

# A Multiple Decision Tree-based Method for Differentiation of a Split First Heart Sound from a Fourth Heart Sound and Ejection Click

*Antonis Stasis, Sotiris Pavlopoulos, Euripides Loukis\**

School of Electrical and Computer Engineering, National Technical University of Athens, \*University of Aegean, Dept. of Information and Communication Systems Engineering, Samos, Greece.

## ABSTRACT

**Objective:** Differentiating a fourth heart sound (S4), from a split first heart sound (SP1), or ejection click (EC), is often difficult particularly for inexperienced clinicians. The objective of this study was to develop and evaluate a computer-assisted classification tool to aid in this difficult differentiation problem, and in general for heart sound differentiation and diagnosis.

**Design:** Developmental study.

**Methods:** Emphasis was given to the selection of appropriate features that are adequately independent from the heart sound signal acquisition method. Relevance analysis was initially performed to identify the features of the heart sound most relevant to aiding diagnosis of S4, SP1 and EC. To detect and differentiate S4, SP1 and EC, a detection decision tree (DeDT) and a differentiation decision tree (DiDT) were used independently and also together in a multiple decision tree architecture. The DeDT provides three suggestions for each heart sound pattern, whereas the DiDT provides one. The MuDT analyses the suggestions of both decision trees to provide one final suggestion for each sound pattern.

**Results:** Relevance analysis on the different heart sound features demonstrated that the most relevant features for aiding diagnosis of S4, SP1 and EC are the frequency features and the morphological features that describe S1. The DeDT architecture demonstrated an average classification accuracy of 80.56%, sensitivity of 70.93%, and specificity of 83.42%, but provided more than one suggestion for many cases. The DiDT architecture demonstrated an average classification accuracy of 66.46%, a sensitivity of 66.15% and a specificity of 82.15%, and only provided one suggestion for each case. The MuDT architecture slightly improved performance compared to the DiDT architecture. Average classification accuracy was improved by 2.79%, classification sensitivity by 2.73% and classification specificity by 1.26%

**Conclusions:** The present work has demonstrated that decision tree algorithms can be successfully used as the basis for a decision support system to assist inexperienced clinicians in heart sound diagnosis. Further work is currently in progress to improve the accuracy, specificity and sensitivity of the system.

---

**Correspondence and reprint requests:** Dr. Sotiris Pavlopoulos, Biomedical Engineering Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, Zografou Campus, Athens 15773, Greece. E-mail: spav@biomed.ntua.gr.

## INTRODUCTION

Auscultation of the heart is a cheap screening method for cardiac pathology and is performed as part of the routine clinical examination. Experienced physicians can often diagnose cardiac pathology on the basis of auscultation alone. Its importance as a diagnostic tool has, however, declined as echocardiography is routinely used today to investigate patients with suspected cardiac pathology. Echocardiography can provide both anatomical and physiological information and its value in aiding accurate diagnoses is well-established. However, it should be remembered that the initial request for an echocardiogram or a cardiologist consultation is usually based on the initial auscultatory findings. This is frequently performed by the patient's general practitioner who may not be experienced or confident in cardiac auscultation<sup>1</sup>. For these physicians a decision support system to assist them in diagnosing different heart sounds and helping them differentiate similar heart sounds would be helpful<sup>2</sup>. Such a system would be based on acquiring and codifying the relevant knowledge of experienced cardiologists and making it available to them.

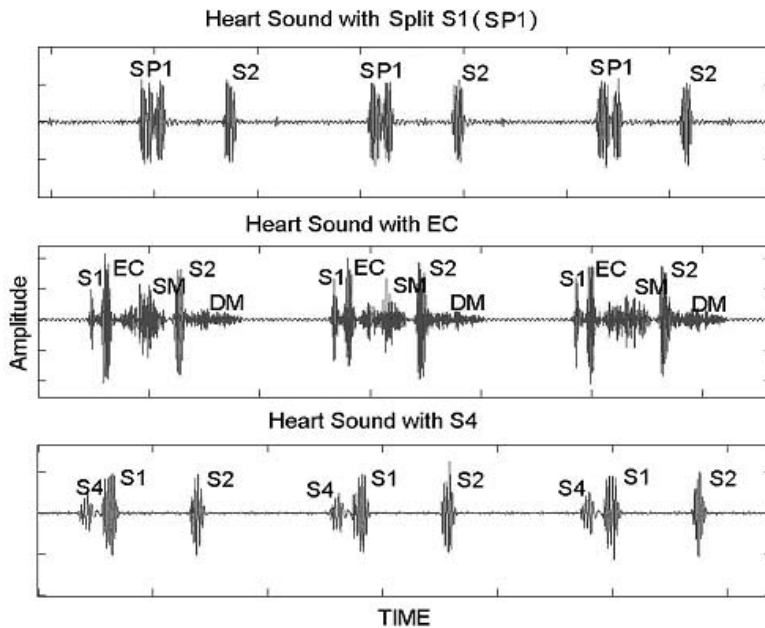
In the past computer-assisted heart sound diagnosis has been treated as a classification problem. Classification algorithms were mainly based on:

