Atrial Fibrillation Detection from Short-Term Recording of a Single Abduction

Abstract

Atrial Fibrillation has always been on the spotlight in the academic research part firstly in medical sciences and therefore, based on technological advancements, in computer sciences as well. It is type of heart beat rate inconsistency in which the human body's heart rate does not function as most of the human's heart beats do. It affects a very small part of the mankind globally, around 1.5%, but its consequences are very serious. This work focuses on the detection and therefore prediction of atrial fibrillation heart beats. At first, atrial fibriallation is explained, giving an example, the dataset is described which consists of four different types of heart beat rates and thereafter the results and outcomes are presented.

Keywords: `Atrial Fibrillation`, Short-Term Recording, Single Abduction, Multi-class classification supervised machine learning model, Model Accuracy

1. Introduction

ECG examinations firstly appeared back the 19th century. Willem Eingthofen was the first one who proposed and received the Nobel Prize in 1924. ECG is a way to examine and diagnose heart beat rates and does not actually by no means interfere into the human body. It helps diagnose the normal heart beat rate like stroke, dementia, anomalies and fibrillations [1, 2]. An ECG is based on measurements emphasizing the P, Q, R, S, T curve spots. The results extracted from an ECG define the conditions and the situation in which the

---

⋆Fully documented templates are available in the elsarticle package on CTAN.

patient is currently in. For example, results are currently based on the season.

People who are in charge, especially mostly and hopefully only doctors in this case acknowledge the fact that the ECG received presents an ECG illustration susceptible to mistakes, computer-wised such as noise [3, 4, 5] due to computer, accuracy error.

Figure 3 illustrates a heart input, the so-called Atreus in which contractions happen and the blood flows from the inner part into the ventricular system. R-wave is the ventricular system in which there are two large inputs of repolarization that creates the central circulatory system. T-wave is step right before of the heart repolarization that happens in order to restart the heart rate and the next heart beat comes forward clock-wise. The time being between the P-wave and the QRS complex is the time that before the advent of the heart contraction atrium up until the heart ventricle. The particular time is named after P-Q period of time. Well, the normal P-Q period of time is all but 0.16 seconds. That period of time is sometimes called P-R because there are times when Q is absent. The ventricle contraction practically leads from the start of Q wave up until the end of T wave. That period of time is named after Q-T of time and the normal period of time is 0.35 seconds. The frequency of heart beat rate is defined easily from an ECG instance, because that period of time that interference between two consecutive heart beats is the inverse of the heart frequency. If the period of time that interferes between two consecutive heart beats, as defined with enumeration lines, is equal to 1 second, the heart frequency is equal to 60 heart beats per minute. The normal period of time that interferes between QRS complex is equal to 0.83 seconds. Therefore, this heart frequency in this case is equal to 72 heart beats per minute. Heart beat fibrillation are classified based on originality, beat, period of time and/or creation mechanism. The originality category is classified into ventricular fibrillation, ventricle fibrillation and fibrillation that origins from AV circularity. The heart beat is classified in bradycardia (rate $\leq$ 60 beats/min) and tachycardia (rate $\geq$ 100 beats/min). The period of time is clssified into the chronic heart beat ($\geq$ 30 seconds) and the non-chronic one ($\leq$ 30 seconds). The creation and the
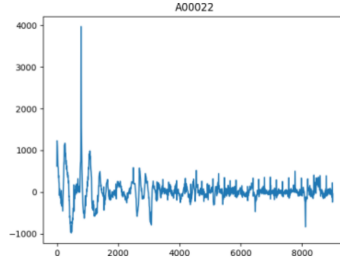
2

Figure 1: ECG presenting Noise.

beginning of a heart rate is classified into the polarization starts automatically without cells being bother one another, metapolarization in which primary polarization leads into emergent contractions due to elecrical instability into the membrane and reentering in which the current follows a circular movement and repolarizes circularly the cells that have been polarized 1.

