A decision tree-based approach for cardiovascular dysautonomias diagnosis

Ilham KADI and Ali IDRI
Software Project Management Research Team
ENSIAS, University Mohammed V
Rabat, Morocco
ilham.kadi@um5s.net.ma, idri.ali123@gmail.com

***Abstract*— Terms as knowledge Discovery from Databases (KDD), Data Mining (DM) and Machine Learning (ML), gain from day to day, an increasing significance in medical data analysis. They permit the identification, evaluation, and quantification of some less visible, intuitively unpredictable, by using generally large sets of data. Researchers have long been concerned with applying DM tools to improve data analysis on large data sets. DM has been increasingly used in medicine, particularly in cardiology. In fact, data mining applications can greatly benefits all parts involved in cardiology. Autonomic nervous system (ANS) is the part of the nervous system that is involved in homeostasis of the whole body functions. A malfunction in this system can lead to a cardiovascular dysautonomias. Thereby, a set of dynamic tests are adopted in ANS units to diagnose and treat patients with cardiovascular dysautonomias. In this paper, a case study was performed in order to construct a cardiovascular dysautonomias prediction system using data mining techniques and a dataset collected from an ANS unit of the Moroccan university hospital Avicenne. The prediction system is a decision tree-based classifier that was developed using C4.5 decision tree algorithm to automate the analysis procedure of ANS's test results and make it easier for specialists. The performance of the generated decision trees was evaluated and the results obtained achieved high accuracy rates which were very promising. In addition, a clinical validation of the developed system was carried out on new patients. In fact, a prototype of the developed system was implemented on JEE platform and deployed in the ANS unit so as to be validated clinically. The results were analyzed and thus the prototype was approved to be highly accurate, interpretable, time saving and easy to use.**

I. INTRODUCTION

DM is an analytic process designed to explore large datasets in hidden and previously unknown patterns, relationships and knowledge. It can be considered as knowledge discovery from data including three phases namely data pre-processing, data modeling and data post-processing [19]. Recent researches have shown that application of data mining in several fields is growing such as education [20, 21], clinical medicine [22] and financial fraud detection [23]. Classification is one of the main tasks of DM. classification techniques are capable of processing a large amount of data. They may predict categorical class labels and classifies data based on a training set [24]. Decision tree (DT) algorithms are considered as one of the popular classification and regression techniques. The goal of DT is creating a model that predicts the value of a target variable by learning simple decision rules inferred from the data features [25]. There are many specific decision-tree algorithms. Notable ones include: ID3, C4.5 [7], and CART [26].

C4.5 algorithm is one of the well-known decision tree algorithms because of its efficiency and comprehensive features [7]. Thereby, data miners have used this algorithm in different disciplines of medical field including cardiology [9]. Pavlopoulos, Stasis and Loukis [10] used the C4.5 algorithm to analyze different heart sound features, which assist clinicians to make a better diagnosis in CHD. Mašetić and Subasi have evaluated the effect of C4.5 decision tree in creating a model that will detect and separate normal and congestive heart failures (CHF) on the long-term ECG time series. Experimental results showed that C4.5 algorithm has significant role in identification and classification of ECG heartbeat signals with an accuracy of 99.86% [11]. Overall, the results obtained by studies applying C4.5 algorithm in cardiology were satisfactory. However, to the best of our knowledge, there is no existing study that applies data mining techniques and particularly C4.5 algorithm in the ANS field.

The autonomic nervous system (ANS) is the designation applied by John Langley [27] to a complex network of peripheral nerves and ganglia. It is the part of the nervous system that is involved in homeostasis by coordinating internal functions of the body and regulating unintentionally and automatically different organs including the cardiovascular system [1]. However, ANS is frequently subject to malfunctions that can cause serious problems and can be, in some cases, life-threatening. This is why, a set of dynamic tests are identified to allow the diagnosis of autonomic disorders. In fact, autonomic testing is designed to determine how well the body is regulating the blood pressure (BP) and heart rate (HR). In this paper, a case study was performed in the ANS unit of the Moroccan university hospital Avicenne. This unit is specialized in diagnosing patients with cardiovascular dysautonomias by performing several dynamic tests and provides them the appropriate treatment. The tests that are performed by specialists in this unit are: deep breathing [4], hand grip [5], mental Stress [5], and orthostatic test [6]. However, the analysis process in this unit is done manually by the specialists which can be hard and challenging, especially in the presence of several cases at one time. Thus, in order to help

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.121

those specialists, the aim of this study is to build a cardiovascular dysautonomias prediction system by applying C4.5 decision tree algorithm to a dataset of the ANS unit of the hospital Avicenne. This dataset contains the records of heart rate and blood pressure of the ANS unit's patients for the several dynamics tests. The data mining model was built and evaluated. Moreover, a clinical validation of the developed model was performed in the ANS unit. In fact, a prototype of the developed model was implemented on the JEE platform and deployed in the ANS unit so as to be validated clinically.