- i) Discriminant analysis<sup>3</sup>
- ii) Nearest neighbour<sup>4</sup>
- iii) Bayesian networks<sup>5</sup>
- iv) Neural networks<sup>6,7</sup>
- v) Rule-based methods<sup>2,8,9</sup>

These different approaches have been necessary because heart sounds have more than one characteristic morphology, e.g. timing in the cardiac cycle, duration and character of murmurs, and different pathologies, e.g. aortic stenosis and mitral regurgitation, can produce similar heart sounds.

A normal heart sound consists of four components. These are the first heart sound (S1), the systolic phase, the second heart sound (S2) and the diastolic phase. Additional sounds, such as murmurs or click-like sounds are heard in patients with a variety of heart diseases.

Normally the mitral and tricuspid valves close simultaneously and are heard as a single first heart sound. If for any reason closure of the tricuspid valve is delayed, the two components of the first heart sound will be heard separately and this is referred to as a split first heart sound (SP1). Delayed closure of the tricuspid valve may, for example, occur with right bundle branch block (delayed contraction of the right ventricle) or an atrial septal defect (increased blood flow through the right ventricle)<sup>10</sup>. Ejection clicks are often heard shortly after S1. They are often caused by valve abnormalities, e.g. aortic or pulmonary stenosis<sup>10</sup>. The fourth heart sound (S4) is a click-like sound that is heard at the end of the diastole, just before S1. S4 is thought to be due to forceful atrial contraction and occurs in conditions when



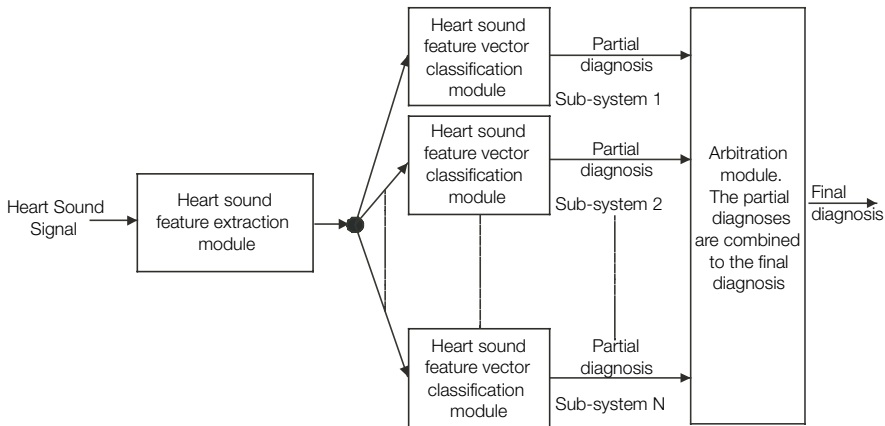
**Figure 1.** Timing of split first heart sound (SP1), ejection click (EC), fourth heart sound (S4), diastolic murmur (DM) and systolic murmur (SM) in the heart cycle

ventricular compliance is impaired, e.g. ventricular hypertrophy or fibrosis<sup>11</sup>. S4 is not heard in normal subjects.

Figure 1 shows the timing of S4, SP1 and EC in relation to the cardiac cycle, and Table 1 lists some of the common condition in which they occur. It should be noted that in many of these cases the patient is asymptomatic and abnormal cardiac auscultation is an incidental finding. However, detection of this abnormal heart sound is important to ensure early diagnosis and optimal management of

**Table 1.** Clinical conditions in which SP1, EC and S4 can be heard

Clinical conditions with a split first heart sound (SP1)	Clinical conditions with an ejection click (EC)	Clinical conditions with a 4th heart sound (S4)
Atrial septal defect	Aortic Stenosis	Aortic stenosis
Right bundle branch block	Bicuspid aortic valve	Severe systemic hypertension
Left ventricular ectopics	Aortic regurgitation	Pulmonary hypertension
Tricuspid stenosis	Pulmonary stenosis	Hypertrophic cardiomyopathy
Coarctation of the aorta	Eisenmenger's syndrome	Ventricular hypertrophy
Normal (i.e no cardiac pathology)	Pulmonary hypertension	Ventricular fibrosis
		Myocardial ischaemia
		Myocardial infarction



**Figure 2.** *Integrated decision support system architecture for heart sound diagnosis*<sup>9</sup>

abnormal cardiac conditions, e.g. antibiotic prophylaxis for dental procedures in a patient with a bicuspid aortic valve.

