is mostly paroxysmal. Each record is 10 hours in duration and contains two ECG signals sampled at 250 samples per second with $12-bit$ resolution over a range of ± 10 millivolts. The signals files `.dat` are available only on 23 records. But all of the records have `.atr` and `.qrs` annotations files. The former contains information about the kind of rhythm: atrial fibrillation, atrial flutter, junctional rhythm or other rhythms. The latter contains unaudited beat prepared using an automated detector and have not been corrected manually. In some cases, manually corrected beat annotations files `.qrsc` are present.

2.2 The best approach proposed in literature

Most of the algorithms work on the processing of the ECGs components (P wave, QRS complex, ...) and the poorly coordinate atrial activation (AA) of

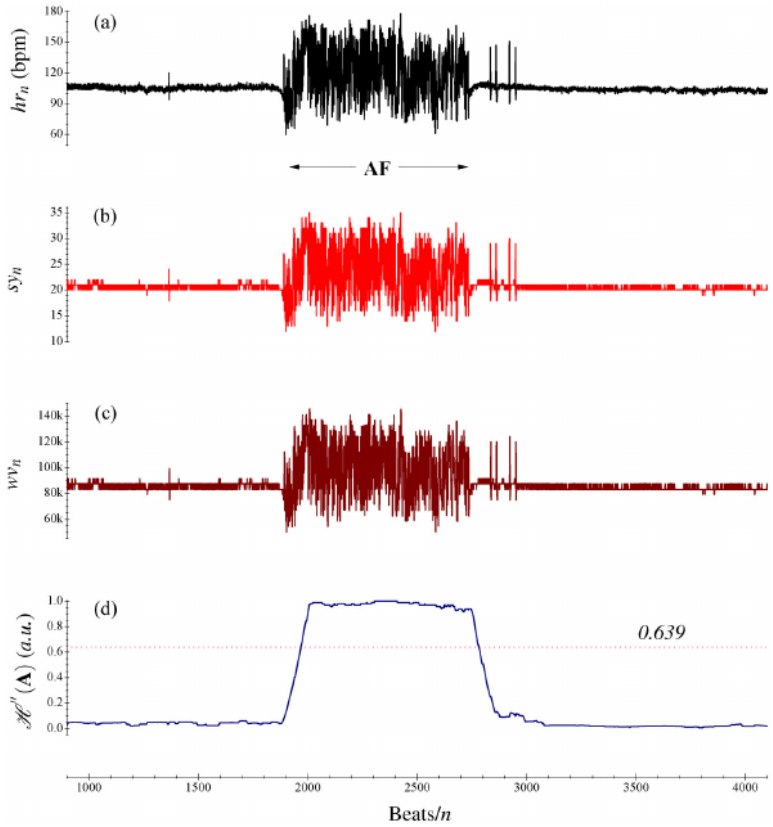

Figure 2.1: Application of the method to detect AF. (a) is the original sequence hr_n; (b) is the symbolic dynamic sy_n; (c) the word sequence wn_n; (d) the distribution of $\mathcal{H}''(\mathbf{A})$.

heart and rapid cardiac beating. Although these pieces of information can lead to the identification of Atrial Fibrillation, noise must be taken into consideration. Especially with P waves which in general is of very low-intensity magnitude. Whereas the approaches based on the RR interval (R wave peak to R wave peak) irregularity, nonetheless the component is a more prominent feature of ECG and thus less subject to noise, tend to be quite complicated and not so efficient to make them suitable for real-time applications [1, p. 2]. Examples of noteworthy methods based on RRI are the Petrėnas, et al [21, 2015] and Lee, et al [22, 2013]. The former is characterized by the use of ectopic beat filtering, bigeminal suppression and signal fusion, while the latter focus on time-varying coherence functions and Shannon entropy.

A real-time and low-complexity but robust method for the discrimination of AF episodes is taken as reference. The algorithm is composed of three steps and is based on the heart rate (HR) [1, p. 2].

2.2.1 Heart rate sequence

Let hr_n be the heartbeat rate sequence obtained from,

$$hr_n = 60\,\mathrm{s} \cdot \frac{250}{R_n - R_{n-1}} \tag{2.1}$$

where 60 are the seconds, R_n is the sequence that denotes the R peak in the QRS complex and 250 is the number of samples per second. From an implementation point of view, here one bpm is lost. The `.qrs` files contain a registration where the first sample does not represent the R peak, therefore all the part before the first R peak cannot be used to compute the first bpm. An example of the sequence hr_n can be found in (Figure 2.1 (a)). The following implementation of the function is done in Python 3.7

```
def compute_bpm(qrs_ann_list, frequency=250):
 bpm_list = list()

 for i in range(0, len(qrs_ann_list) - 1):
   dist = qrs_ann_list[i + 1] - qrs_ann_list[i]
   bpm = 60 / (dist / frequency)
   bpm_list.append(bpm)

 return bpm_list
```

2.2.2 Symbolic dynamics of hr_n sequence

Let sy_n denote a symbolic dynamics that encodes the information of hr_n to a series with fewer symbols, where the mapping function is given by [1, p. 3],

$$sy_n = \begin{cases} 63 & \text{if } hr_n \geq 315 \\ \lfloor hr_n/5 \rfloor & \text{otherwise} \end{cases} \tag{2.2}$$

where $\lfloor \cdot \rfloor$ is a floor operator. In this way the raw sequence hr_n is transformed in a sequence $sy_n \in [0, 63]$, with 64 instantaneous states (Figure 2.1 (b)). Here below the implementation

```
def compute_sy(bpm_list):
  return list(map(lambda x: 63 if x >= 315 else floor(x / 5),
                                  bpm_list))
```

2.2.3 History sequence of sy_n

A 3-symbols template can be applied to get a window of information that acts as a history (Figure 2.1 (c)), in this case on 3 successive symbols. Through

a novel operator-defined below [1, p. 3], the word value can be calculated.