2. Atrial Fibrillation

Atrial Fibrillation is the most ordinary heart beat inconsistent form. It appears to 1-2 % of the total worldwide. Nowadays, almost 12 million Europeans and Americans are diagnosed with urgent Atrial Fibrillation. Atrial Fibrillation increases as long as human being ages. It hardly is diagnosed because it is sudden. Figure 2 illustates two graphs in which the first one presents the pattern of the heart beat progress along with normal heart beat and the second one on the bottom of the graph presents an ECG instance diagnosed with Atrial Fibrillation. X-axis shows time in seconds and y-axis shows Amplitude in mV.
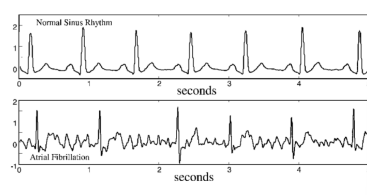

Figure 2: Normal ECG Example with wave representations

The goal of this work is dataset classification based on heart beat rate using classification supervised machine learning methods which is one of the machine learning model categories. There are four different heart beat categories, the first one is the normal one (N), the second one is the one of Atrial Fibrillation (A), a random noise (O) and noise ( ). The given, corresponding data set consists of 8.528 instances and they are patient recordings. They are publicly available for a contest that took place in 2017 for the dataset training part using machine learning methods for the best accuracy model extraction to predict the heart beat type [6].

3. Research Methodology

In order to extract a classification machine learning model, our data first and foremost should be analyzed and pre-processed. The data set consists of 8.528 instances. Each one of the instance belongs to one of the four classes, in total. The annotation of each class is as follows. The letter N corresponds to the class in which the heart rate is normal, the letter A corresponds to the class in which the heart rate belongs to the one classified as atrial fibrillation, the letter O corresponds to the class in which the heart rate belongs to an arbitrary heart rate and the ~corresponds to the class in which there is noise in the ECG sample-instance examined. One of the basic steps of pre-processing is to find the total number of observations that corresponds to each of the four classes. Firstly, classes have been decoded in a computer-wise manner. The class with the letter A takes the value of 0, the class N takes the value 1, the class O takes the value of 2 and the class that corresponds to the letter ~takes the value 3.

The given data set is imbalanced, according to Table 1. That means that there is skewness in the data classification. Specifically, the proportion of the observations, instances of the class that corresponds to the letter N and the class that corresponds to the class with the symbol ~is 1:18.

In order to face the imbalance data into our data set, the particular framework oversamples the sample, meaning that it dupliactes the instances from the
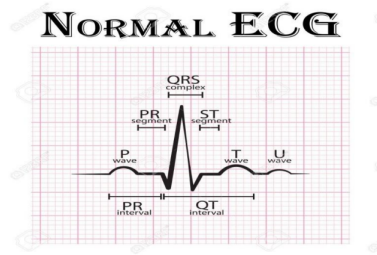
4

Figure 3: Normal ECG Example with wave representations

Figure 4: Normal Heart Rate
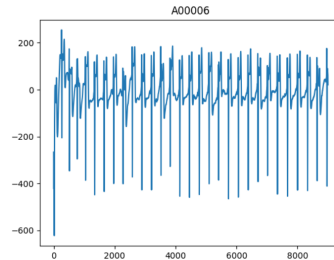
Figure 5: Normal Heart Rate

Figure 6: Atrial Fibrillation

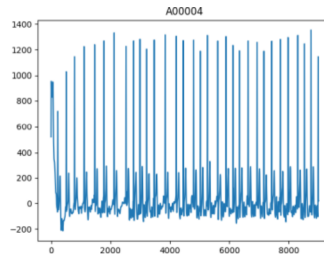class which has the least number of observations into the training data set. The goal is to avoid any error caused because out model is about to overfit. In order to solve such issues, the reduction or even the removal of some instances from the classes that might have more observations than the ones which have less, so that the number of observations is close or the same with the one which has
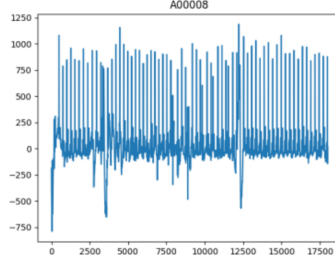
Figure 7: Arbitrary Heart Rate
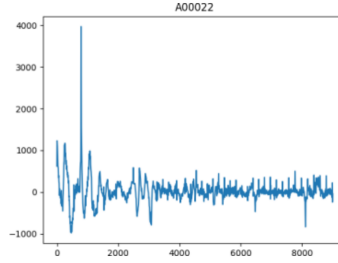
Figure 8: Noise

Table 1: Imbalances dataset enumeration

Class	No. of Observations
'A'	738
'N'	5050
'O'	2456
''	284

more instances. Therefore, the data set consists of 1.136 instances in which each class has equal number of observations corresponding to 285 observations.