From Figure 1 the difficulty in clearly differentiating sounds heard around S1 can be appreciated. In this work we propose a method that uses time-frequency features and decision tree classifiers for addressing this problem. We have attempted to develop a computer-based assisted system for analysing the morphological characteristics of heart sounds and in particular for detecting and differentiating a fourth heart sound (S4), from a split first heart sound (SP1) or ejection click (EC).

The approach adopted has been to divide heart sound diagnosis into a number of simpler sub-problems, each of them dealing either with a morphological characteristic of the heart sound signal, e.g. timing of murmur, or frequency (tone) of the murmur<sup>9</sup>. Each of these sub-problems is dealt with using a method or algorithm which is most appropriate to analysing it, e.g. decision trees or neural networks. An arbitration module then processes and combines the partial diagnoses of these specialised sub-systems, to make a final diagnosis. All the above specialised sub-systems and the arbitration module incorporate and are based on expert knowledge. Their combination can lead to an integrated decision support system architecture for heart sound diagnosis, as shown in Figure 2.

## METHODS

### Preprocessing of Heart Sound Signals

The characteristics of the heart sound signal are significantly affected by factors related to the signal acquisition and preprocessing method. Therefore, a heart sound diagnosis algorithm should be tested in heart sound signals from different sources and recorded with different acquisition methods for objective evaluation. For this purpose we collected heart sound signals from nine different heart sound

sources (see Appendix) and created a “global” heart sound database. Because these sources were intended for training purposes they included heart sounds representative of all heart diseases. These heart sounds files had already been diagnosed and linked to a specific heart disease; therefore, they incorporate the knowledge of numerous experts in this area. From the available heart sound signals of this database we chose the ones containing either S4, SP1 or EC. This resulted in a total of 100 heart sound signal files.

Each of these heart sound signal files was initially pre-processed and then converted to the corresponding heart sound feature vector, following a previously described method<sup>2,9</sup>. In particular, the pre-processing comprised of an initial normalisation of each signal in order to account for the amplitude variations among the signals due to different acquisition and recording methods. A set of six processing stages were then performed to identify S1 and S2 and their boundaries.

### Calculation of Feature Vectors

Following the preprocessing tasks, the corresponding feature vector was calculated for each heart sound signal. The selection of the features was based on the technique used by experienced clinicians for analysing heart sounds to make a diagnosis or differential diagnoses. These include features such as the timing of the additional noise in the cardiac cycle (i.e. whether it occurs in diastole or systole) and other characteristics of the sound e.g. its duration, and the frequency of its tone. For these

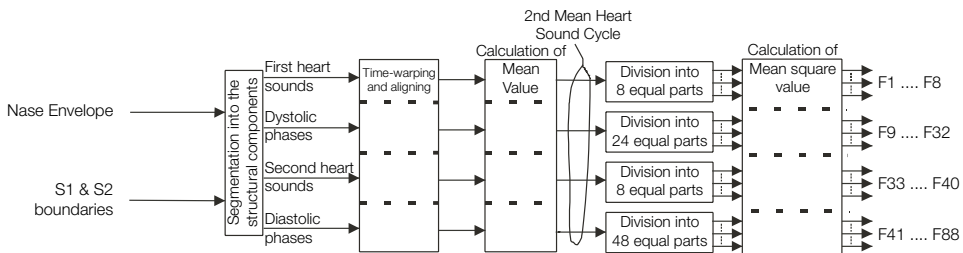


Figure 3. Calculation of the 88 morphological features (F1–F88)

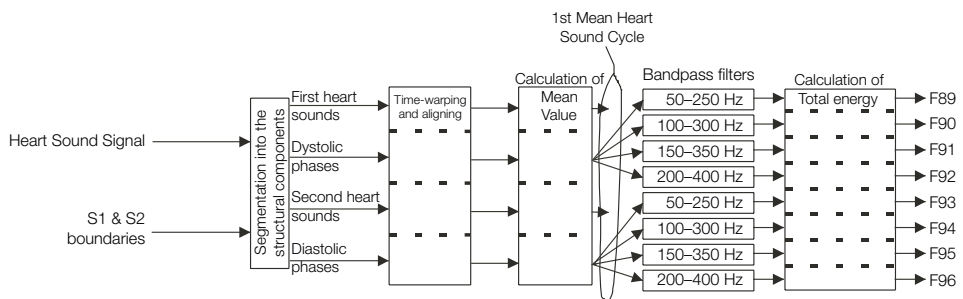


Figure 4. Calculation of the 8 frequency features (F89–F96)

reasons we decided to use a set of time domain morphological features that cover and describe the whole heart cycle, in combination with a set of frequency domain features concerning the energy of the systolic and the diastolic phase in four significant frequency zones. These steps divide the signal into 88 morphological features and 8 frequency features as shown in Figures 3 and 4

