Multi-arrhythmias detection with an XML rule-based system from 12-Lead Electrocardiogram

Abdeldjalil Khelassi[1], Sarra-Nassira Yelles-chaouche[2], Faiza Benais[2]

[1] Informatics Department, Sciences Faculty, Abou Beker Belkaid University of Tlemcen, Tlemcen, Algeria
[2] Informatics Resources Center, University of Tlemcen, Tlemcen, Algeria

Type of article: Original

Abstract

Background: The computer-aided detection of cardiac arrhythmias stills a crucial application in medical technologies. The rule based systems RBS ensure a high level of transparency and interpretability of the obtained results.

Aim: To facilitate the diagnosis of the cardiologists and to reduce the uncertainty made in this diagnosis.

Methods: In this research article, we have realized a classification and automatic recognition of cardiac arrhythmias, by using XML rules that represent the cardiologist knowledge. Thirteen experiments with different knowledge bases were realized for improving the performance of the used method in the detection of 13 cardiac arrhythmias. In the first 12 experiments, we have designed a specialized knowledge base for each cardiac arrhythmia, which contains just one arrhythmia detection rule. In the last experiment, we applied the knowledge base which contains rules of 12 arrhythmias. We used, for the experiments, an international data set with 279 features and 452 records characterizing 12 leads of ECG signal and social information of patients. The data sets were constructed and published at Bilkent University of Ankara, Turkey. In addition, the second version of the self-developed software "XMLRULE" was used; the software can infer more than one class and facilitate the interpretability of the obtained results.

Results: The 12 first experiments give 82.80% of correct detection as the mean of all experiments, the results were between 19% and 100% with a low rate in just one experiment. The last experiment in which all arrhythmias are considered, the results of correct detection was 38.33% with 90.55% of sensibility and 46.24% of specificity. It was clearly show that in these results the good choice of the classification model is very beneficial in terms of performance. The obtained results were better than the published results with other computational methods for the mono class detection, but it was less in multi-class detection.

Conclusion: The RBS is the most transparent method for cardiac arrhythmias detection and multi arrhythmias detection. It improves an exceptional recognition of arrhythmias, but due to conflicts between rules, multi-arrhythmias and uncertainty of measures, the rate of correct classification was less than the other methods.

Keywords: Rule based system RBS, Extensible Markup Language XML, Cardiac Arrhythmias, Electrocardiogram ECG

1. Introduction

In recent years, the automatic detection of the cardiac arrhythmias established a domain among which several works and studies were developed for the discrimination between normal rhythm and its irregularity (1-3). The Electrocardiogram ECG represents the mechanical activity of the heart (the contraction and the relaxation, Slackening). Also, it represents the main technical source of information used for cardiac diagnosis. The detection of multi-arrhythmias from physiological signal stills is a complex task and needs more achievements and progress. This is due to several reasons such as conflicts, uncertainty and the weaknesses of the developed approaches. The aim of this study is to detect automatically the cardiac arrhythmia from an important stream of data (social information and 12 ECG leads). In our study, we tried to finalize a classification with rules written in Extensible

Markup Language XML by using the second version of the classifier XMLRULE. The classification was between twelve classes of cardiac arrhythmias adding to the normal case. This classification was for all classes, at the same time aiming to return the diagnosis more interpretable by an explicit system; our work allows bringing effective solutions to these types of problems. This classification uses 279 attributes extracted from twelve leads of the physiological signal ECG. The used data are extracted from a database published in UCI Machine Learning Repository (4) which contains patient's record with multi-arrhythmias problem (1). The database was published by the International Turkish Datum and the researchers H. Altay Guvenir, Burak Acar, Haldun Muderrisoglu, from the department of the engineering of computing and the sciences of the information of the university, Ankara, Turkey (1). The general objective of this research was to present a transparent method for cardiac arrhythmia detection. The specific objectives were 1) to apply XML rules which offer interpretation of obtained results after detection, 2) to ameliorate the performance of detection, and 3) to compare the rule based system with the other applied methods

2. Material and Methods
2.1. Study design
In this study, we have developed an interpretable knowledge base presented on XML rules inferred by self-developed software "XMLRULE". The process of our study was divided into several tasks:
1) The first one consists of collecting the clinical data.
2) The second one was to extract the maximum of the clinical information used by physicians to detect the arrhythmias.
3) The third one was to define a rational model for representation of the extracted knowledge in a standardized language XML.
4) The last one was to test the rational model by using the collected clinical data.