$$wv_n = (sy_{n-2} \times 2^{12}) + (sy_{n-1} \times 2^6) + sy_n \tag{2.3}$$

A major intrinsic property to be explained is the following,

$$|wv| = |sy| - 2 = |hr| - 2 = |qrs| - 3 \tag{2.4}$$

where wv denotes the set of words in a specific moment. The set qrs is composed of all the QRSs complexes provided by the ECG, in this case in the `.qrs` files. To sum up, a tiny bit of information is lost, precisely 1 R peak to compute the heartbeat sequence hr and 2 bpm to compute the word sequence wv. Here follows the implementation of the above

```
def compute_wv(sy_list):
  wv_list = list()

  for i in range(2, len(sy_list)):
    wv = (sy_list[i - 2] << 12) + (sy_list[i - 1] << 6) +
                                   sy_list[i]
    wv_list.append(wv)

  return wv_list
```

2.2.4 Shannon entropy

A coarser version of Shannon entropy is employed to discriminate the AF arrhythmias (Figure 2.1 (d)). Without loss of generality, let $\mathbf{A} = (A|P)$ denote a dynamic system. The unique elements in this set can be defined as $A = \{a_1, \ldots, a_k\}$ with the interrelated probability set $P = \{p_1, \ldots, p_k\}(1 \leq k \leq N)$, where N is the total number of elements and k are the unique elements in space $\mathbf{A}$. Each element a_i has the probability $p_i = N_i/N(0 < p_i \leq 1, \sum_{i=1}^{k} p_i = 1)$, where N_i is the total number of the specific element a_i in space $\mathbf{A}$. Hence the coarser version of Shannon entropy can be defined to quantitatively calculate the information size of wv_n,

$$\mathcal{H}''(\mathbf{A}) = -\frac{k}{N\, log_2 N} \sum_{i=1}^{k} p_i\, log_2 p_i \tag{2.5}$$

The dynamic $\mathcal{A}$ is characterized by a bin size of $N = 127$ consecutive word elements from wv_{n-126} to wv_n. By defining the characteristic set A and the corresponding probability set P, the entropy $\mathcal{H}''(\mathbf{A})$ can be calculated. A specific cardiac beat hr_n is labelled as AF if the coarser entropy meets or

exceeds a discrimination optimal threshold equal to 0.639. The threshold was obtained through an investigation of various thresholds in the range $[0.0, 1.0]$ with an increment of 0.001 from the receiver operating characteristic (ROC) on training databases. The computational challenges that are found in the Equation 2.5 can be overcome with a pre-calculated map of $-\frac{1}{log_2 N} p_i \, log_2 pi_i$ [1, p. 4]. Here follows an implementation where a constant value 1000000 it's used to get decimal floating points as integers through a floor operator.

```
def get_pimap(n, cons):
  pi = lambda i: int(-cons / log(n, 2) * i * log(i, 2))
  return [pi(p / n) for p in range(1, n + 1)]

pi_map = get_pimap(127, 1000000)

def compute_entropy(wv_list):
  nu = list([0] * 127)
  sh2 = list()
  for i in range(0, len(wv_list)):
    nu.pop(0)
    nu.append(wv_list[i])
    # Number of occurrences of an element are counted
    a = dict()
    for j in nu:
      a[j] = (a[j] + 1) if j in a else 1

    k = len(a)
    sh1 = sum([pi_map[a[element] - 1] for element in a])
    sh2.append(k / 127000000 * sh1)

  return sh2
```

2.3 Results and comparisons

Performance metrics

The work under consideration measures the performances using sensitivity (Se), specificity (Sp), positive predictive value (PPV), and overall accuracy (ACC) [1, p. 6].

$$\begin{aligned} Se &= \frac{TP}{TP+FN}, & PPV &= \frac{TP}{TP+FP} \\ Sp &= \frac{TN}{TN+FP}, & ACC &= \frac{TP+TN}{TP+TN+FP+FN} \end{aligned} \tag{2.6}$$

where TP stands for true positives, TN true negatives, FP false positives and FN false negatives.

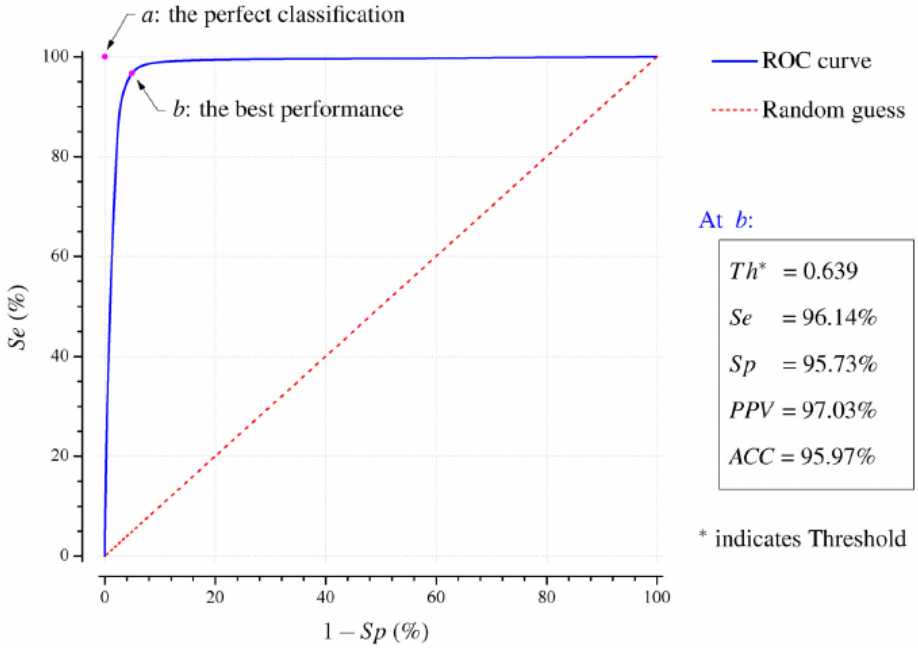

Figure 2.2: Receiver operating characteristic when the training set of LTAFDB database is applied with threshold values from 0.0 to 1.0 in increments of 0.001. The calculated value for the area under the blue curve is 0.9845 [1].