Following the above described procedures, every heart sound signal was transformed into a feature vector (pattern) with dimensions of  $1 \times 96$ . The feature vectors of the initial 100 heart sound signals were stored in a database table with 100 records and 98 attributes-fields: one attribute named ID is the pattern identification code, one attribute named S4\_SP1\_EC is the characterisation (diagnosis) of the corresponding heart sound signal as having S4, SP1 or EC, while the remaining 96 attributes are the above 96 heart sound features (F1–F96).

### Analysing the Heart Sounds

Before constructing and utilising the decision tree classifiers, we used relevance analysis<sup>12,13</sup>, to find the most suitable and relevant features for this classification problem. For this purpose, we used the value of the *uncertainty coefficient*<sup>12,13</sup> of each of the above 96 features, which are the independent variables, for ranking them according to their relevance to the classifying attribute (S4\_SP1\_EC), which is the dependent variable. The calculation of the uncertainty coefficient of an independent variable regarding the dependent variable consists of a number of steps which gives a value between 0% and 100%. A low value (near 0%) of the uncertainty coefficient of an independent variable means that if we use this variable for partitioning the initial set of heart sounds there will be only a low increase in homogeneity regarding the dependent variable (and, therefore, low increase in classification rules accuracy), and, therefore, the relevance between this variable and the dependent variable is low. On the contrary a high value (near 100%) of the uncertainty coefficient of an independent variable indicates a high relevance with the dependent variable. In order to examine the relevance and the contribution to the differentiation of S4, SP1 and EC of each of the above mentioned 96 heart sound features, the uncertainty coefficients were calculated for each of them considering the S4\_SP1\_EC field as the classifying attribute – dependent variable.

### Decision Tree Classifiers

A decision tree is a classification tree for classifying new instances (e.g. new heart sound feature vectors) into one of the categories of an important target attribute-dependent variable based on a number of other attributes constituting the independent variables<sup>13–15</sup>. To construct a decision tree we used a training data set of instances, for which we had the values of both the attributes that constitute the independent variables and the targeted attribute that constitutes the dependent attribute. We determined the best test (= attribute + condition) for splitting

the training data set which creates the most homogeneous subsets regarding the dependent variable and, therefore, gives the highest classification accuracy.

We used three different types of decision trees to analyse the heart sounds.

*Detection Decision Tree (DeDT) Architecture.* The detection decision tree (DeDT) architecture treats the problem of differentiation of S4, SP1 and EC as three two-category classification sub-problems: existence of S4 or not, existence of SP1 or not and existence of EC or not. Each of these simpler two-category sub-problems is handled by a separate decision tree, which aims to detect whether the corresponding morphological characteristic exists or not in the examined heart sound.

*Differentiation Decision Tree (DiDT) Architecture.* A differentiation decision tree (DiDT) treats the problem of differentiation of S4, SP1 and EC as a three-category classification problem, i.e. it classifies a feature vector-pattern as having either S4, SP1 or EC.

*Multiple Decision Tree (MuDT) Architecture.* The multiple decision tree architecture combines the DiDT and DeDT architectures to exploit the advantages of both. The suggestions made by these two decision trees are analysed by an arbitration module that makes the final decision on which of these suggestions should be accepted and which of them should be rejected (see Figures 2 and 7).

In order to examine the generalisation capabilities of the constructed decision tree structures, the available feature vectors-patterns set was divided in two subsets. The first subset included 60% of the records of each class of the heart sound patterns set (S4, SP1 and EC classes), which were randomly selected and were used as the training set. The other subset consisted of the remaining patterns (40% of the records of each class) and were used as the test set. In this way the first training test (60%a–40%a) set scheme was formed.

For the second scheme (60%b–40%b) the same proportions (60% training set–40% test set), were kept but random different patterns were selected for the training set. In the same way two more schemes were created (70%a–30%a and 70%b–30%b) with a different proportion (70% training set–30% test set).