⁹⁰ The second number of observations of preprocessing is the removal of continuous current that is created into the ECG instance examination. So, in order to make sure that there is such a noise from the power supply, DC bias is calculated which gives the mean of each wavelength. If the mean of each wavelength equals to zero, then there is no DC bias. In this particular case, there is DC bias

and there will always be on most of the cases. In order to deal with such a pre-processing data cleaning, the mean of each ECG is calculated and subsequently, the values are normalized in the range of between -1 and 1, [-1,1].

The third step is the feature extraction through the ECG cardiographs. Using the appropriate libraries in Python programming language such as scipy and biosppy and in accordance with the help of basic mathematical calculations such as mean, median, minimum and maximum values. The calculation of the feature extraction outputs the following features from the medical point of view. PR interval shows the time between the starting point of P-wave up until the starting point of QRS curve, the maximum value of P-wave, time between P and R waves, time between P and W, the part between ST curves, the maximum value of T, a value corresponding before the maximum value of R wave, a value corresponding after the maximum value of R wave, the time being between T and R waves, time between T and S waves, time between P and T waves, skewness in P wave, curve in T wave, time being in QRS curves, time being in QS curves, kurtosis in QRS, skewness in QRS and the difference between the maximum and the minimum value of QRS wave that appears in an ECG instance.

Data pre-processing task is completed and therefore, the implementation and evaluation of machine learning model in 4.

4. Research Analysis and Preprocessing metrics

The particular machine learning used belongs to a classification supervised machine learning categories. A lof ot applications in machine learning are based on multi-class classification models. The challenge and work examined in this paper belongs to multi-class classification machine learning models because the number of classes supervised is greater than 1.

So, in a binary classification problem, there are two classes often called P and N, named after Positive and Negative. In such a problem, the confusion matrix extraction is a good tactic, in which there are two rows and columns respectively.

7

In multi-class classification problem such that, the terms of Precision and Recall are presented. Precision term is referred to True Positive quantity divided by the addition of True and False Positive. From the mathematical point of view, Precision metric is described as follows 1:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Precision metric shows the proportion of the predicted values which are truly positive and not predicted as false though they were true. A second metric examine is the Recall metric. This particular, mathematically speaking, is expressed in 2.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

The Recall metric is referred to True Positive divided by the addition of the ones which are True Positive and False Negative. Recall metric covers the question of the proportion of the True Positive are correctly classified as given in 3.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

So, it is the summation of the diagonal values of the confusion matrix divided by the total values of the matrix.

An additional metric that needs to be mentioned is the F1-score. This particular metric outputs a value between Precision and Recall metrics. It is not considered a reliable one, though many researcher still use it to compare machine learning models [7]. F1-score is expressed in equation 4.

$$F1_score = 2 \times \frac{precision * recall}{precision + recall} \tag{4}$$

The reason that the particular metric won't be taken into consideration is because it emphasizes mostly on lower values that Precision and Recall metrics output. For instance, if Precision is 100% and Recall 0%, then F1score equals to 0% and not 50%.

The machine learning model studies, examined and implemented belongs to supervised machine learning models and is the Random Forest one. Shortly, decision trees from the given data set are created so that the decision is taken into consideration from each one of them. The Random Forest algorithm samples a particular number of instances from the current data set. The algorithm constructs a decision tree for each sample and receives an output from each of these samples. Each sample output is configured whether it is an appropriate one or not for classification prediction.

Experiments have been conducted successfully adjusting the appropriate parameters of the machine learning models so that the best optimal accuracy is succeeded and as a result, the selection of the best machine learning model. The best optimal result is chosen with the one that consists of the most final predicted/supervised summation output.

Machine learning training processed had been performed firstly tuning the hyperparameters of the machine learning model so that the best accuracy score succeeded and consequently the best model is chosen. There are three metrics examined for the evaluation and the comparison amongst these models. The metrics are the Mean Absolute Error (MAE), Mean Squared Error(MSE) and Root Mean Squared Error (RMSE). Mean Absolute Error (MAE) is mathematically described in equation 5.

$$MAE = \sum_{i=1}^{n} \frac{||y_i - \lambda(x_i)||}{n} \tag{5}$$

Mean Squared Error is described in equation 6.

$$MSE = \frac{1}{n} \times \sum_{i=1}^{n} (y_i - \lambda(x_i)) \tag{6}$$

Root Mean Square Error is described in equation 7.

$$\sqrt{\frac{\sum_{i=1}^{n} (y_i - \lambda \times (x_i))}{n}} \tag{7}$$

in which n shows the total number of the sample of y_i which shows the real ones and λ

9

5. Research outcomes

The parameters in which the evaluation of the model is considered and subsequently the accuracy score of the particular work is estimated to be the number of the decision trees of the Random Forest algorithm (n_estimators) and the minimum number of the sample at each leaf of the random forest. It is considered that the weight of the leaves is equal to all of them.

