A Decision Tree-Based Method, Using Auscultation Findings, for the Differential Diagnosis of Aortic Stenosis from Mitral Regurgitation

AC Stasis[1], EN Loukis[2], SA Pavlopoulos[1], D Koutsouris[1]

[1]Biomedical Engineering Laboratory, National Technical University of Athens, Greece,
[2] Dept. of Information and Communication Systems Engineering, University of Aegean, Samos, Greece

Abstract

In this study, Decision Tree algorithms are used with promising results in a crucial and at the same time complicated classification problem concerning differential diagnosis of heart sounds.

Decision Tree structures are constructed, using data mining/distillation methods and then are used to classify heart sounds that were recorded from patients that have either Aortic Stenosis (AS) or Mitral Regurgitation (MR).

Emphasis is given on the selection of the appropriate features that are adequately independent from the heart sound signal acquisition method.

The differentiation capabilities and the classification performance of the fully expanded Decision Tree classifiers and the pruned Decision tree classifiers are studied for this problem. For each constructed Decision Tree classifier the partial classification accuracies for the AS and MR auscultation findings are also estimated.

1.   Introduction

The heart sound diagnosis problem consists in the diagnosis from heart sound signals a) whether the heart is healthy, or not and b) if it is not healthy, which is the exact heart disease. Although heart sounds can provide low cost screening for pathologic conditions, the internal medicine and the cardiology training programs underestimate the value of cardiac auscultation and the new clinicians are not well trained in this field [1]. Therefore efficient Decision Support Systems would be very useful for supporting clinicians to make better heart sound diagnosis, especially in rural areas, in homecare in primary healthcare and generally wherever operationally complex and high cost medical equipment (like CT, MRI, Echocardiography) is not available [2].

In Aortic Stenosis (AS) the aortic valve is thickened and narrowed. As a result, it does not open fully during cardiac contraction in systolic phase, leading to abnormally high pressure in the left ventricle and producing a systolic murmur that has relatively uniform frequency and rhomboid shape in magnitude [3].

In Mitral Regurgitation (MR) the mitral valve does not close completely during systole due to tissue lesion and there is blood leakage back from the left ventricle to the left atrium. MR is producing a systolic murmur that has relatively uniform frequency and magnitude slope [3]. The spectral content of the MR systolic murmur has faintly higher frequencies than the spectral content of the AS systolic murmur. The closure of aortic valve affects the second heart sound and the closure of the mitral valve affects the first heart sound [4].

It is obvious from the above description that the AS systolic murmur has very similar characteristics with the MR systolic murmur and therefore the differentiation between these two diseases is a difficult problem in heart sound diagnosis, especially for the young inexperienced clinicians. The correct discrimination between them is of critical importance for the determination of the correct treatment that has to be recommended.

2.   Methodology

In this paper a rule-based method, based on Decision Trees, is developed for the problem of differential diagnosis between the AS and the MR using auscultation findings. The problem of differentiation between AS and MR has been investigated in [3] and [5]. The method proposed in [3] was based on the different statistic values that exist in the spectrogram of the systolic murmurs that these two diseases produce. The method proposed in [5] was based on frequency spectrum analysis and a filter bank envelope analysis of the first and the second heart sound. Our work aims to investigate whether Decision Tree-based classifier algorithms can be a trustworthy alternative for this heart sound diagnosis problem.

2.1.   Preprocessing of heart sound data

Heart sound diagnosis is quite complicated, because it depends not only on the heart sound signal but also on other physical examination variables. Furthermore, the quality of the heart sound signal has great variability, affected by mans factors associated with the acquisition method. Therefore a heart sound diagnosis algorithm should be tested in heart sound signals from different

sources and recorded with different acquisition methods. For this purpose we tried to collect heart sound signals from different heart sound sources (the same heart sound sources as the ones described in [2]) and create a "global" heart sound database. The heart sound signals were collected from educational audiocassettes, audio CDs and CD ROMs; they had already been diagnosed and related to a specific heart disease. We chose 41 heart sound signals with AS systolic murmur and 43 ones with MR systolic murmur.

This set of 41+43=84 heart sound signals was converted into a set of heart sound feature vectors using the pre-processing method described in [2]. Three categories of features were extracted:

1. The statistical features: these are the standard deviations of the durations, the S1 and the S2 peak values of all the heart cycles, the heart sound signal includes, and the mean heart rate. These are the first four scalar features (F1-F4) of the feature vector of the heart sound signal.

2. The morphological features: Each component of the characteristic heart sound signal envelope was divided into equal parts and then the mean square value of each part was calculated and this value was used as a feature in the corresponding heart sound vector. In this way we calculated 8 scalar features for S1 (F5-F12), 24 scalar features for the systolic period (F13-F36), 8 scalar features for S2 (F37-F44) and 48 scalar features for the diastolic period (F45-F92).

3. The frequency features: The systolic and diastolic phase components of the characteristic heart sound signal were also passed from four bandpass filters (50-250Hz, 100-300Hz, 150-350Hz, 200-400Hz). For each of these 8 outputs, the total energy was calculated and was used as a feature in the heart sound vector (F93-F100).

These pre-processed data feature vectors were stored in a database table. This table had 84 records, i.e. as many records as the available heart sound signals; each record describes a heart sound feature vector and has 102 fields. Each field corresponds to one feature of the feature vector or in other words to one attribute of the heart sound. One field named ID (used as the primary key of this database table), one attribute named hdisease for the characterization of the specific heart sound signal as MR or AS and 100 attributes for the above 100 heart sound features (F1-F100).