A detailed explanation of the electrocardiogram was presented in (2, 3, 5). Also in Figure 1, we present an example of normal ECG from 6 leads.
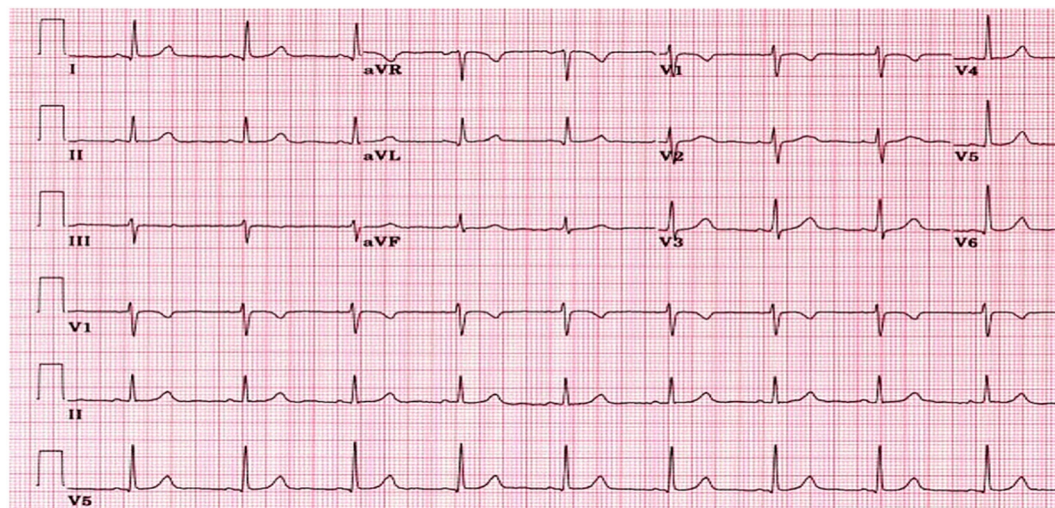

Figure 1. An example of normal ECG from 6 leads.

2.2. Clinical data collection:
The collected clinical data was constructed from a well-known dataset and published at UCI Machine Learning Repository (4), which belongs to Turk researchers H. Altay Guvenir, Burak Acar, Haldun Muderrisoglu, from the department of computer engineering and information sciences of the university in Ankara (1). The database contains 279 attributes including social information and ECG measures from all 12 leads. The used population contains 452 patients the date of data collection was not declared. Each record was diagnoses by cardiologists and the decision has been recorded in the database. Table1 shows the defined arrhythmias recorded in the data base. Several artificial intelligence approaches are used for the automatic detection and published in (1), these results were reused for comparing our developed approach.

2.3. Rational model

Several published books and articles review the cardiology and cardiac arrhythmias (5-10). The ECG is still the most important source of clinical information in cardiology although there are several alternatives and additional sources such as genetics (5, 6, 8) and medical imaging (7, 8). The Electrocardiogram ECG (Figure 1) represents the electrophysiological signal generated by the heart. All physicians apply the same rules for arrhythmias detection. Table 2 describes the used rules by physicians collected from literature (9, 10).

The used classifier, XMLRULE (Figure 2), is java self-developed software with five aims:

1) To facilitate the interaction between users and the knowledge base.
2) To ensure inferring and verifying the results.
3) It also allows personalization of the rules.
4) Also, to make the connection by visualizing the obtained results and the used given base and this by editing rules in XML.
5) It also allows classifying a base of datum which contains a large number of requests.

The applied engine of rules is "JRuleEngine1.3 which uses the "Jsr94-1.1". JRuleEngine is an engine of rules, which can also be considered as an "if/then" resolver based on production rules. The states "if/then" are interpreted and called rules. The parts" if' of rules contain the conditions. The parts "then" of rules contain actions, (Figure 3). Jsr94-1.1: "Java Specification Request " developed by Java Community Process of program (JCP), defines an API of execution Java for the engines of rules by supplying a simple API to reach an engine of rules from a platform Java Standard Edition J2SE or a platform Java Enterprise Edition J2EE.