The article is organized as follow: Section 2 provides some details about the different tools used in this study. Section 3 presents a description of the experimental design. Section 4 presents and discusses the results obtained. Finally, the conclusion and future work are presented in Section 5.

II. Background

In this section, a detailed description of the ANS is presented. Thereafter, a brief presentation of C4.5 decision tree algorithm is introduced.

A. Autonomic Nervous System

The autonomic nervous system is a complex network of peripheral nerves and ganglia. This system is the part of the nervous system that controls involuntary actions, such as the beating of the heart and the widening or narrowing of the blood vessels. It controls, in particular, smooth muscle, heart muscle, some endocrine glands and the majority of exocrine glands. Thus, the ANS system controls blood pressure, heart and breathing rates, body temperature, digestion, metabolism (thus affecting body weight), the production of body fluids (saliva, sweat, and tears), and other processes [1]. The ANS is divided into two complementary systems: the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The balance of these two systems provides the balance of physiological functions [2].

The ANS is frequently subject to malfunctions that are called dysautonomias. Autonomic nervous system disorders can occur alone or as the result of another disease, such as Parkinson disease, alcoholism and diabetes [3]. This disorder can cause serious problems, including: Blood pressure problems, Heart problems, Trouble with breathing and swallowing and others. When they affect the breathing or heart function, these disorders can be life-threatening.

Doctors can check for signs of autonomic disorders during the physical examination. They measure blood pressure and heart rate while a person is lying down or sitting and after the person stands. In this paper, a case study was conducted in the ANS unit of the Moroccan university hospital Avicenne. This unit is specialized on performing ANS tests to diagnose patients with cardiovascular dysautonomias and provide them the appropriate treatment. According to the tests results, a set of preliminary conclusions is deducted. These conclusions are analyzed by the specialists to provide a global synthesis and diagnosis of the patient's state. The tests conducted by this unit are:

- Deep breathing (DB) [4]: it has a major interest in the determination of the vagal response (VR). It assesses autonomic function by measuring changes in HR in response to a deep breath. The calculation of (VR) is obtained by means of Eq. 1.

$$VR = \frac{HR\ \max - HR\ \min}{HR\ \min} * 100 \qquad (1)$$

- Hand Grip (HG) [5]: This is a manual effort contraction performed to determine changes in the BP in static effort. In normal condition, muscle contraction causes a rise in HR and BP. In this test, two values are measured: VR, by the same method as Deep breathing test, and Peripheral sympathetic alpha activity by means of Eq. 2.

$$\text{PSR}\ \alpha = \frac{BP\ \max - BP\ \min}{BP\ \min} * 100 \qquad (2)$$

- Mental Stress (MS) [5]: The patient performs mental arithmetic calculations. The result is an increase in BP and in HR by activation of the central sympathetic nerve. In mental stress, the central sympathetic nerves activities "α" was evaluated by measuring the variations of BP using Eq. 3:

$$\text{CSR}\ \alpha = \frac{BP\ \max - BP\ \min}{BP\ \min} * 100 \qquad (3)$$

  The central sympathetic nerves activities "β" was evaluated by measuring the variations of HR using Eq. 4:

$$\text{CSR}\beta = \frac{HR\ \max - HR\ \min}{HR\ \min} * 100 \qquad (4)$$

- Orthostatic test (Ort) [6]: it aims at measuring HR and BP variations in different positions: stand up and rest. In fact, the transition from rest position to a standing position causes a variety of physiological processes of adaptation in normal subjects and a variation in HR and BP. Thereby, several measures of HR and BP are taken in orthostatic test including: VR, basal state and supine position.

B. C4.5 decision tree algorithm: an overview

C4.5 is a standard algorithm for inducing classification rules in the form of decision tree. It was introduced by Quinlan [7]. It is an extension of the basic ID3 algorithm used to overcome its disadvantages. C4.5 algorithm made several improvements in order to enhance the ID3 algorithm. Some of these are:

1) Choosing an appropriate attribute selection measure.
2) Handling training data with missing attribute values.
3) Handling attributes with differing costs.
4) Pruning the decision tree after its creation.
5) Handling continuous attributes.

C4.5 algorithm builds a decision tree from a set of training data similar to the ID3 algorithm, using the concept of information entropy. It uses the divide-and-conquer approach to decision tree induction. The algorithm uses a selected criterion to build the tree. It works top–down, seeking at each stage an attribute to split on that which best separates the classes, and then recursively processing the sub problems that result from the split [8]. In addition, C4.5 algorithm uses heuristics for pruning derived based on the statistical significance of splits.