For the DiDT architecture, *classification accuracy* was calculated as the ratio of the number of the correctly classified patterns to the total number of patterns of the test data set. For the DeDT architecture, classification accuracy was calculated using the following equation:

$$\text{Accuracy} - X = \frac{\text{correctly classified (X sounds + non X sounds)}}{\text{tested (X sounds + non X sounds)}}$$

where X stands for either S4, SP1 or EC

For each of the three morphological characteristics (S4, SP1, EC) *classification sensitivity* and *classification specificity* were also calculated. The classification

sensitivity for a morphological characteristic is defined as the ratio of the number of the patterns correctly classified as having this morphological characteristic to the total number of patterns having this morphological characteristic of the test data set. Similarly, the classification specificity for a morphological characteristic is defined as the ratio of the patterns correctly classified as not having this morphological characteristic to the total number of patterns not having this morphological characteristic of the test data set.

RESULTS

For each heart sound feature we calculated the uncertainty coefficient separately from the training data set of each of the above four data schemes. Then, based on these four values, we calculated the average value and the standard deviation of the uncertainty coefficients. The average values and the standard deviations of the uncertainty coefficient for the most important features are shown in Figure 5.

The graph demonstrates that the most relevant features of the classifying attribute S4\_SP1\_EC are the frequency features, i.e. high frequency energy and medium frequency energy in the diastolic and systolic phases, and also the morphological features that describe the first heart sound. These results are compatible with our physical understanding of the problem that S4, SP1 and EC click-like sounds appear almost simultaneously with S1. Also each of these click-like sounds is usually related to specific heart diseases that have heart sound murmurs in the systolic and the diastolic phase. The standard deviation values are generally smaller than 7%, showing that the uncertainty coefficients calculated from each scheme separately, especially the ones of the most relevant features, are similar and consistent.

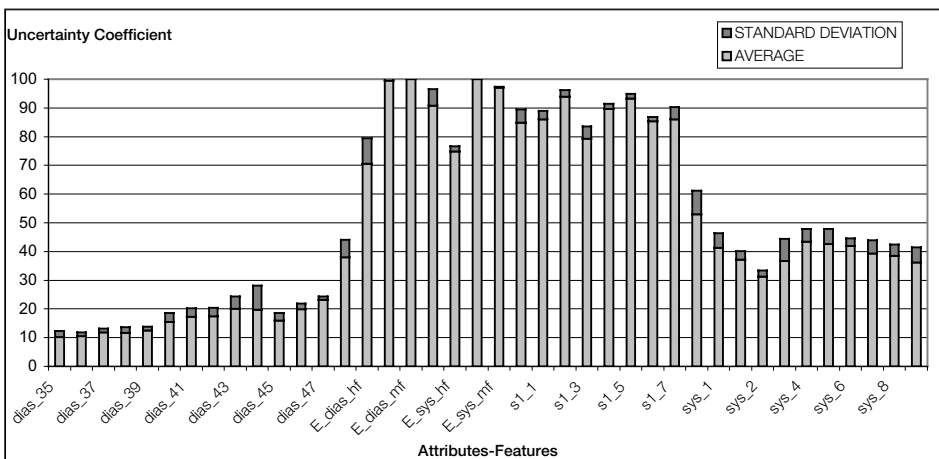
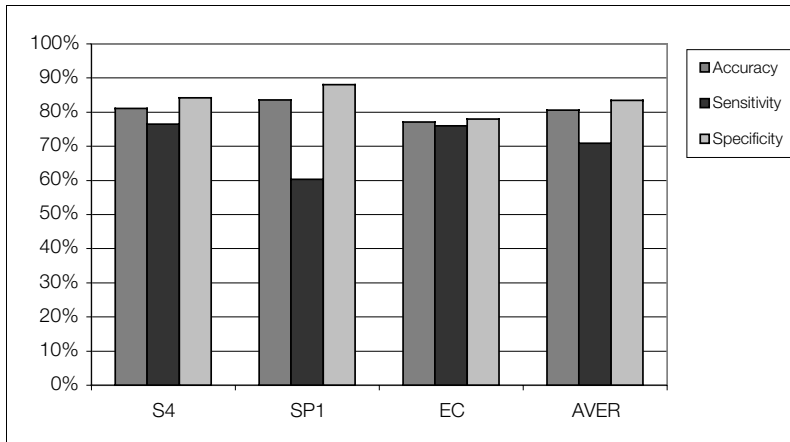


Figure 5. Average values and standard deviations of the uncertainty coefficient for the most important features regarding S4\_SP1\_EC as the classifying attribute





**Figure 6.** Classification accuracy, sensitivity and specificity results for the detection decision tree architecture (DeDT)

**Table 2.** Classification accuracy, sensitivity and specificity results for the differentiation decision tree (DiDT)

Test data set	40%a	40%b	30%a	30%b	Average
Classification accuracy	67.50%	65.00%	70.00%	63.33%	66.46%
Classification sensitivity	65.34%	68.29%	71.54%	59.42%	66.15%
Classification specificity	83.56%	79.83%	83.72%	81.49%	82.15%

Figure 6 gives the results obtained with the DeDT architecture. This shows an average classification accuracy of 80.56%, an average classification sensitivity of 70.93% and average classification specificity of 83.42%. The major drawback with the DeDT architecture is that for many cases it gives more than one suggestion.