Training phase

The training phase was executed on LTAFDB which consists of 84 long-term ECG recordings (commonly 24 to 25 hours duration) of patients with paroxysmal or persistent/permanent AF. The number of cardiac beats is around 9 million of which 59.2% are annotated as AF. The threshold, as already mentioned in the previous section 2.2.4, is tested from 0.0 to 1.0 with an increment of 0.001. ROC curve is defined based on the metrics Se and $1-Sp$, where a point on the graph is formed by the couple $(Se; 1-Sp)$ of a specific threshold. Thus the best-case scenario, namely perfect classification, is the couple $a = (Se = 1; 1 - Sp = 0)$, hence the best performance can be found trivially using the Euclidean distance of a generic point from the point a. In (Figure 2.1), the best performance point b is found at threshold 0.639, with distance 0.0576 from the perfect classification, an area under the ROC curve that is 0.9845 and the corresponding values of Se, Sp, PPV and ACC are 96.14%, 95.73%, 97.03% and 95.97% respectively [1]. In other words, a slight improvement is made compared to RRI based method Zhou, et al[23].

Testing phase

This phase uses the threshold 0.639 across all testing databases: AFDB, MITDB[24] and NSRDB[25] set. A complete overview of the results of the state of the art method explained and others can be found in Table 2.1. AFDB[1] is experimentation without the records `00735` and `03665` for which the `.hea` files are not existent. While AFDB[2] excludes the record `04936` and `05091` because many incorrect manual AF annotations are contained [22]. To be sure about experimentation, edge cases are needed. Hence dataset like MITDB which contains many coexisting various types of complex arrhythmias and NSRDB without any AF annotation, are perfect for this purpose.

Table 2.1: Classification performance of different methods based on three different testing databases [1, p. 8].

Method	**Feature**	**Year**	**Database**	**Results**			
				SE(%)	**SP**(%)	**PPV**(%)	**ACC**(%)
Zhou, et al[1]	HR	2015	AFDB	97.37	98.44	97.89	97.99
			AFDB[1]	97.31	98.28	97.89	97.84
			AFDB[2]	98.43	98.46	97.92	98.45
			MITDB	97.83	87.41	47.67	88.51
			NSRDB	NA	99.68	NA	NA
Petrênas, et al[21]	RRI	2015	AFDB	97.12	98.28	-	-
			AFDB[1]	97.1	98.1	-	-
			AFDB[2]	98.0	98.2	-	-
			MITDB	97.8	86.4	47.67	88.51
			NSRDB	NA	98.6	NA	NA
Zhou, et al[23]	RRI	2014	AFDB	96.89	98.25	97.62	97.67
			AFDB[1]	96.82	98.06	97.61	97.50
			AFDB[2]	97.83	98.19	97.56	98.04
			MITDB	97.33	90.78	55.29	91.46
			NSRDB	NA	98.28	NA	NA
Lee, et al[22]	RRI	2014	AFDB[2]	98.22	97.68	-	97.91
			MITDB	91.1	89.7	-	-
			NSRDB	NA	99.7	NA	NA

[1] Records `00735` and `03665` excluded.
[2] Records `04936` and `05091` excluded.
'NA' indicates not applicable because there is no beat with AF reference annotation in this database.

Thus the method performs statistically better than the others [1, p. 11] with a very low computational complexity [1, p. 14]

Chapter 3

Global prediction

3.1 Introduction

In this Chapter, an attempt to improve the state of the art through Machine Learning (ML) techniques is made. Specifically different algorithms are applied on datasets enriched by morphology data to detect atrial fibrillation.

3.1.1 State of the art replication

Before proceeding with the actual experiment, a replication of Zhou, et al[1] is needed. Unfortunately the final results don't match between the original work and the replication. The latter was made with Python 3.7 and the most popular libraries for data science, such as Numpy and Pandas. Wfdb library to manipulate the different databases introduced in Section 2.1. The core part of the work has been shown in pieces in Chapter 2 that explains the steps of the algorithm. In Table 3.1 the performance of the replication is

Table 3.1: State of the art algorithm replication performance.

Method	**Database**	**Results**			
		SE(%)	**SP**(%)	**PPV**(%)	**ACC**(%)
Zhou, et al[1]	AFDB	97.37	98.44	97.89	97.99
A	AFDB	96.03	97.49	96.59	96.87
	AFDB[3]	96.04	97.50	96.60	96.88
B	AFDB	95.99	97.50	96.60	96.86
	AFDB[3]	96.00	97.50	96.62	96.86
C[4]	AFDB	96.03	97.53	96.64	96.89
	AFDB[3]	96.04	97.53	96.66	96.90

[3] File `.qrsc` (qrs complexes corrected manually) used when available.
[4] Hybrid heartbeats rate were introduced. 584 *hr* not classified.

Table 3.2: Number of beats comparison between state of the art and replication.

Method	Database	AF	NON-AF	TOTAL	Difference from SOA*
Zhou, et al[1]	AFDB	519687**	701887**	1221574	0
B	AFDB	516515	704969	1221484	−90 beats
B	AFDB[3]	518082	705013	1223095	+1521 beats

* Difference from state of the art method.
** [1, p. 9].

reported. The predicted values were compared with an oracle. To define the matching oracle oa_n of a specific record, a binary sequence bs_k was used to keep track of the samples that are AF (bit 1) and non-AF (bit 0), between the peaks R_i and R_{i+1}. The correct labels were obtained from the `.atr` files. Then the percentage of AF bit in the interval RR was counted,

$$\text{AF}\% = \frac{\#\text{ of ones}}{RR\text{ length}} \tag{3.1}$$

In order to be able to carry out as complete trial as possible, the oracle oa_n was defined based on the percentage in three different ways:

- Method A: oa_n is AF $\iff$ $AF\% = 1$, else non-AF
- Method B: oa_n is AF $\iff$ $AF\% > 0.5$, else non-AF
- Method C: oa_n is $AF\%$

In method C 584 beats were not classified, because of hybrids (not 1 and not 0). In Table 3.2 the number of AF and non-AF beats classified per method are shown. An artefact was introduced in the implementation of the algorithm or the definition of the oracle. Further investigation is needed, but the difference between the methods applied to the same database is negligible.