Sample algorithm

- Check for base cases.
- For each element x, discover the normalized information gain ratio from splitting on x.
  - **Gain ratio** applies a kind of normalization to information gain using a **split information** value
  - **Split information value:** represents the potential information generated by splitting the training data set D into v partitions, corresponding to v outcomes on attribute A:

$$\text{SplitInfo}_A(D) = -\sum_{j=1}^{v} \frac{|Dj|}{|D|} * \log_2 \left(\frac{|Dj|}{D}\right) \quad (5)$$

  - The gain ratio is defined as:

$$\text{GainRatio}(A) = \frac{Gain\ (A)}{SplitInfo\ \ (A)} \quad (6)$$

  - The attribute with the maximum gain ratio is selected as the splitting attribute
- Select the best x, attribute that has highest information gain.
- Create a decision node that breaches on best of x.
- Recurs on the sub lists obtained by splitting on best of x, and add those nodes as children of node.

III. Material and methods

The development of our data mining system went through the following major phases: data pre-processing, modeling, evaluation and validation. In this section, the tasks performed in each phase will be presented and explained.

A. Medical dataset description

A total of 178 records with 66 features were collected from the ANS unit belonging to the cardiology department of university hospital Avicenne in Morocco. Some of these features provide general and administrative information about the patient such as: name of patient, file reference, date of consultation and the attending physician. These data do not affect the patient diagnosis which is why they were discarded. Only the attributes judged by the specialist to be required for the diagnosis of patients were selected. Table I provides a brief description of each attribute as well as some statistics such as mean, max, and the min of each selected attribute. According to Table I, the patients diagnosed by ANS units are from all generations including the children and the oldest persons. For the other attributes, the age of patients is considered as the main factor influencing the results interpretation. In fact, a normal value of an ANS test result differs according to age category.

an example, for the VR_DB attribute, a normal value should be near to 30% for adult patients (18<Age<60). However, for the elderly, a normal value is around 25%; and for children and young patients, a normal value is close to 60%. Thereby, the same principle applies to VR_HG, PSR α, CSR α, CSR β and VR_Ort attributes except that the critical values differ from one test to another. Nevertheless, we notice that the average value for VR_DB is 46.23% which shows that a lot of patients suffer from difficulties in case of breathing efforts. In general, a normal HR value should be between 60 beats/min and 80 beats/min, and a normal value of BP should be between 100 and 140 for systolic values. However, according to Table I, the min and max values detected have far exceeded the normal one for both HR and BP which shows that there are some patients that are suffering from serious problems that need to be treated urgently.

TABLE I. DESCRIPTION AND STATISTICS OF SELECTED ATTRIBUTES

Symbol	Quantity	Mean	Min	Max
Age	Age of the patient	42.3	7	84
VR_DB	Vagal response measured using HR values in DB test	46.2	4	155
VR_HG	Vagal response measured using HR values in HG test	19.9	0	66
PSR α	Peripheral sympathetic response α measured using BP values in HG test	23.3	1	72
CSR α	Central sympathetic response α measured using BP values in MS test	17.1	2	67
CSR β	Central sympathetic response β measured using BP values in MS test	18.6	1	95
VR_Ort	Vagal response measured using HR values in Ort test	21.2	1	80
HR_{min}	Minimum heart rate measured in Ort test	61.5	17	104
HR_{max}	Maximum heart rate measured in Ort test	70.2	38	165
BP_{min}	Minimum blood pressure measured in Ort test	114.4	84	185
BP_{max}	Maximum blood pressure measured in Ort test	125.9	89	193

B. Pre-processing

Data preprocessing is a very important step in a data mining process. It is a critical step which deals with the preparation and transformation of the initial data. In fact, analyzing data that has not been carefully screened can produce misleading in results. Thereby, the quality and representation of data is first and foremost before running an analysis [13]. An initial dataset can generally gather several problems such as: missing values, noisy data and inconsistency [14]. For this reason, several methods were developed to solve these problems and improve the data quality. These methods can be divided in [15]:

1) Data cleaning: Fill in missing values, smooth noisy data, identify or remove outliers, and resolve inconsistencies.
2) Data integration: Integration of multiple databases, data

cubes, or files.

3) Data transformation: Normalization and aggregation.
4) Data reduction: Obtains reduced representation in volume but produces the same or similar analytical results.

The database adopted in this study, contained a few missing values that did not exceed 4%. Thus, a data cleaning process is required. After examining the nature of the missing values, we noticed that these latter were a result of one of the mathematical formulas aforementioned. Thereby, the missing values were filled manually by performing the required calculations. Furthermore, several attributes contained numeric values including real numbers which were all transformed to integer values. This transformation was carried out by keeping just the integer part in order to simplify the generation of decision trees. On the other hand, the classes identified to generate the decision tree with C4.5 algorithm were coded. In fact, to run C4.5 algorithm, three files are required: names file, data file and test file [7]. The data and test files contain the training dataset and test dataset respectively. The names file includes details about the identified classes and attributes: their names and types (discrete or continuous). However, this file requires that the attributes have to be coded: discrete attributes are coded by numbers and continuous attributes are only identified as continuous. Thus, the classes identified for this experiment were coded and included in the names file as follow: high: 2, normal: 1 and low: 0. Regarding the other input attributes of Table II, they are not identified as discrete attributes so they were indicated as continuous in the names file.