Table 2 shows the results for classification accuracy, sensitivity and specificity for each of the four datasets using the differentiation decision tree (DiDT) architecture. The average accuracy and sensitivity was 66% and the specificity 82%. It should be mentioned that all the calculated classification sensitivity and specificity values presented in this paper are based on a lower number of patterns than the corresponding classification accuracy values (i.e. while the classification accuracy is based on all the patterns of the test data set, the classification sensitivity is based only on the ones having the specific morphological characteristic, and the classification specificity is based only on the ones not having the morphological characteristic). It should also be mentioned that the classification accuracy for the training data set was 100% for all the examined cases with both the DiDT and the DeDT.

The multiple decision tree architecture combines the DiDT and DeDT architectures and exploits the advantages of both (Figure 7). Because the DeDT architecture

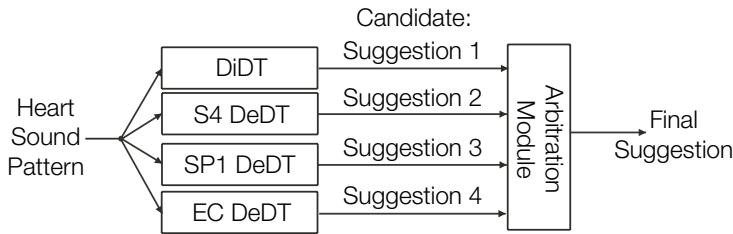


Figure 7. Multiple decision tree architecture

Table 3. Classification accuracy, sensitivity and specificity results for the multiple decision tree architecture

Test data set	40% a	40% b	30% a	30% b	Average
Classification accuracy	72.50%	68.50%	70.00%	66.50%	69.25%
Classification sensitivity	69.90%	69.42%	71.54%	64.33%	68.80%
Classification specificity	81.47%	85.43%	83.72%	83.00%	83.41%

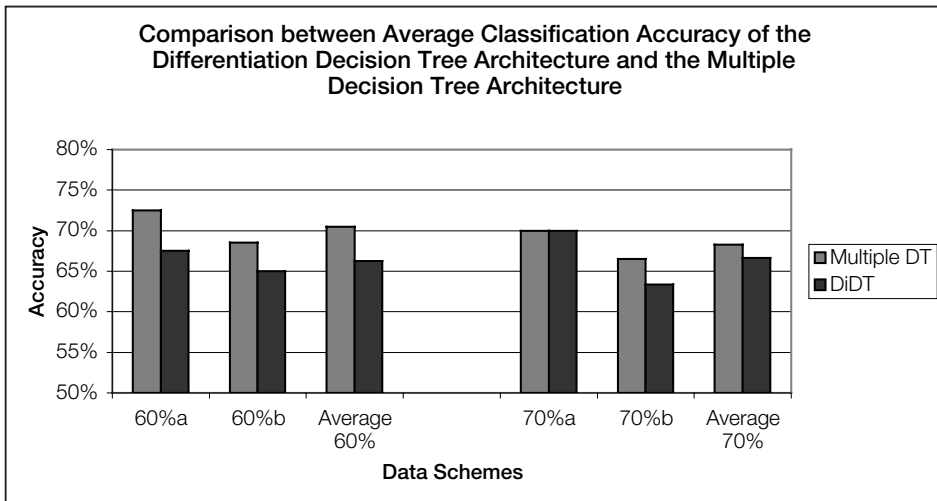


Figure 8. Comparison between DiDT architecture and MuDT architecture

has better classification performance for this problem than the DiDT architecture, the arbitration module can be based on the following rule:

*“If only one DeDT detects its corresponding morphological characteristic, then the final suggestion is the one of this DeDT, otherwise the final suggestion is the one of the DiDT”.*

The results concerning the classification performance achieved using the MuDT architecture, in combination with the above arbitration rule, are shown in Table 3 for all four data schemes, while a comparison with the results of the DiDT architecture are shown in Figure 8. We can see that the MuDT architecture results in improvement of the classification performance in comparison with the DiDT with an increase of 2.79% in classification accuracy, 2.73% in classification sensitivity and 1.26% in classification specificity.