3.2 Applying machine learning

During this phase, machine learning algorithms were applied on enriched feature datasets using Weka 3.8. Method B on the corrected AFDB is the base of the experiment. Furthermore, to make the comparison more immediate, Matthews correlation coefficient (MCC), a measure of the quality of binary classification, was introduced

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{3.2}$$

From here on, bespoke entropy be_n is introduced on each the dataset and is defined as follows. Let hr_n be the heartbeat rate sequence.
1 step. Filter the hr_n sequence by labelling beat 1 if is stable otherwise 0 or 2:

$$\mathrm{x_n} = \begin{cases} 0, & \text{if } hr_n \leq 50 \\ 1, & \text{if } 50 < hr_n < 120 \\ 2, & \text{otherwise} \end{cases} \tag{3.3}$$

2 step. Count number of stable beats (1) in a window of 10 elements $[x_{n-9}, x_n]$:

$$\mathrm{be_n} = \sum_{i=n-9}^{n} [x_i = 1] \tag{3.4}$$

Explicit and encoded entropy dataset

As already explained in 2.2.4, hr_n is labelled as AF if the coarser entropy meets or exceeds a discrimination optimal threshold equal to 0.639. This will be referred as encoded entropy, while the value itself of the Shannon entropy is the explicit entropy.

The first experiment consisted of using a dataset with encoded/explicit entropy, custom entropy and of course the oracle as a label (Table 3.3). In the case of the explicit dataset (explicit and custom entropy with the oracle), the overall performance was lower than in the case of the replication. But an outstanding increment was obtained in the case of the `logistic` algorithm in term of SE around (+0.92). The MCC though of the `logistic` was quite low compared to the replica. All in all, the expected result, since the threshold was not used to discriminate against the Shannon entropy.

As for the encoded case, an increase in performance was hypothesised, which proved to be true. Five out of seven algorithms were able to obtain an increase in MCC (+0.14) compared to the replica. It is important to underline that `j48`, `ibk`, `random forest` and `reptree` had the same performances and `logistic` in contrast to the first case, was the algorithm with the worst performance.

Instead for the last case where both entropies were used, a performance to report is certainly that of the algorithm `bayesnet` with an increment in terms of MCC equal to (+0.06). In any case, the experiment in question was placed in the middle between the two previous ones.

Experimentation with the encoded dataset can be considered successful. If we consider that it has been applied on a basis that does not reach Zhou, we can assume that applying it on the latter, we should achieve improvements.

Table 3.3: Machine learning algorithms applied on dataset with explicit and encoded entropy, compared with the replication of Zhou and Zhou, et al [1] itself.

Dataset	**Algorithm**	**Results**				
		SE(%)	**SP**(%)	**PPV**(%)	**ACC**(%)	**MCC**(%)
Explicit	j48	96.01	97.19	96.21	96.69	93.22
	ibk	94.95	96.09	94.74	95.60	91.01
	logistic	**96.93**	96.58	95.47	96.73	93.33
	bayesnet	96.05	96.72	95.61	96.43	92.71
	adaboostm1	95.57	97.34	96.39	96.59	93.01
	randomforest	95.19	96.10	94.78	95.71	91.24
	reptree	95.38	97.32	96.36	96.49	92.83
Encoded	j48	96.48	97.27	96.33	96.93	**93.73**
	ibk	96.48	97.27	96.33	96.93	**93.73**
	logistic	96.03	97.04	96.01	96.61	93.06
	bayesnet	96.07	97.45	96.56	96.86	93.59
	adaboostm1	96.03	97.53	96.66	96.89	**93.64**
	randomforest	96.48	97.27	96.33	96.93	**93.73**
	reptree	96.48	97.27	96.33	96.93	**93.73**
Both	j48	95.98	97.19	96.21	96.68	93.20
	logistic	96.67	96.91	95.87	96.81	93.48
	bayesnet	96.05	97.53	96.65	96.90	**93.65**
	adaboostm1	95.57	97.34	96.39	96.59	93.01
	randomforest	95.15	96.05	94.71	95.66	91.14
	reptree	95.32	97.28	96.31	96.45	92.73
AFDB	**replication**	96.01	97.51	96.62	96.87	93.59

'NA' indicates not applicable because there the metric is not offered by the reference [1].

Fast Fourier Transform and AR Coefficients

The explicit and encoded datasets of the previous experimentation were used as a base, on which signal analysis was done excluding records `03665` and `00735` (no `.hea` file). Each interval $[R_i, R_{i+1}]$ was divided into two blocks. For every block, Fast Fourier Transform was applied to obtain 16 values and 4 Autoregressive model's coefficients estimated through Yule–Walker equations.

Table 3.4 shows results of the process. First clarification to do is that some algorithms were removed for performance issues.

In the case of explicit FFT dataset, 16 values were added to the previous dataset with explicit entropy and custom entropy. A several drop in overall performance, compared to the explicit dataset, occurred for most of the algorithms except for `logistic` and `adaboostm1` where it was just slight. The `logistic` tended to find a great number of true positives and in this case it reached its best performance in terms of SE that was equal to 97.00%, facilitated by a dataset richer in information.

But as for encoded FFT dataset, the state of the art is outdated again but the increment in terms of MCC was just a marginal (+0.06) with the

Table 3.4: Machine learning algorithms applied on the dataset with Fast Fourier Transform and AR coefficients, compared with the replication of Zhou and Zhou, et al [1] itself.

Dataset	**Algorithm**	**Results**				
		SE(%)	**SP**(%)	**PPV**(%)	**ACC**(%)	**MCC**(%)
Explicit FFT	j48	84.38	95.48	93.88	90.47	80.93
	logistic	**97.00**	95.73	94.92	96.30	92.56
	adaboostm1	95.49	97.06	96.39	96.35	92.63
	randomforest	88.78	96.82	95.83	93.19	86.36
	reptree	84.82	95.06	93.39	90.44	80.82
Encoded FFT	j48	89.01	93.36	91.68	91.40	82.61
	logistic	96.27	96.61	95.89	96.46	92.85
	adaboostm1	95.97	97.26	96.65	96.68	**93.29**
	randomforest	92.44	94.95	93.78	93.82	87.51
	reptree	90.85	93.40	91.89	92.25	84.34
Explicit FFT with AR	j48	86.74	94.96	93.40	91.25	82.40
	logistic	96.85	95.66	94.83	96.19	92.35
	adaboostm1	95.49	97.06	96.39	96.35	92.63
	randomforest	92.30	97.18	96.42	94.98	89.89
	reptree	89.86	95.23	93.94	92.81	85.49
AFDB[1]	**replication**	95.94	97.23	96.61	96.65	93.23

[1] Records `00735` and `03665` excluded.
'NA' indicates not applicable because there the metric is not offered by the reference [1].