Figure 2. Second version of XMLRULE: the rule-based classifier

```
<rule-execution-set>
    <name>RuleExecutionSet1</name>
    <description>Rule Execution Set</description>
    <synonymn name="ApplyRule" class="xmlrule.ApplyRule"/>
  - <rule name="Rule1" description="Normal">
       <if leftTerm="ApplyRule.getdureep" op="<" rightTerm="110.0"/>
       <if leftTerm="ApplyRule.getpr" op=">" rightTerm="120.0"/>
       <if leftTerm="ApplyRule.getpr" op="<" rightTerm="200.0"/>
       <if leftTerm="ApplyRule.gethrate" op=">" rightTerm="60.0"/>
       <if leftTerm="ApplyRule.gethrate" op="<" rightTerm="100.0"/>
       <if leftTerm="ApplyRule.getqt" op=">" rightTerm="350.0"/>
       <if leftTerm="ApplyRule.getqt" op="<" rightTerm="390.0"/>
       <then method="ApplyRule.printresultN"/>
    </rule>
  - <rule name="Rule1.1" description="maladie coronarienne lesion sous epicardique">
       <if leftTerm="ApplyRule.getSwave" op="<" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTwave" op="<" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTangle" op=">" rightTerm="180"/>
       <then method="ApplyRule.printresultCOR"/>
    </rule>
  - <rule name="Rule1.2" description="maladie coronariennelesion sous endocardique">
       <if leftTerm="ApplyRule.getSwave" op=">" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTwave" op=">" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTangle" op="<" rightTerm="180"/>
       <then method="ApplyRule.printresultCOR"/>
    </rule>
  - <rule name="Rule1.3" description="maladie coronarienne ischemie sous endocardique">
       <if leftTerm="ApplyRule.getTwave" op=">" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTangle" op="=" rightTerm="90"/>
       <then method="ApplyRule.printresultCOR"/>
    </rufe>
  - <rule name="Rule1.4" description="maladie coronarienne ischemie sous epicardique">
       <if leftTerm="ApplyRule.getTwave" op="<" rightTerm="0"/>
       <if leftTerm="ApplyRule.getTangle" op="=" rightTerm="270"/>
       <then method="ApplyRule.printresultCOR"/>
    </rule>
  - <rule name="Rule4" description=" infarctus du myocarde inferieur">
       <if leftTerm="ApplyRule.getSD2" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getSD3" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getSavf" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getQD2" op=">" rightTerm="0.33"/>
       <if leftTerm="ApplyRule.getQD3" op=">" rightTerm="0.33"/>
       <if leftTerm="ApplyRule.getQavf" op=">" rightTerm="0.33"/>
       <then method="ApplyRule.printresultIMI"/>
    </rule>
  - <rule name="Rule3" description=" infarctus du myocarde anterieur">
       <if leftTerm="ApplyRule.getSv1" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getSv2" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getSv3" op=">" rightTer="40"/>
       <if leftTerm="ApplyRule.getSv4" op=">" rightTerm="40"/>
       <if leftTerm="ApplyRule.getQv1" op=">" rightTerm="0.33"/>
       <if leftTerm="ApplyRule.getQv2" op=">" rightTerm="0.33"/>
       <if leftTerm="ApplyRule.getQv3" op=">" rightTerm="0.33"/>
       <if leftTerm="ApplyRule.getQv4" op=">" rightTerm="0.33"/>
       <then method="ApplyRule.printresultIMA"/>
    </rule>
  - <rule name="Rule5" description="Sinus tachycardy">
       <if leftTerm="ApplyRule.getdureep" op="<" rightTerm="110"/>
       <if leftTerm="ApplyRule.getpr" op=">" rightTerm="120"/>
       <if leftTerm="ApplyRule.getpr" op="<" rightTerm="200"/>
       <if leftTerm="ApplyRule.getdureeqrs" op="<" rightTerm="120"/>
       <if leftTerm="ApplyRule.gethrate" op=">" rightTerm="100"/>
       <then method="ApplyRule.printresultTAC"/>
    </rule>
```

Figure 3. Sample of XML rules extracted from the knowledge base

Table 1. The recorded classes in the database.