C. Classifier modeling

The modeling phase is related to the discovery of relationships between various data in order to extract hidden patterns. As previously explained in Section II, a cardiovascular dysautonomias patient's diagnosis is based on several preliminary conclusions of ANS's tests. For each test, one or more preliminary conclusions are identified. Each conclusion requires a set of input attributes and needs to be classified. Thereby, a deep analysis was carried out to determine the input data needed for the generation of each conclusion and identify the predefined classes. This information was required to apply C4.5 algorithm and generate the decision tree so as to produce a decision support system for cardiologist. In fact, to identify the input attributes for this case study, we integrated the ANS unit and attended the elaboration phase of diagnosis and treatment. Thus, through several observations and based on the specialists guidelines, the input attributes were identified and used by C4.5 algorithm.

Table II presents all information extracted in order to carry out a classification with C4.5 algorithm. Table II was designed by analyzing each test separately and identifying all the necessary attributes based on the empirical knowledge of ANS experts. The attributes Age is common between the different tests which show that it has a great influence on the results interpretation. As an example, in mental stress test two important values need to be measured: CSR α and CSR β by means of formulas 3 and 4. These measures are analyzed separately by the specialists depending on the age factor. Thereby, two preliminary conclusions are identified for the mental stress test, one for the CSR α value, and another for the CSR β value. These conclusions identify whether the CSR α and CSR β values depending on the age are interpreted as *high*, *normal* or *low* values; consequently, two decision trees were generated for this test. For the other tests, one or more decision trees were generated. The classes that were identified to be used by C4.5 algorithm are *high*, *normal* or *low* for all tests. As a result, eight decision trees were generated and tested.

TABLE II. DETAILS ABOUT INPUT DATA AND CLASSES FOR EACH ANS TESTS

ANS tests	Measured values	Input attributs	Class
Deep Breathing test	Vagal response	Age VR	High Normal Low
Hand Grip test	Vagal response	Age VR	High Normal Low
	PSR α	Age PSR	High Normal Low
Mental Stress test	CSR α	Age CSR α	High Normal Low
	CSR β	Age CSR β	High Normal Low
Orthostatic test	Vagal response	Age VR	High Normal Low
	SP_FC	Age HR_{min} HR_{max}	High Normal Low
	SP_TA	Age BP_{min} BP_{max}	High Normal Low

D. Cardiovascular dysautonomias prediction system: evaluation and validation

The evaluation phase is required to assess the performance of the generated classifiers. For this reason, the data set was divided into two sets training set and testing set. The decision trees were generated using the training sets and evaluated using the testing sets.

The validation phase is adopted to validate the model clinically. In fact, after evaluating the data mining model, a clinical validation is required to ensure the efficiency and reliability of the developed system. Thereby, a web-based prototype was developed on the JEE platform to enhance visualization and ease of interpretation. The prototype is a web presentation of our classification model which allows experts to enter the HR and BP values and have as an output the preliminary conclusions generated by the decision trees. As aforementioned, The VR, PSR α, CSR α, CSR β values are calculated using the mathematical equations 1, 2, 3 and 4. Thus, the prototype allows entering just the measured values of HR and BP in the different ANS tests. The mathematical calculations and the preliminary conclusions generation are performed automatically by the prototype.

Fig. 1 presents a screenshot of the web based prototype. The screenshot shows the implemented form where input attributes are entered to be processed by the developed system. As shown in Fig. 1, only HR and BP values of ANS tests are required; which will make the analysis of ANS tests results easier for specialists and prevent them from doing all the work manually.