`adaboostm1`. And if we consider the additional layer of complexity and an increment that is smaller than the previous one of (+0.14) achieved by four algorithms, then we can say that is not successful at all.

In the last case, coefficients of the autoregressive model were used in addition to FFT over the explicit dataset. A general increment in the worst-performing algorithms was achieved, but the `logistic` decreased of (−0.21) in terms of MCC and the `adaboostm1` remained completely unchanged.

Thus because of the inconsistency in the statistical margins and the increase in complexity compared to the basic version of the datasets, this path can be considered as a failure.

Transient values

In 2.2.4 Shannon entropy has been defined on a bin with size 127. Therefore, to avoid possible interpretations of the state of the art work, the first 126 transient beats were removed, the $127nth$ value it's completely defined on the previous 126 values, thus it was not removed. The dataset used as a base were the explicit and encoded dataset, because if there are improvements here consequently there are on the experimentations that use them as a starting point. Table 3.5 shows the results obtained.

In the case of the explicit dataset, logistic in terms of SE obtained a remarkable increment of (+0.86). But was not enough to compete with the

Table 3.5: Machine learning algorithms applied on the dataset with explicit and encoded entropy without the transient beats, compared with the replication of Zhou and Zhou, et al [1, p. 7] itself.

Dataset	**Algorithm**	**Results**				
		SE(%)	**SP**(%)	**PPV**(%)	**ACC**(%)	**MCC**(%)
Explicit	j48	96.08	97.19	96.23	96.72	93.29
	ibk	94.99	96.14	94.83	95.65	91.11
	logistic	**96.97**	96.60	95.50	96.76	93.40
	bayesnet	96.12	96.73	95.64	96.47	92.79
	adaboostm1	95.65	97.33	96.39	96.61	93.07
	randomforest	95.23	96.16	94.87	95.77	91.35
	reptree	95.44	97.29	96.33	96.50	92.84
Encoded	j48	96.56	97.26	96.34	96.96	**93.79**
	ibk	96.56	97.26	96.34	96.96	**93.79**
	logistic	96.11	97.53	96.66	96.92	93.71
	bayesnet	96.13	97.45	96.57	96.89	93.64
	adaboostm1	96.11	97.53	96.66	96.92	93.71
	randomforest	96.56	97.26	96.34	96.96	**93.79**
	reptree	96.56	97.26	96.34	96.96	**93.79**
AFDB	**replication**	96.01	97.51	96.62	96.87	93.59
AFDB[1]	**replication**	96.11	97.53	96.66	96.92	93.71

[1] dataset without the 126 transient beats.
'NA' indicates not applicable because there the metric is not offered by the reference [1].

replica results based on $AFDB^1$. If compared with the results in Table 3.3, an overall slight improvement in terms of MCC was achieved and the distance between the replica and the best performance algorithm reduce from (-0.26) to (-0.19).

As for the encoded dataset, the behaviour was quite similar to what happened in 3.3. In fact, `j48`, `ibk`, `random forest` and `reptree` had the same performances. But `logistic` and `adaboostm1` added nothing more to the replica. In terms of MCC compared to the replica, the quartet of algorithms had an increment of around $(+0.08)$. All in all, the experiment without removing the 126 transient beats, obtained greater increments equal to $(+0.14)$ in terms of MCC.

3.3 Improving specific aspects in AF detection

After applying the above techniques, it is necessary to choose the experiment that had the best performance that is that on the encoded entropy dataset. The conclusions on global prediction can be drawn by observing Table 3.6 based on the mentioned dataset. The AFDB data set consists of 25 records, 12 of which were classified better than the state of the art algorithm. As regards the 3 records not reported in the table, there were neither improve-

ments nor worsening compared to the basic version. Generally, however, there is a slight improvement in the number of correctly classified fibrillating beats.

3.4 Further developments

Further developments of the work could be more investigations on the gap of beats between the reply and what is reported on the reference paper. It is also necessary to understand what the artefact is between the performance of the replica and that of the paper. Also, it is possible to use neural networks to obtain exhaustive experimentation of machine learning.

Table 3.6: Records with specific improvements of classification based on encoded entropy experiment.

Record	Algorithm	Confusion matrix				Surpasses
		TP	TN	FP	FN	
05261	j48 & others	**771**	43888	702	**163**	Se
	Zhou et, al	654	44217	380	280	-
07879*	j48 & others	**39995**	16447	102	**40**	Se, Acc, Mcc
	logistic, bayesnet, adaboostm1	**39945**	16487	62	**90**	Se
	Zhou et, al	39943	16496	60	92	-
06453*	All	**126**	**34288**	**94**	**319**	All metrics
	Zhou et, al	117	34272	117	328	-
04043*	j48 & others	**9170**	44060	3211	**5464**	Se, Acc, Mcc
	logistic, bayesnet, adaboostm1	**8693**	44211	3060	**5941**	Se, Acc, Mcc
	Zhou et, al	8544	44321	2957	6090	-
05091	j48 & others	**4**	36615	12	**137**	Se
	Zhou et, al	0	36634	0	141	-
00735	logistic, adaboostm1	275	39840	**51**	57	Sp, Ppv
	Zhou et, al	277	39846	52	55	-
03665	j48 & others	**10995**	41394	304	**62**	Se
	Zhou et, al	10985	41416	289	72	-
08405*	All	**45019**	13743	8	**76**	Se, Acc, Mcc
	Zhou et, al	45003	13758	0	92	-
06995*	j48 & others	**27028**	25553	2110	**488**	Se
	logistic, adaboostm1	**26975**	**25645**	**2018**	**541**	All
	Zhou et, al	26961	25640	2023	562	-
07910	j48 & others	**6504**	29589	230	**266**	Se
	Zhou et, al	6499	29722	104	271	-
04908*	j48 & others	**5526**	**55240**	**700**	**284**	All
	logistic	5446	**55279**	**661**	364	Sp, Ppv, Acc, Mcc
	Zhou et, al	5491	55055	892	319	-
08455*	j48 & others	**44189**	15238	40	**75**	Se, Acc, Mcc
	Zhou et, al	44103	15246	39	161	-
05121	j48 & others	**32592**	**14976**	**1135**	1168	Sp, Ppv
	logistic, adaboostm1	32589	**14986**	**1125**	1171	Sp, Ppv
	bayesnet	32589	**14981**	**1130**	1171	Sp, Ppv
	Zhou et, al	32689	14923	1195	1071	-
08378	bayesnet, adaboostm1	10996	33886	**142**	481	Sp, Ppv
	Zhou et, al	11008	33891	144	469	-
04015*	j48 & others	**499**	40714	2756	**26**	Se, Mcc
	logistic, adaboostm1	483	40807	**2663**	42	Sp, Ppv, Mcc
	bayesnet	**496**	40771	2699	**29**	Se, Ppv, Mcc
	Zhou et, al	485	40812	2665	40	-
06426	logistic, adaboostm1	52014	**799**	**1220**	1112	Sp, Ppv
	Zhou et, al	52095	796	1223	1038	-
04048*	j48 & others	**632**	38909	202	**181**	Se, Ppv, Acc, Mcc
	bayesnet	**556**	38930	181	**257**	Se, Acc, Mcc
	Zhou et, al	419	38982	136	394	-
04936*	All	**33559**	12086	1869	**6122**	Se, Acc, Mcc
	Zhou et, al	32662	12362	1600	7019	-
04126*	j48 & others	**3151**	38664	893	**142**	Se, Acc, Mcc
	logistic, bayesnet, adaboostm1	3017	38789	**768**	276	Sp
	Zhou et, al	3021	38795	769	272	-
07162	All	39198	0	0	**90**	Se, Acc
	Zhou et, al	39198	0	0	97	-
07859	All	61789	0	0	**93**	Se, Acc
	Zhou et, al	61789	0	0	100	-
08219*	j48 & others	**12953**	42511	2578	**1241**	Se, Ppv, Acc, Mcc
	logistic, bayesnet, adaboostm1	12643	**42605**	**2484**	1551	Sp, Ppv, Acc, Mcc
	Zhou et, al	12645	42555	2541	1549	-