So, starting and setting n_estimators = 10 and the number of leaves in the decision trees is from 1 to 10, the best optimal model is about to the one in which the number of leaves is equal to 3. Table 2 displays the results with the best optimal number of leaves in the decision tree to be equal to 3 which the errors are the following ones MAE = 0.51, MSE = 0.97, RMSE = 0.99 and Accuracy Score 0.69.

Table 2: Metric results for n_estimators = 10

Leaf	MAE 5	MSE 6	RMSE 7	Accuracy Score 3
1	0.62	1.23	1.11	0.63
2	0.54	1.06	1.03	0.68
3	0.53	1.04	1.02	0.69
4	0.59	1.15	1.07	0.64
5	0.56	1.09	1.04	0.67
6	0.61	1.22	1.10	0.64
7	0.60	1.18	1.09	0.64
8	0.61	1.20	1.10	0.65
9	0.58	1.15	1.07	0.66
10	0.63	1.23	1.11	0.63

For n_estimators = 10 and the number of leaves in the decision tree starting from 1 to 10, the best optimal number of leaves for the decision tree equals to 3. Table 3 shows the results for the best optimal number of leaves equals to 3 in which the errors are the following ones, MAE $\equiv$ 0.51, MSE $\equiv$ 0.97 and RMSE $\equiv$ 0.99.

Table 3: Metric results for n_estimators = 100

Leaf	MAE 5	MSE 6	RMSE 7	Accuracy Score 3
1	0.54	1.09	1.05	0.69
2	0.52	0.99	1.00	0.69
3	0.51	0.97	0.99	0.69
4	0.53	1.02	1.01	0.68
5	0.53	1.03	1.01	0.68
6	0.56	1.08	1.04	0.67
7	0.55	1.05	1.02	0.67
8	0.55	1.07	1.03	0.67
9	0.56	1.11	1.06	0.68
10	0.55	1.05	1.02	0.67

For n_estimators = 1.000 and for the number of leaves in the decision tree starting from 1 to 100, the best optimal number of leaves in the decision tree equals to 5. Table ?? displays the results with the best optimal number of leaves in the decision tree to be equal to 5 which the errors are the following ones MAE ≡ 0.55, MSE ≡ 1.07 and RMSE ≡ 1.04.

Table 4: Metric results for n_estimators = 1000

Leaf	MAE 5	MSE 6	RMSE 7	Accuracy Score 3
1	0.55	1.08	1.04	0.68
2	0.55	1.09	1.04	0.67
3	0.57	1.11	1.05	0.67
4	0.55	1.11	1.05	0.68
5	0.55	1.07	1.04	0.68
6	0.56	1.11	1.05	0.67
7	0.57	1.12	1.06	0.66
8	0.56	1.10	1.05	0.67
9	0.57	1.11	1.05	0.66
10	0.56	1.09	1.04	0.66

For n_estimators = 10.000 and for the number of leaves in decision tree starting from 1 to 10, the best optimal number of leaves for the decision tree equals to 9. In table 5 shows the results with the best optimal number of leaves to be equal to 9 with the following errors, MAE ≡ 0.55, MSE ≡ 1.06 and RMSE ≡ 1.03.

Table 5: Metric results for n_estimators = 10000

Leaf	MAE 5	MSE 6	RMSE 7	Accuracy Score 3
1	0.55	1.07	1.03	0.67
2	0.56	1.08	1.04	0.67
3	0.57	1.11	1.05	0.67
4	0.57	1.11	1.06	0.66
5	0.56	1.09	1.04	0.67
6	0.55	1.08	1.04	0.67
7	0.55	1.08	1.04	0.67
8	0.56	1.10	1.05	0.67
9	0.55	1.06	1.03	0.67
10	0.55	1.07	1.03	0.67

According to Tables 2, 3, 4, 5 which correspond to the best optimal results based pn the calculated errors. Table 6 shows that the best machine learning model is equal to the number of trees to be equal to 100 and the number of leaves of the decision tree to be equal to 3.