The classes	Arrhythmias
1	Normal
2	Ischemic changes (Coronary Artery Disease)
3	Old Anterior Myocardial Infarction
4	Old Inferior Myocardial Infarction
5	Sinus tachycardia
6	Sinus bradycardia
7	Ventricular Premature Contraction (PVC)
8	Supraventricular Premature Contraction (ESA)
9	Left bundle branch block
10	Right bundle branch block
14	Left ventricular hypertrophy
15	Atrial Fibrillation or Flutter
16	Others

Table 2. The used rules for arrhythmias diagnosis

Arrhythmias	Rules
Normal	P<110 ms, 60<hrate<100mn, 120<PR<200ms, 350<QT<390ms
Ischemic changes (Coronary Artery Disease)	//
*COR (epicardial Lesion)	S Wave<0, T Wave<0, T angle>180°
*COR (Endocardial lesions)	S Wave>0, T Wave>0, T angle<180°
*COR (Ischemia epicardia)	T Wave<0, T angle=270°
*COR (Endocardial Ischemia)	T Wave>0, T angle=90°
Old Anterior Myocardial Infarction	Q>1.3 ms, S large>40 ms en V1, V2, V3, V4
Old Inferior Myocardial Infarction	Q>1.3 ms, S large>40 ms en D2, D3, AVF
Sinus tachycardia	P<110 ms, 120<PR<200ms, QRS<120 ms, hrate>100mn
Sinus bradycardia	120<PR<200ms, QRS<120 ms, hrate<40mn
Ventricular Premature Contraction (PVC)	P=0, PR>140ms, QRS<120 ms
Supra ventricular Premature Contraction (ESA)	PR>120 ms, 60<QRS<100 ms
Left bundle branch block	T negative<0, QRS>120 ms
Right bundle branch block	QRS >0.12ms, onde T negative <0, SV6 large>50 ms
Left ventricular hypertrophy	Index of Socolow= Onde S at V1+ Onde R at V5 or V6>= 35mm, T wave negative, QRS>120ms
Atrial Fibrillation or Flutter	R=0, 400<hrate<600mn, 100<QRS<150ms
Others	Elsewhere

3. Results

After thirteen experiments with different knowledge bases, and with the same test data sets presented in section 2.1, the obtained results are resumed in Table 3. We observed in the first experiment, when we handled a normal case, an acceptable rate of 56.66% from correct and a rate of 43.33% error. In the second experiment, we find the coronary diseases, 43% as rates of correct and 57% rate of error, which were the lowest accurate results. The rest of experiments, in which we presented every arrhythmias independently of the others, we noticed a very high rate of correct recognition: concerning certain cases of arrhythmia: Old anterior myocardial infarction, old inferior myocardial infarction, sinus tachycardia, sinus bradycardia, ventricular premature contraction (PVC), block of branch left, block of branch right, left ventricular hypertrophy, atrial fibrillation or flutter which have the following values respectively : 97%, 95%, 99%, 96.33%, 96.66%, 100%, 90%, 99%, 99% and a rate of 19% of correct and 81% of error concerning the supra-ventricular premature contraction (ESA). The mean rate of recognition in 12 first experiments was 82.80% of correct detection. It was between 19% and 100% with a low rate in just one experiment. In the last experiment which all rules recognized 13 arrhythmias, the rate of correct classification was decreased until 38.33% of correct and 54.66 % of error, and reached a rate of sensibility 90.55%, and a rate of specificity 46.24%.

Table 3. The rate of correct diagnosis and faults.

Classes	Corrects (%)	Errors (%)
Normal	56.66	43.33
Ischemic changes (Coronary Artery Disease)	43	57
Old Anterior Myocardial Infarction	97	3
Old Inferior Myocardial Infarction	95	5
Sinus tachycardia	99	1
Sinus bradycardia	96.33	3.66
Ventricular Premature Contraction (PVC)	99.66	0.33
Supraventricular Premature Contraction (ESA)	19	81
Left bundle branch block	100	0
Right bundle branch block	90	10
Left ventricular hypertrophy	99	1
Arterial Fibrillation or Flutter	99	1
all arrhythmias	38.33	54.66