* record classified better than the state of the art.
NOTE: `j48`, `ibk`, `random forest` and `reptree` had the same performances. First one is reported.

Chapter 4

Local prediction

4.1 Introduction

A further trial to improve the performances of the algorithm analysed in Chapter 2, is to carry out a local prediction approach during the training. That is, instead of finding the discriminating threshold that works for all patients, a specific threshold per patient is found. Afterwards, the records are grouped according to a common criteria. The reason behind this is the physiological difference among each patient, which can lead to the use of an unsuitable discriminatory threshold.

4.2 Finding threshold

The procedure to complete the task consists in brutally testing all the thresholds in the range $[0.0, 1.0]$ with an increase of 0.001. But unlike what was done in Section 2.2.4, the method acts on a single record of a specific patient. The choice of the right threshold among all the others, can be done by optimizing different parameters and metrics (equation 2.6 for reference).

4.2.1 Receiver operating characteristic

In the first experiment, the ROC was employed to optimize in the same way as Zhou, et al[1] did, but for every single record. To decide if the performances were remarkable, the MCC metric was calculated on the total of the values of the confusion matrix. Then it was compared with the replica realized for the global prevision. On the AFDB dataset the final MCC was 90.05% while the replica's one was 93.59% and the average of all the thresholds was 0.461. Instead on the AFDB dataset without the 126 transient values the

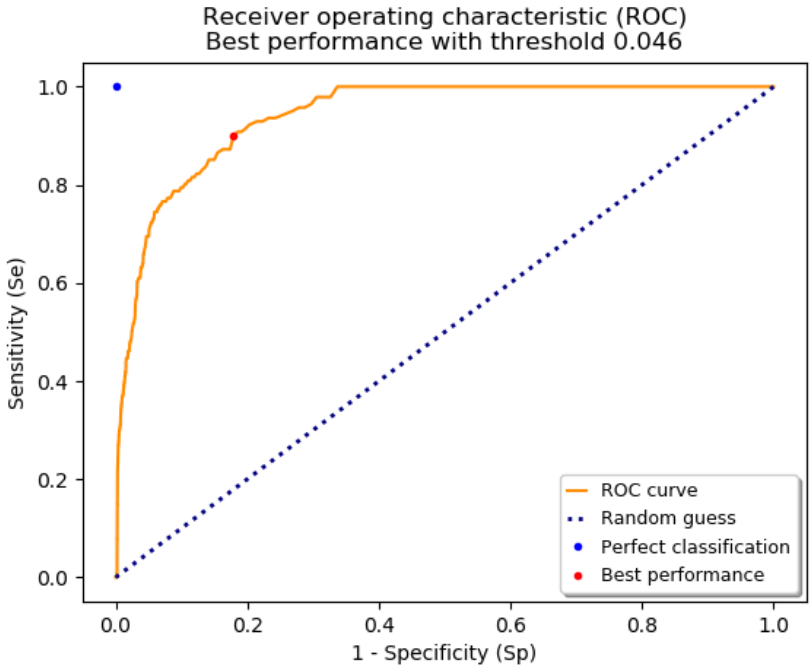

Figure 4.1: Example of optimisation through ROC. For the record 05091, the best performance is obtained with threshold 0.046. $Se = 0.90$ and $Sp = 0.82$.

MCC was 90.12% while the one of the replica was 93.71% and the average threshold 0.522. In both cases, the optimisation carried out was not sufficient to achieve good performance.

4.2.2 Maximizing the number of TP and TN

In this trial the optimization was directed on the number of beats correctly classified, i.e., true positives and true negatives. That is to find that threshold with the highest accuracy (ACC). This method on the AFDB dataset yield to a aggregate MCC of 95.00% compared to the 93.59% of the replica. In this case the threshold average lied around 0.578. On the other hand, the AFDB reduced of the 126 transient beats, had a MCC of 95.04% while the replica 93.71%. The average of the threshold was 0.637. This method was indeed the best one (Table 4.1).

4.3 Clustering

The optimal thresholds were obtained in 4.2.2, the next step to aggregate patients with a similar ECG from a physiological point of view is to create

Table 4.1: Summary of local prediction best threshold per record.