Table 6: Best results of MAE, MSE and RMSE

Leaf	n_estimators	MAE 5	MSE 6	RMSE 7
3	10	0.53	1.04	1.02
3	100	0.51	0.97	1.00
5	1000	0.55	1.07	1.04
9	10.000	0.55	1.06	1.03

The next step is the confusion matrix extraction to be estimated accurately:

• signal that have already been categorized as Atrial Fibrillation compared with the ones categorized as normal and

• signals that have already been categorized as Atrial Fibrillation compared with the ones categorized that they do not belong to the Atrial Fibrillation category.

Table 7 shows the confusion matrix results.

Table 7: Confusion Matrix

	A	N	O	˜
A	63	0	11	10
N	2	52	13	7
O	25	21	39	8
˜	1	3	5	81

Rows horizontally represent the predicted values, whilst the columns vertically represent the real ones.

Assuming it was a binary classification, meaning that we only predicted whether the ECG derives from Atrial Fibrillation data set or not, the numbers are as follows, TP= 52, FP = 0, TN = 63 and FN = 2, Precision = 1, Recall = 0.96 and Accuracy = 0.98.

As a multi-class classification problem as it is, Precision = 0.75, Recall = 0.89 and Accuracy = 0.68.

6. Conclusions

Atrial Fibrillation is defined to be as an atrial fibrillation in which the heart rate speed is not synchronized normally based on consecutive normal beats. That happens to 1-2% of the total population and is relevant to mortality and morbidity. The goal of this work is to score based on accuracy metrics which they separately characterized as Atrial Fibrillation, normal ones and the ones that they don't belong to any of the class. Results have shown that the first binary classification category reaches an accuracy score of 98% , whilst the

second multi class one reaches up to 68%. The difference between these two cases is equal to 30%. That difference is expected because, in the first case the machine learning model is a binary supervised one that belongs to Random Forest category and in the second case, the machine learning model is a multi class one that consist of four classes, the first one is the normal one, the second one is the atrial fibrillation one, the third one does not belong to any of the aforementioned classes and the last one is described as noise.

References

[1] V. Punia, R. Vakani, R. Burgess, S. Hantus, Electrographic and Clinical Natural History of Lateralized Periodic Discharges, Journal of Clinical Neurophysiology 35 (2017) 1. `doi:10.1097/WNP.0000000000000428`.

[2] L. Staerk, J. A. Sherer, D. Ko, E. J. Benjamin, R. H. Helm, Atrial Fibrillation, Circulation Research 120 (9) (2017) 1501–1517. `doi:10.1161/CIRCRESAHA.117.309732`.

[3] G. Osman, R. Rahangdale, S. Hantus, A. Herlopian, S. Hocker, J. Britton, J. W. Lee, B. Legros, M. Mendoza, V. Punia, A. Wallace, M. B. Westover, E. Gilmore, L. Hirsch, N. Gaspard, Electrographic Predictors of Electrographic Seizures and Electrographic Status Epilepticus in Patients with Bilateral Independent Periodic Discharges: A Retrospective Multi-Center Study (Dec. 2018).

[4] G. Lippi, F. Sanchis-Gomar, G. Cervellin, Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge, International Journal of Stroke: Official Journal of the International Stroke Society 16 (2) (2021) 217–221. `doi:10.1177/1747493019897870`.

[5] W. Chua, L. Di Biase, A. Luna, C. David, D. Haase, G. Hindricks, K. Hauesler, R. Al-Hussein, J. Piccini, L. Mont, J. Nielsen, L. Escobar, L. Bono, L. Fabritz, P. Kirchhof, Dynamic changes of cardiovascular biomarkers after ablation for atrial fibrillation: Observations from

AXAFA-AFNET5, European Heart Journal 42 (Oct. 2021). `doi:10.1093/eurheartj/ehab724.0538`.

[6] D. T. Linker, Accurate, Automated Detection of Atrial Fibrillation in Ambulatory Recordings, Cardiovascular Engineering and Technology 7 (2) (2016) 182–189. `doi:10.1007/s13239-016-0256-z`.

[7] P. Vittorini, S. Menini, S. Tonelli, An ai-based system for formative and summative assessment in data science courses, International Journal of Artificial Intelligence in Education 31 (12 2020). `doi:10.1007/s40593-020-00230-2`.

255