## DISCUSSION

In this paper we investigated the use of decision trees for the differentiation of S4, SP1 and EC, which is a difficult and challenging problem in cardiac auscultation. In this direction several decision tree structures and architectures have been constructed and evaluated as to their classification accuracy, sensitivity and specificity for diagnosing these different heart sounds.

We chose to use decision trees classification algorithms because the knowledge representation model that they produce is compatible with the practices followed by clinicians in making differential diagnoses. Decision trees do not work as a 'black box' but offer a full justification for their suggestions. Using decision trees, clinicians can trace back the model and either accept or reject the proposed suggestion, thus increasing their confidence about the final diagnosis. In contrast, neural networks and algorithms that need a lot of iterations in order to converge on a solution do not offer a justification of their suggestions and are regarded as 'black boxes' by clinicians.

The DiDT architecture provides one final suggestion for each heart sound pattern, but its classification performance is lower in comparison with the DeDT architecture. The DeDT architecture has better classification performance, but provides three suggestions for each heart sound pattern; if these three suggestions are not consistent, the result can be confusing and probably less useful to the clinician. The multiple decision tree (MuDT) architecture achieves higher classification performance than the DiDT and also provides a single suggestion.

This work has demonstrated that decision tree algorithms can be used as a basis for decision support systems to assist inexperienced clinicians with heart sound diagnosis. Decision trees can be very useful knowledge management tools in this area. They codify and effectively incorporate the knowledge of numerous highly specialised and experienced doctors, making them a valuable and useful tool for the exploitation and dissemination of knowledge. Such computer-based support can play a role in improving the quality and effectiveness of primary care, particularly in small and remote areas. In these places it may help reduce unnecessary patient travel for specialist consultations and investigations.

Different decision tree structures and architectures were constructed and tested on various training and test data sets. Their performance on the training data sets

was 100% successful, while their performance on the test data sets, which is an indicator of their generalisation capabilities, was satisfactory. For this difficult and complicated differentiation problem our decision tree structures achieved classification accuracy and sensitivity levels of almost 70% and classification specificity levels greater than 80%. These results are encouraging, taking into account the limited amount of data available for this study, and also the existing possibilities for classification performance improvements.

Relevance analysis can be used to determine a small critical subset of the initial set of features that contains most of the information required for heart sound diagnosis.

Further improvement in the classification performance of the examined decision tree structures and architectures is necessary. We believe that this is possible by:

- a) Using more heart sound signals with these morphological characteristics to give us larger training and test data sets.
- b) Developing more sophisticated MuDT architectures. For example, we can improve the architecture shown in Figure 7 by adding three two-categories DeDTs in the first (suggestion) stage. Each of these will be trained and become specialised (and, therefore, more efficient than the three-categories DiDT) in differentiating between two of the three targeted morphological characteristics (i.e. one DiDT for differentiating between S4 and SP1, one for differentiating between SP1 and EC, and one for differentiating between S4 and EC). If two DeDTs detect the corresponding morphological characteristics, then the arbitration module will use the output of the corresponding two-categories DiDT as the final suggestion; if all three DeDTs detect the corresponding morphological characteristics, then the arbitration module will use the output of the three-categories DiDT as the final suggestion.

Along these directions, further research is already in progress. Additional research is being conducted concerning the use of neural network architectures for this differentiation problem, and the comparison of their classification performances with those of the decision trees architectures. Initial results from such studies have shown that neural network architectures can provide small improvements in classification performance (2–3% increase in classification accuracy). This improvement in performance compensates for the disadvantage of being a ‘black box’ that does not provide justification for its suggestions.

Further research is required for the development of a systematic methodology for designing arbitration rules. In addition the design of an appropriate MuDT architecture for achieving the best classification performance for a specific problem, which could possibly include as ‘nodes’ not only decision trees but also other types of classifiers as well, is an open research question. Finally, the proposed decision trees structures and architectures can be applied to other heart sound diagnosis (or medical diagnosis in general) problems and should be further evaluated and improved.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr D. E. Skarpalezos for his clinical support, Dr G. Koundourakis, and Neurosoft S.A. for their support and provision of Envisioner, a data-mining tool that was used to execute algorithms related to the decision trees.