Method	Dataset	MCC	Threshold average
ROC	AFDB	90.05%	0.461
ROC	AFDB[1]	90.12%	0.461
TP & TN	AFDB	95.00%	0.578
TP & TN	AFDB[1]	95.04%	0.637
Replica	AFDB	93.59%	0.639
Replica	AFDB[1]	93.71%	0.639

[1] AFDB without the 126 transient beats.

different groups, i.e. clusters. The thresholds found through the optimization of max accuracy on AFDB without the 126 transient beats was taken as reference. From here on, there are two main ways forward.

4.3.1 Through numeric intuition

The first is to use numerical intuition. Depending on how optimization is defined, different results can be obtained. For example, taking the first value that meets a specific criterion leads to different results than taking the last one. To be thorough in the experimentation, both methods have been reported in 4.2 that shows the clustering. Albeit only the records where the threshold can move along a wide interval were the one that made big leap. Hence for those records, was not really important the picked up threshold, but they are still to be kept in mind because they can lead to interesting implications. The remaining records were completely independent of the method used and clung more or less to the same clusters. Especially the one in the interval $[0.4, 0.8]$ where most of the records tend to stay around those groups.

4.3.2 Through machine learning

The second way to find clusters is through machine learning techniques. A dataset to perform a leave one person out experiment was made. It was formed by groups of first 30 beats with atrial fibrillation and their optimal threshold of belonging record. The latter was the label at the base of the supervised learning. A `gls` function was used, which fits a linear model using generalized least squares. Unfortunately most of the predicted thresholds were really similar (Table 4.3), thus near the average of 0.54632782. A further experiment with a time window of 126 heartbeats, a length equal to the state of the art window, was performed but led to almost similar results. The

Table 4.2: Numerical clustering applied on the optimal thresholds found maximizing accuracy on AFDB without the 126 transient values. Method 1 takes the first thresholds that meets a criteria, Method 2 takes the last one.

Method 1			**Method 2**		
Cluster block	Record	Threshold	Cluster block	Record	Threshold
$[0.0, 0.1)$	7162	0	$[0.1, 0.2)$	8215	0.197
	7859	0		8455	0.199
$[0.1, 0.2)$	8455	0.194	$[0.3, 0.4)$	8378	0.354
	8215	0.195	$[0.4, 0.5)$	7910	0.47
$[0.3, 0.4)$	8378	0.353		4746	0.471
$[0.4, 0.5)$	4746	0.434		7879	0.489
	7879	0.456	$[0.5, 0.6)$	4936	0.514
	7910	0.47		4043	0.515
$[0.5, 0.6)$	4936	0.514		5121	0.557
	4043	0.515		6426	0.576
	5121	0.557	$[0.6, 0.5)$	8219	0.623
	6426	0.576		735	0.632
	735	0.583		8405	0.632
$[0.6, 0.7)$	5091	0.622		7859	0.687
	8219	0.623		4908	0.693
	8405	0.628	$[0.7, 0.8)$	8434	0.71
	4908	0.693		3665	0.713
$[0.7, 0.8)$	8434	0.71		6453	0.726
	3665	0.713		6995	0.75
	6453	0.721		5261	0.758
	6995	0.75	$[0.8, 0.9)$	7162	0.803
	5261	0.758		4126	0.899
$[0.8, 0.9)$	4126	0.899	$[0.9, 1.0)$	4048	0.979
$[0.9, 1.0)$	4048	0.969		4015	0.999
	4015	0.99	$[1.0, -)$	5091	1

problem was the over-representation of thresholds in the range $[0.55, 0.65]$ as it can be seen in (Figure 4.2).

4.4 Further developments

An overall good method was not found during the local prediction analysis. Future developments should aim to find a model with remarkable performance. For example, possible methods that might prove interesting are genetic algorithms that are more complex than linear regression. Subsequently, certain thresholds should be applied to certain groups of records, i.e. patients with a common physiology, and thus better performance should be achieved.

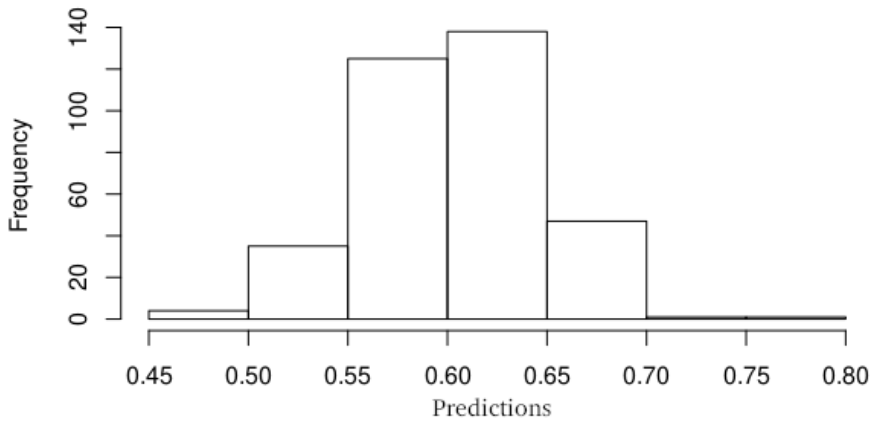

Figure 4.2: Histogram of prediction for the record 08455

Table 4.3: GLS function applied on a dataset with a window of the first 30 af beats. The output of the function is a threshold.

Record	**Optimal threshold**	**Predicted threshold**	**Difference**
00735	0.632	0.5490468	0.08295317
03665	0.713	0.5397742	0.1732258
04015	0.999	0.5435027	0.4554973
04043	0.515	0.5476209	0.03263931
04048	0.979	0.5394849	0.4395151
04126	0.899	0.5246542	0.3743458
04746	0.471	0.5518012	0.08080118
04908	0.693	0.5429615	0.1500385
04936	0.514	0.5499928	0.03600446
05091	1	0.5503535	0.4496465
05121	0.557	0.5460103	0.01135475
05261	0.758	0.5456952	0.2123048
06426	0.576	0.5437635	0.03224319
06453	0.726	0.5489861	0.1770139
06995	0.75	0.5348914	0.2151086
07162	0.803	0.4498696	0.3531304
07859	0.687	0.5280188	0.1589812
07879	0.489	0.5547526	0.06575262
07910	0.47	0.548922	0.07892204
08215	0.197	0.6375596	0.4405596
08219	0.623	0.5442554	0.07874457
08378	0.354	0.5520026	0.1980026
08405	0.632	0.5341969	0.09780312
08434	0.71	0.5460842	0.1639158
08455	0.199	0.6039946	0.4049946

Bibliography

[1] Xiaolin Zhou, Hongxia Ding, Wanqing Wu, and Yuanting Zhang. A real-time atrial fibrillation detection algorithm based on the instantaneous state of heart rate. *PLOS ONE*, 10(9):1–16, 09 2015.