4. Discussion
The results generated from ten experiments were very high and proved the accuracy of RBS. In these experiments, the detection of just one arrhythmia by the modeled cardiologist's rules has ensured the top-quality diagnosis. In the rest of the experiments, 1) The normal case contains a very large number of rules which can be confused with other arrhythmias, also 2) experiments on ischemic changes and 3) ESA give a low rate of recognition. The above experiments demonstrate the significant source of conflicts which is due by normal, ischemic changes and ESA classes. The low rate of diagnosis in the global experiment is close to the mean, but it was less than the obtained in non-symbolic approaches (1). These weaknesses may have been due to the uncertainty of the values, missed data, or the conflict between rule symmetry of the parameters. In either case, the patient can generate several arrhythmias in parallel. As a related works (Table 4) VF15 is an algorithm of learning, overseen and inductive to lead the knowledge of classification from example. The representation of the knowledge is based on a technique called Intervals of Function, where a concept is represented by the projections of the cases of training on every function separately. Classification in VF15 is based on a majority voting among the predictions of class made by every option separately. The comparison of the algorithm of VF15 indicates a 62% rate and that it surpasses other standard algorithms such as Bayesian naïve Network which reached a rate of 50% and closest to classifiers Neighbors of a rate which reached 53 %, all results are published in (1).

Table 4. Comparison of the full rules experiments with published works.

Authors		CC (%)	SE (%)	SP (%)	The applied method	Results interpretation
This application	RBS	38.33	90.55	46,24	XML rules based system	Yes
- Burak Acar, Haldun Muderrisoglu and - H. Altay Guvenir (1)	VF15	62	//	//	Supervised and inductive learning algorithm	No
	NBN	50	//	//	Supervised and probabilistic learning algorithm	
	NN	53	//	//	Similarity based learning algorithm	

5. Conclusions
Cardiac diseases represent the first cause of death in the world. Several works were dedicated to decreasing cardiac risk. In this study, we were initially interested in the detection of the cardiac arrhythmias by implementing a system with rules in XML with the values collected from an electrocardiogram (ECG). Our finding in this study achieves to the influence of conflicts in the detection of multi-arrhythmias. It demonstrates also that the symbolic methods (rule based systems), through a graphical interface give a good classification, and the obtained results are clearer and transparent and more interpretable with regard to the non-symbolic methods.

Conflict of Interest:
There is no conflict of interest to be declared. We declare that this article is published for free funded by the journal because it wins the price of best presentation at ICHSMT'16

Authors' contributions:
All authors contributed to this project and article equally. All authors read and approved the final manuscript.

References:
1) Guvenir HA, Acar B, Demiroz G, Cekin A. A Supervised Machine Learning Algorithm for Arrhythmia Analysis. Proceedings of the Computers in Cardiology Conference: Lund, Sweden; 1997.
2) Khelassi A, Chick MA. Fuzzy knowledge-intensive case based classification for the detection of abnormal cardiac beats. Electron Physician. 2012; 4(2): 565-71.

3) Khelassi A, Chikh MA. Cognitive amalgam with a distributed fuzzy case-based reasoning system for an accurate cardiac arrhythmias diagnosis. International Journal of Information and Communication Technology. 2015; 7(4-5): 348-65.

4) UCI Machine learning Repository, depository of Center for Machine Learning and Intelligent Systems, Available from: http://cml.ics.uci.edu/

5) Swerdlow DI, Humphries SE. Genetics of CHD in 2016: Common and rare genetic variants and risk of CHD. Nat Rev Cardiol. 2017; 14(2): 73-4. doi: 10.1038/nrcardio.2016.209. PMID: 28054577.

6) Wang F. Genetics of coronary artery disease. Current Trends in Cardiology. 2017; 1(1): 1-2.

7) Delgado V, Gaemperli O, Lombardi M, Kaufmann PA, Bax JJ. The year in cardiology 2016: imaging. European Heart Journal. 2017.ehw633.

8) Li Y, Fang KL, Huang Z, Lu Y, Zhang B, Yao Y. Advancements in Cardiovascular Diagnostics. In Emerging Applications, Perspectives, and Discoveries in Cardiovascular Research. 2017; 194-211. doi: 10.4018/978-1-5225-2092-4.ch011.

9) Khelassi A. Distributed Case Based Reasoning for Cardiac Arrhythmias. Health sciences application. Lambert Academic publishing: LAP; 2012.

10) J Gay, Benoit P, Desnos M. L'électrocardiogramme, savoir l'interpréter: 460 tracés commentés et figures. Paris: Frison-Roche; 1990.