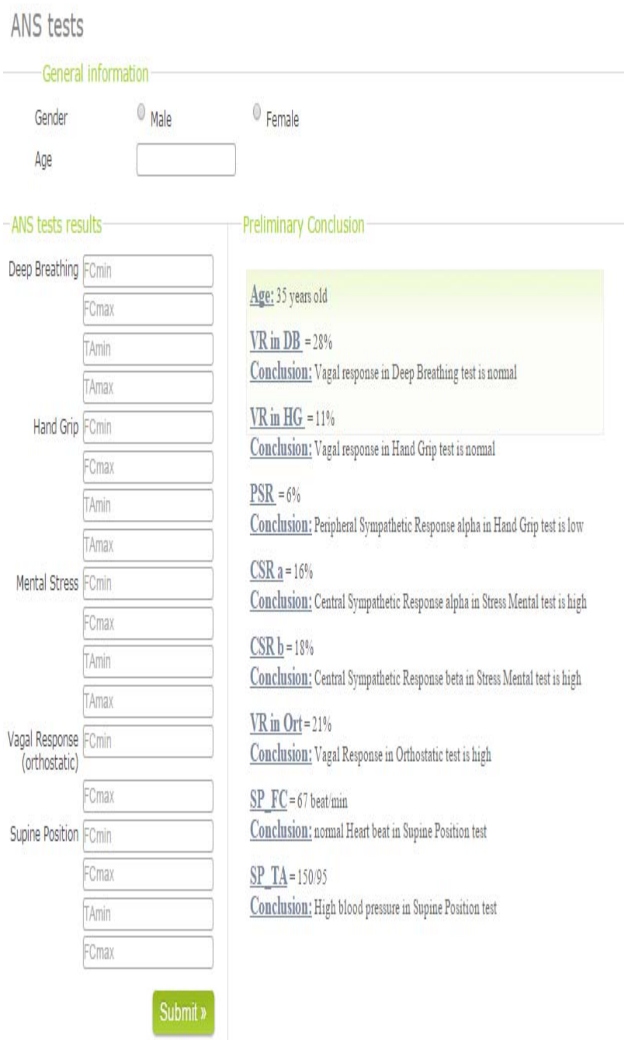

Fig. 1. Output of preliminary conclusions for a 35 years old patient

IV. Results and Discussion

A. Evaluation of the data mining model

In order to evaluate the efficiency of the generated decision trees, the data set was divided into two sets training (123 records) and testing set (55 records). The decision trees were generated using the training set and validated using the testing set. In fact, C4.5 algorithm was executed under the Ubuntu distribution of Linux operating system using a C4.5 software release 8. More details about this C4.5 software are available in the following website [12] where the download link and the instructions for use are provided. The names, data and test files required for the execution of C4.5 algorithm were constructed. Then, the decision trees were generated through several commands. 10 trials were carried out in this experimentation. Data and test files were changed in each trial. Fig. 2 presents an example of a generated decision tree regarding DB test. As previously explained, the age attribute have a great influence on the interpretation of the VR values which is why the age was identified as the root node of the generated decision tree.

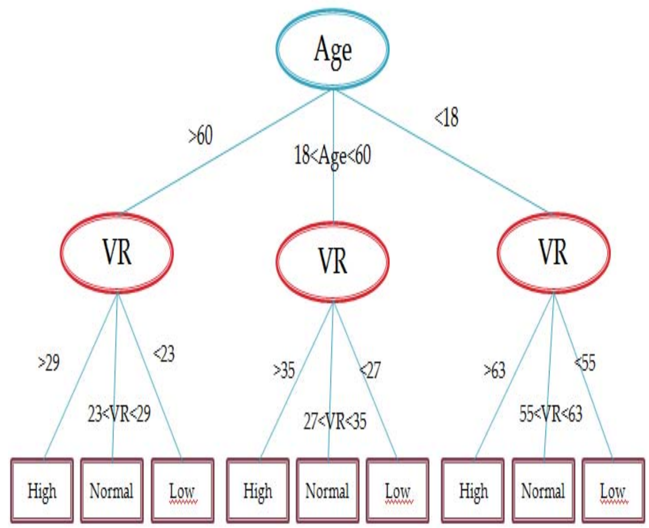

Fig. 2. Example of a generated decision tree regarding DB test

1) *Learning phase :*

In this experimentation, 10 trials were carried out to generate and test the decision trees for each ANS test. In each trial, the training sets were changed and the error rate measured was recorded. Table III presents the performance results in terms of error rate on the training set for each ANS test. Class distribution for all ANS tests was evaluated and recorded; thus, the approximate rate of each class in the different tests was provided as follow: 41% of the data were identified as *high* class, 31% as *normal* class and 28% as *low* class.

TABLE III. ERROR RATES OF THE GENERATED CLASSIFIERS IN TRAINING SET

ANS tests	Phase	Mean error rate
Deep Breathing	Vagal response	2.15%
Hand Grip	Vagal response	3.99%
	PSR α	0.81%
Mental stress	CSR α	0.85%
	CSR β	0%
Orthostatic	Vagal response	0%
	SP_FC	1.28%
	SP_TA	2.54%

According to these results, we can notice that all results were close and so there is no majority class and. The values of Table III are the mean of error rate values obtained in the

10 trials for each generated decision tree. According to the results of Table III, the mean values of the error rate are low, which contributes to the increase of the accuracy rate up to 98.54%. These results may be explained by the fact that input features required for the construction of each decision tree did not include a lot of input attributes. In fact, as shown in Section 4, the number of input attributes did not exceed four attributes which help to produce classifiers with high accuracy rates [7, 8].

2) Testing phase :

The generated decision trees were tested using testing sets which are different of training set. These testing sets were also changed in each trial. The error rates measured for these sets were recorded. Table IV summarize all the results obtained regarding the testing phase and presents the mean value of error rate obtained in the 10 trials for each generated decision tree. The results obtained in the testing phase were also satisfactory and the values of the error rates recorded were low. Thereby, the classifiers of this study achieved high accuracy rates up to 98.54% for training set and 97.76% for testing set respectively.