## REFERENCES

- 1 Criley SR, Criley DG, Criley JM. Beyond heart sound: an interactive teaching and skills testing program for cardiac examination. Blaufuss Medical Multimedia, San Francisco, CA, USA, *Computers in Cardiology* 2000; 27; 591–94.
- 2 Stasis AC, Loukis EN, Pavlopoulos SA, Koutsouris D. Using decision tree algorithms as a basis for a heart sound diagnosis decision support system. *Proc of the 4th Annual IEEE Conf on Information Technology Application in Biomedicine*. UK, 2003, pp. 354–57.
- 3 Leung T, White P, Collis W, Brown E, Salmon A. Analysing paediatric heart murmurs with discriminant analysis. *Proceedings of the 19th Annual conference of the IEEE Engineering in Medicine and Biology Society*. Hong Kong, 1998, pp. 1628–31.
- 4 Durand L, Guo Z, Sabbah H, Stein P. Comparison of spectral techniques for computer-assisted classification of spectra of heart sounds in patients with porcine bioprosthetic valves. *Med Biol Eng Comput* 1993; 31: 229–36.
- 5 Wu CH. On the analysis and classification of heart sounds based on segmental Bayesian networks and time analysis. *Journal of the Chinese Institute of Electrical Engineering, Transactions of the Chinese Institute of Engineers* 1997; Series E, 4: 343–50.
- 6 DeGross CG, Bhatikar S, Hertzberg J, Shandas R, Valdes-Cruz L, Mahajan RL. Artificial neural network-based method of screening heart murmurs in children *Circulation* 2001; 103; 2711–16.
- 7 Leung T, White P, Collis W, Brown E, Salmon A. Classification of heart sounds using time-frequency method and artificial neural networks. Engineering in Medicine and Biology Society. *Proceedings of the 22nd Annual International Conference of the IEEE*, 2000, Vol. 2, pp. 988–91.
- 8 Sharif Z, Daliman S, Sha'ameri AZ, Salleh SHS. An expert system approach for classification of heart sounds and murmurs. *Signal Processing and its Application, 6th International Symposium*. IEEE, 2001, vol. 2, pp. 739–40.
- 9 Stasis A. Decision support system for heart sound diagnosis, using digital signal processing algorithms and data mining techniques. Phd Thesis; National Technical University of Athens; 2003.
- 10 Epstein EJ. *Cardiac Auscultation*. Oxford: Butterworth–Heinemann, 1991.
- 11 Baracca E, Scorzoni D, Brunazzi MC, et al. Genesis and acoustic quality of the physiological fourth heart sound. *Acta Cardiologica* 1995; 50: 23–28.
- 12 Kamber M, Winstone L, Gong W, Cheng S, Han J. *Generalisation and decision tree induction: efficient classification in data mining. Proc. of 1997 Int'l Workshop on Research Issues on Data Engineering (RIDE'97)*. Birmingham, England, 1997, pp. 111–20.
- 13 Koundourakis G. EnVisioner: a data mining framework based on decision trees, October 2001, Doctoral Thesis, University of Manchester Institute of Science and Technology.
- 14 Mitchell T. *Machine Learning*. Mc-Graw Hill Companies Inc., 1997.
- 15 Han J, Kamber M. *Data Mining: Concepts and Techniques*. Morgan Kaufman Publisher, 2001.

APPENDIX: RECORDED HEART SOUND SOURCES

- A1 Karatzas N, Papadogiann D, Spanidou E, Klouva F, Sirigou A, Papavasiliou S. *Twelve Recorded Lessons with heart Sound Simulator*. Merk Sharp & Dohme Hellas, Medical Publications Litsas, Athens 1974.
- A2 Criley JM, Criley DG, Zalace C. *The Physiological Origins of Heart Sounds and Murmurs*. Harbor UCLA Medical Center, Blaufuss Medical Multimedia, 1995.
- A3 Littman. *20 Examples of Cardiac & Pulmonary Auscultation*.
- A4 Cable C. *The Auscultation Assistant*, 1997. <http://www.wilkes.med.ucla.edu/intro.html>.
- A5 Glass L, Pennycook B. *Virtual Stethoscope*. McGill University, Molson Medical Informatics Project 1997. <http://www.music.mcgill.ca/auscultation/auscultation.html>.
- A6 VMRCVM Class of 2002. Virginia Maryland regional College of veterinary medicine. <http://students.vetmed.vt.edu/2002/cardio/heartsounds.html>.
- A7 Kocabasoglu YE, Henning RH. *Human Heart Sounds*. [http://www.lf2.cuni.cz/Projekty/interna/heart\\_sounds/h12/index.html](http://www.lf2.cuni.cz/Projekty/interna/heart_sounds/h12/index.html).
- A8 Frontiers in Bioscience. *Normal and Abnormal EKGs and Heart Sounds*. <http://www.bioscience.org/atlas/heart/sound/sound.htm>.
- A9 Student Internal Medicine Society. *Class 2000 Cardiology Heart Sounds Wave Files*. American College of Physicians, Baylor College of Medicine. <http://www.bcm.tmc.edu/class2000/sims/HeartSounds.html>.