[2] Atrial fibrillation fact sheet|data & statistics|dhdsp|cdc, Aug 2017.

[3] Philip A. Wolf, Robert D. Abbott, and William B. Kannel. Atrial Fibrillation: A Major Contributor to Stroke in the Elderly: The Framingham Study. *Archives of Internal Medicine*, 147(9):1561–1564, 09 1987.

[4] Thomas M. Munger, Li-Qun Wu, and Win K. Shen. Atrial fibrillation. *Journal of biomedical research*, 28(1):1–17, Jan 2014. 24474959[pmid].

[5] Sumeet S. Chugh, Rasmus Havmoeller, Kumar Narayanan, David Singh, Michiel Rienstra, Emelia J. Benjamin, Richard F. Gillum, Young-Hoon Kim, John H. McAnulty, Zhi-Jie Zheng, Mohammad H. Forouzanfar, Mohsen Naghavi, George A. Mensah, Majid Ezzati, and Christopher J.L. Murray. Worldwide epidemiology of atrial fibrillation. *Circulation*, 129(8):837–847, 2014.

[6] Types of atrial fibrillation: Persistent, paroxysmal & permanent afib, Jun 2018.

[7] H. S. Abed and G. A. Wittert. Obesity and atrial fibrillation. *Obesity Reviews*, 14(11):929–938, 2013.

[8] David Tonelo, Rui Providência, and Lino Gonçalves. Holiday heart syndrome revisited after 34 years. *Arquivos brasileiros de cardiologia*, 101(2):183–189, Aug 2013. 24030078[pmid].

[9] Xin Du, Jianzeng Dong, and Changsheng Ma. Is atrial fibrillation a preventable disease? *Journal of the American College of Cardiology*, 69(15):1968 – 1982, 2017.

[10] Caroline S. Fox, Helen Parise, Ralph B. D'Agostino, Sr, Donald M. Lloyd-Jones, Ramachandran S. Vasan, Thomas J. Wang, Daniel Levy, Philip A. Wolf, and Emelia J. Benjamin. Parental Atrial Fibrillation as a Risk Factor for Atrial Fibrillation in Offspring. *JAMA*, 291(23):2851–2855, 06 2004.

[11] Laila Staerk, Jason A. Sherer, Darae Ko, Emelia J. Benjamin, and Robert H. Helm. Atrial fibrillation: Epidemiology, pathophysiology, and clinical outcomes. *Circulation research*, 120(9):1501–1517, Apr 2017. 28450367[pmid].

[12] Ernest Noble Chamberlain, David Gray, and Andrew R. Houghton. *Chamberlains symptoms and signs in clinical medicine: an introduction to medical diagnosis*. Hodder Arnold, 2010.

[13] Richard L Page, Thomas W Tilsch, Stuart J Connolly, Daniel J Schnell, Stephen R Marcello, William E Wilkinson, and Edward LC Pritchett. Asymptomatic or "silent" atrial fibrillation: frequency in untreated patients and patients receiving azimilide. *Circulation*, 107(8):1141–1145, 2003.

[14] Atrial fibrillation, Jun 2019.

[15] L.S. Lilly and Harvard Medical School. *Pathophysiology of Heart Disease: A Collaborative Project of Medical Students and Faculty*, page 74. Wolters Kluwer, 2015.

[16] Valentin Fuster, Lars E. Rydén, David S. Cannom, Harry J. Crijns, Anne B. Curtis, Kenneth A. Ellenbogen, Jonathan L. Halperin, Jean-Yves Le Heuzey, G. Neal Kay, James E. Lowe, S. Bertil Olsson, Eric N. Prystowsky, Juan Luis Tamargo, Samuel Wann, null null, Sidney C. Smith, Alice K. Jacobs, Cynthia D. Adams, Jeffery L. Anderson, Elliott M. Antman, Jonathan L. Halperin, Sharon Ann Hunt, Rick Nishimura, Joseph P. Ornato, Richard L. Page, Barbara Riegel, null null, Silvia G. Priori, Jean-Jacques Blanc, Andrzej Budaj, A. John Camm, Veronica Dean, Jaap W. Deckers, Catherine Despres, Kenneth Dickstein, John Lekakis, Keith McGregor, Marco Metra, Joao Morais, Ady Osterspey, Juan Luis Tamargo, and José Luis Zamorano. Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation. *Circulation*, 114(7):e257–e354, 2006.

[17] Z.F. Issa, J.M. Miller, and D.P. Zipes. *Clinical Arrhythmology and Electrophysiology: A Companion to Braunwald's Heart Disease*, page 221. Companion to Braunwald's Heart Disease Series. Saunders, 2009.

[18] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):E215–220, Jun 2000.

[19] G. MOODY. A new method for detecting atrial fibrillation using r-r intervals. *Computers in Cardiology*, pages 227–230, 1983.

[20] Simona Petrutiu, Alan V. Sahakian, and Steven Swiryn. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. *EP Europace*, 9(7):466–470, 05 2007.

[21] Andrius Petrėnas, Vaidotas Marozas, and Leif Sörnmo. Low-complexity detection of atrial fibrillation in continuous long-term monitoring. *Computers in biology and medicine*, 65:184–191, 2015.

[22] J. Lee, Y. Nam, D. D. McManus, and K. H. Chon. Time-varying coherence function for atrial fibrillation detection. *IEEE Transactions on Biomedical Engineering*, 60(10):2783–2793, Oct 2013.

[23] Xiaolin Zhou, Hongxia Ding, Benjamin Ung, Emma Pickwell-MacPherson, and Yuanting Zhang. Automatic online detection of atrial fibrillation based on symbolic dynamics and shannon entropy. *BioMedical Engineering OnLine*, 13(1):18, 2014.

[24] George B Moody and Roger G Mark. Mit-bih arrhythmia database, 1992.

[25] The Arrhythmia Laboratory The Beth Israel Deaconess Medical Center. The mit-bih normal sinus rhythm database, 1990.