TABLE IV. ERROR RATES OF THE GENERATED CLASSIFIERS IN TESTING SET

ANS tests	Phase	Mean error rate
Deep Breathing	Vagal response	0.38%
Hand Grip	Vagal response	2.15%
	PSR α	3.01%
Mental stress	CSR α	0%
	CSR β	1.72%
Orthostatic	Vagal response	0%
	SP_FC	8.96%
	SP_TA	1.66%

3) Comparison with other classifiers :

After evaluating the decision trees generated using testing sets. A comparison between the accuracy rates obtained using C4.5 algorithm, K-NN and Naïve Bayes (NB) classifiers was carried out in order to assess the performance of our system,. The K-NN and NB classifiers were performed using the Tanagra 1.4 software. Table V shows the results obtained when applying K-NN and Naïve Bayes classifiers on our data set. These classifiers were applied on training and testing sets. The mean, max and min values of accuracy rates for each classifier were recorded. In fact, when running the predefined classifiers, several trials were conducted for each ANS test to identify the appropriate neighborhood size for K-NN and the Lambda parameter for Naïve Bayes. Thus, the best results were obtained using a neighborhood size between 3 and 10, and a default Lambda parameter equals to 1.0. According to Table V, C4.5 have presented the best accuracy rates comparing to K-NN and Naïve Bayes that did not exceed 97.56% and 93.18% respectively for training sets and 92.73% and 89.79% respectively for test sets. These classifiers have achieved good performance but still lower in comparison with the performance of C4.5 algorithm.

TABLE V. COMPARISON OF ACCURACY RATES OBTAINED USING C4.5, K-NN AND NAIVE BAYES CLASSIFIERS

Classification techniques	Training sets			Test sets		
	Mean (%)	Min (%)	Max (%)	Mean (%)	Min (%)	Max (%)
C4.5	98.5	96.01	100	97.7	91.1	100
K-NN	96.4	95.1	97.5	85.8	83.3	92.7
Naïve Bayes	89.1	85.2	93.1	85.7	83.8	89.8

B. Validation of the prototype cardiovascular dysautonomias prediction system

After assessing the developed data mining model, we moved to the clinical validation phase using the web prototype. In fact, once the prototype was implemented, it was approved first by the cardiologist responsible for ANS unit. Then, the prototype was deployed in the ANS unit so as to be evaluated on a set of new patients different from the dataset used in the evaluation phase. For this purpose, a set of 85 patients who went to ANS unit in the period between July and November 2014 was selected. These patients were subject of cardiovascular dysautonomias. Thus, they performed the different dynamic tests required to diagnose this kind of autonomic disorder. The data set selected include patients of different ages from children to older people which enable to assess the prototype on different possible cases. The ANS tests results were analyzed and interpreted by the specialists. The preliminary conclusions identified by the specialists were compared to those generated from the implemented prototype. The results of this comparison were analyzed based on three main goals namely: accuracy, interpretability and usability.

1) Accuracy :

Accuracy is one of the most popular performance evaluation criteria. It aims at measuring the rate of correct classification [9]. For classification models, error rate is used as accuracy metric to evaluate their performance. Thereby, in order to measure the error rate of the implemented prototype when validating it in ANS unit, the preliminary conclusions generated by each decision tree were analyzed separately to determine if it complies with specialists conclusions. The result of each comparison was recorded whether or not there was compliance. This operation was repeated for each patient record. Eventually, the non-compliance cases recorded in each decision tree were collected to calculate the error rate. Table VI summarizes the results of the error rates obtained in clinical validation. According to the results presented in Table VI, the error rates recorded in clinical validation are low which contributes to the increase of the accuracy rate up to 99.12% which is very promising.

However, as shown in Table VI, there were some cases of non-compliance between the preliminary conclusions generated by the prototype and those identified by the specialists; which can be explained by the fact that some patients were suffering from several dysfunctions at the same time which affects the results interpretation. These exceptions are considered as critical cases and needs a deeper analysis by the specialists to produce a correct conclusion. As an example, a 34 years old patient was suffering from diabetes type 1 for 20 years. She had instable diabetes with unexplained hypoglycemia and hyperglycemia. As a result, the patient was subject of an autonomic dysfunction and underwent the different tests of the ANS units. The results obtained were entered and processed by the implemented prototype. The preliminary conclusions generated by the prototype were compared to those identified by the specialists. These latter were consistent with the conclusions obtained by the prototype except for CSR β phase of mental stress test. In fact, the patient recorded 12% regarding the central sympathetic response β which is normally interpreted as a normal value for adult patients. However, the specialists interpreted this value as a high central response β which is due to the functional symptoms presented in this case. Thus, a non-compliance case between the conclusions of the prototype and the specialists was recorded.

TABLE VI. ERROR RATES OF THE GENERATED CLASSIFIERS IN CLINICAL VALIDATION

ANS tests	Phase	Mean error rate
Deep Breathing	Vagal response	0.29%
Hand Grip	Vagal response	1.89%
	PSR α	0.97%
Mental stress	CSR α	0.71%
	CSR β	1.12%
Orthostatic	Vagal response	0.32%
	SP_FC	0.67%
	SP_TA	1.02%

2) Interpretability :

In many domains, interpretability of a predictive model is a fundamental quality characteristic [16]. Several experts tend to prefer models that are more transparent rather than black-box predictive models, because they provide a clear sight on which factors were used to make a particular prediction. Interpretable models can be very convincing, particularly when only a few key factors are used, and each of them is meaningful. When we first started developing our decision support system, we aimed to build classification models that are accurate as well as easy to interpret. Thereby, we transformed the decision trees models previously generated to the natural language statements that consists of a series of *if... then...* statements where the *if* statements define a partition of a set of features and the *then* statements correspond to the outcome of interest. Using this method, the classification models have become easier to understand and interpret. These models were validated and approved to be interpretable by the experts. Fig. 3 presents an example of a part of the decision tree generated regarding Deep breathing test for a patient who is 35 years old using *if... then...* statements. Fig. 3 shows the different possible interpretation according to VR value for a 35 years old patient which greatly facilitates the deduction of preliminary conclusions. Obviously, this structure differs according to the patient age.

if VR >35 **then** vagal response for Deep Breathing test is high

else if VR<27 **then** vagal response for Deep Breathing test is low

else vagal response for Deep Breathing test is normal

Fig. 3. Example of a part the decision tree regarding Deep breathing test for a patient who is 35 years old.

3) Usability :

Usability is a quality attribute that assesses how easy user interfaces are to use. It also refers to methods for improving ease-of-use during the design process [17]. In other words, usability lies in the interaction of the user with the product or system and can only be accurately measured by assessing user performance, satisfaction and acceptability. A product is not itself usable or unusable, but a set of metrics are identified to determine the usability for a particular user, task and environment. ISO 9126 defines Usability in terms of five sub-characteristics: Understandability, Learnability, Operability, Attractiveness, and Usability Compliance [18]. In this study, we tried to assess the usability of the implemented prototype especially learnability metric. Learnability means the capability of the software component to enable the user to learn the application. It enables to assess how long it takes system developers to learn how to use particular functions. In fact, the prototype was designed with the aim to facilitate as much as possible the work of specialists. For this reason, the prototype requires only entering the measured values of HR and BP. The mathematical calculations and the preliminary conclusions are generated automatically and presented as a final result for this stage. Thus, the prototype provides a time saving and ease of use in comparison with the existing method where all the work is done manually.

V. CONCLUSION

In this paper, a case study about the application of C4.5 decision tree algorithm was conducted using a dataset extracted from the ANS unit of university hospital Avicenne in Morocco. The objective of this study was to produce a decision support system to automate the analysis procedure of the ANS's test results and make it easier for specialists. In fact, these tests are adopted by specialists to check for signs of autonomic disorder. The results obtained need to be analyzed to deduce preliminary conclusions. These preliminary conclusions are analyzed with other parameters by the specialists to provide a diagnosis of the patient’s state and prescribe the appropriate treatment. Thereby, C4.5 decision tree algorithm was used to define a set of rules helping to generate the preliminary conclusions in the first stage of ANS procedure. The performance of the generated decision trees was measured by calculating error rates values obtained in training and testing sets. Thereby, the classifiers of this study achieved high accuracy rates up to 98.54% for training set and 97.76% for testing set respectively.

In addition, a comparison between the accuracy rates obtained using C4.5 algorithm, K-NN and Naïve Bayes classifiers was carried out so as to assess the performance of our system. These classifiers have achieved good performance but still lower comparing to the performance reached by C4.5 algorithm.

Moreover, a clinical validation of the cardiovascular dysautonomias prediction system on new patients was carried out. Thereby, a JEE platform was adopted to develop the prototype which was deployed in the ANS unit of Avicenne hospital. Several tests were carried out in order to compare the preliminary conclusions provided by the specialists and those generated by our prediction system. The results were analyzed based on three main goals namely: accuracy, interpretability and usability. Thus, the prototype was approved to be highly accurate, interpretable, time saving and easy to use. In fact, the prediction system developed can automate the analysis procedure of the ANS's test results and make it easier for specialists. It can also provide decision support for cardiologists to assist them and help them to make better clinical decisions or at least provide them a second opinion. Besides, it can serve as a training tool for nurses and medical students to train them diagnose patients with cardiovascular dysautonomias.

As a limitation of this system, the size of the dataset used in this research is still quite small. A large dataset would definitely give better results. Moreover, the current cardiovascular dysautonomias prediction system is not yet accomplished. The second stage of the ANS procedure needs to be automated and integrated in the system. In fact, the prototype was deployed in the ANS unit just for validation purposes. Thereby, the system will not be adopted by the specialists until the whole system is completed. For future work, the second stage of the ANS procedure needs to be automated using classification and association techniques. Thereby, a complete cardiovascular dysautonomias prediction system that provide a diagnosis for patients and suggest the appropriate treatment will be produced.

ACKNOWLEDGMENT

This research is part of the project "mPHR in Morocco" financed by the Ministry of High education and Scientific research in Morocco 2014-2016. Our gratitude to Professor Halima Benjelloun, Head of the Cardiology Department in Avicenne hospital, for providing us data to perform this case study.

REFERENCES

[1] S. D. Kreibig, "Autonomic nervous system activity in emotion: a review". Biological Psychology, vol.84, pp. 394–421, April 2010.

[2] E. Benarroch, "The central autonomic network: functional organization, dysfunction and perspective". Mayo Clinic Proceedings, vol.68, pp. 988–1001, October 1993.

[3] B. P. Grubb, B. Karas, "Clinical Disorders of the Autonomic Nervous System Associated With Orthostatic Intolerance: An Overview of Classification, Clinical Evaluation and Management". Pacing and Clinical Electrophysiology, vol. 22, pp. 798–810, May 1999.

[4] R. W. Shields, "Heart rate variability with deep breathing as a clinical test of cardiovagal function". Cleveland Clinic Journal of Medicine, pp. 37–40, April 2009.

[5] T.L. Johansen, G. Kambskar, J. Mehlsen, "Heart rate variability in evaluation of the autonomic nervous system". Ugeskr Laeger, vol. 159, pp. 6666–6671, 1997.

[6] A. R. Mejía-Rodríguez, M. J. Gaitán-González, S. Carrasco-Sosa, and A. Guillén-Mandujano, "Time varying heart rate variability analysis of active orthostatic and cold face tests applied both independently and simultaneously". Computers in Cardiology7, C4.5 Programs for Machine Learning, Morgan Kaufmann publisher, pp. 361–364 2009.

[7] J. R. Quinlan, "C4.5 Programs for Machine Learning" in CA: Morgan Kaufmann, 1993.

[8] J. Han, M. Kamber, "Data Mining, Concepts and Techniques". Morgan Kaufmann publisher, 2001.

[9] N. Esfandiari, M. R. Babavalian, A. E. Moghadam, V. Tabar, "Knowledge discovery in medicine: Current issue and future trend". Expert Systems with Applications, vol. 41, pp. 4434–4463, July 2014

[10] S. A. Pavlopoulos, A. C. Stasis, E. N. Loukis, "A decision treebased method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds". BioMedical Engineering OnLine, June 2004

[11] Z. Mašetić, A. Subasi, "Detection of congestive heart failures using C4.5 Decision Tree". Southeast Europe Journal of Soft Computing, vol. 2, pp. 74–77, 2013.

[12] http://www2.cs.uregina.ca/~dbd/cs831/notes/ml/dtdtre/c4.5/tutorial.html

[13] J. Han, M. Kamber, J. Pei, "Data preprocessing". The Morgan Kaufmann Series in "Data Management Systems", Morgan Kaufmann Publishers. 2011.

[14] I. H. Witten, E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations". San Mateo, CA: Morgan Kaufmann, 1999.

[15] A. Familia, W. M. Shenb, R. Weberc, E. Simoudis, "Data preprocessing and intelligent data analysis". Intelligent Data Analysis, vol. 1, pp 3–23, 1997.

[16] A. Vellido, J. D. Martin-Guerrero, P. J. G Lisboa, "Making machine learning models interpretable". in Proc. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2012.

[17] N. Bevan, J. Kirakowski, J. Maissel, "What is Usability?." in Proceedings of the 4th International Conference on HCI, Stuttgart, 1991.

[18] M. F. Bertoa, A. Vallecillo, "Usability metrics for software components". 8th International Workshop on Quantitative Approaches in Object-Oriented Software Engineering, Oslo, Norway, 2004.

[19] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, "From data mining to knowledge discovery in databases". AI Magazine, vol. 17, pp. 37–54, 1996.

[20] C. Romero, S. Ventura, "Educational data mining: A review of the state of the art". IEEE Transaction on Systems, Man and Cybernetics Part C, vol. 40, pp. 601–618, October 2010.

[21] C. Romero, S. Ventura, E. Garcia, "Data mining in course management systems: Moodle case study and tutorial". Computers & Education, vol. 51, pp. 368–384, August 2008.

[22] R. Bellazzi, B. Zupan, "Predictive data mining in clinical medicine: Current issues and guidelines". International Journal of Medical Informatics, vol. 77, pp. 81–97, 2008.

[23] E. Kirkos, C. Spathis, Y. Manolopoulos, "Data mining techniques for the detection of fraudulent financial statements". Expert Systems with Applications, vol. 32, pp. 995–1003, 2007.

[24] R. Aparna, G. Bincy, T. Mathu, "Survey on common data mining classification Technique". In International Journal of Wisdom Based Computing, 2012.

[25] C. Apté, S. Weiss, "Data mining with decision trees and decision rules". Future Generation Computer Systems, vol.13, pp 197–210, November 1997

[26] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, "Classification and regression trees". Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, pp. 246–280, January 1984.

[27] J. N. Langley, "The Autonomic Nervous System". in Cambridge Heffer, 1921

[28] P. A. Low, "Laboratory evaluation of autonomic function". In Clinical Autonomic Disorders. Evaluation and Managment, 1997.