2.2. The decision tree-based method

Before constructing the Decision Tree Classifiers we tried, using Relevance Analysis ([6]), to improve the classification efficiency by removing the less useful (for the classification) features and reducing the amount of input data to the classification stage. The value of the Uncertainty Coefficient ([6], [7]) of each of the above 100 features was used to rank these features according to

their relevance to the classifying attribute, which in our case is the hdisease attribute.

In order to construct the Decision Trees we used a training data set and the algorithms described in [8]. Starting from the root node we determined the best test (= attribute + condition) for splitting the training data set, which created the most homogeneous subsets concerning the classifying attribute. Each of these subsets can be further split in the same way etc. In order to determine the best split test in a node we evaluated all possible split tests for each attribute using the Entropy Index ([7]) as split index. We finally selected the split test with the lowest value for Entropy Index to split the node.

The further split of the subsets into smaller subsets can continue until we have subsets-nodes with records having all the same value of the classification attribute. This kind of expansion usually over-fits the decision tree to the training data set and to the noise this data set contains and therefore this decision tree usually does not have good generalization capabilities. Pruning therefore is adopted in order to prevent such an over-fitting ([9]).

In our work for the pruned Decision Tree structure a node was not split if it had less records than a percentage of the initial training data set records (least node support).

2.3. Selection of training and test pattern sets

In order to create the first training-test scheme the complete pattern set was divided in two subsets. The first subset included 50% of each class (AS and MR) of the heart sound patterns set, randomly selected. This subset was the training set. The other subset included the remaining patterns (50% of each class) was used as the test set. The division of the pattern set was repeated, keeping the same proportions (50% training set - 50% test set), but using different patterns giving a second scheme. In the same way were created two more schemes with different proportion (60% training set-40% test set).

3. Results.

3.1. Relevance analysis

In order to search the relevance and the contribution to the differential diagnosis of AS from MR for each of the above-mentioned 100 heart sound features, the uncertainty coefficients were calculated for each one of them considering the hdisease field as the classifying attribute. The calculation was made separately for the training data set of each of the 4 data schemes and for each heart sound feature. Then we calculated the average value and the standard deviation of the Uncertainty Coefficients taking into account the 4 values that were calculated from these 4 schemes (Figure 1).
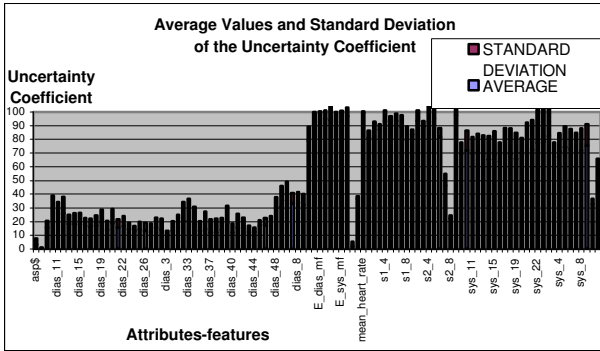
Figure 1: Average values and Standard Deviations of the Uncertainty Coefficient for the most important features regarding the hdisease as classifying attribute.

We remark that the most relevant features are the frequency features (e.g. E_dias_hf, E_sys_lf, etc), the morphological features that describe the S1 (s1_1…s1_8), the S2 (s2_1…s2_8) and the systolic murmur (sys1, … sys24). These results are compatible with our physical understanding of the problem; the AS and MR systolic murmurs have different frequency content and different envelope shape. On the contrary in the diastolic phase there is no murmur in both diseases, therefore the diastolic phase of heart sound signals cannot contribute to the differentiation between AS and MR. Additionally the closure of Mitral valve affects the S1 and the closure of Aortic valve affects the S2.

The standard deviation values are generally small, showing that the Uncertainty Coefficients calculated from each scheme separately, especially the ones of the most relevant features, are similar and consistent.

3.2.   Decision trees

Initially we constructed the Fully Expanded Decision Tree structure (without pruning). The training data set we used was from the first scheme 50% and had 42 heart sound patterns. The rest 42 patterns were used as a test set. The decision tree was constructed based on the training data set and afterwards using this decision tree the patterns of the test data set were classified in order to investigate the generalization capabilities of the decision tree. Then the percentage of the correctly classified patterns of the test data set (according to the hdisease attribute) was calculated. The classification accuracy for the first 50% scheme was 95.24% (=40/42*100).

In order to exploit the results of the Relevance Analysis, we constructed one more Decision Tree using only the heart sound features that have Uncertainty Coefficient above 90% (i.e. 16 out of 100 features). This Decision Tree was also used for classifying the patterns of the same test data set, with results identical to the ones described above. Therefore using only 16 features we can get identical levels of classification accuracy with much less computational effort. For each of the available data

schemes we repeated the same procedure and calculated the corresponding classification accuracy (Figure 2).

We remark from Figure 2 that the classification accuracy results for the schemes with the largest training data sets (=60%a and 60%b) have a consistent (=the same for both schemes) level of classification accuracy about 92.56%. For the 50%a and 50%b schemes, the classification accuracies have differences i.e. 75% and 95% respectively. This is probably due to the small size of the corresponding training sets. Therefore it is concluded that decision trees if and only if properly trained can give a high and consistent level of classification accuracy concerning the differentiation of AS from MR, using auscultation findings. The classification accuracy for the training data set was 100% for all the examined cases.

In order to investigate the performance of the Decision Tree classifier in more detail we calculated, in addition to the above total accuracy, also the partial accuracy for AS heart sounds, and the partial accuracy for MR heart sounds (Figure 2). The conclusion we derive from Figure 2 concerning the partial classification accuracies is that the MR classification accuracy is significantly higher than the AS classification accuracy; therefore generally the classification performance of the decision trees for the MR heart sounds is higher than it is for AS heart sounds. Again the classification total accuracy, AS_Accuracy and MR_Accuracy for the training data set was 100% for all the examined cases.
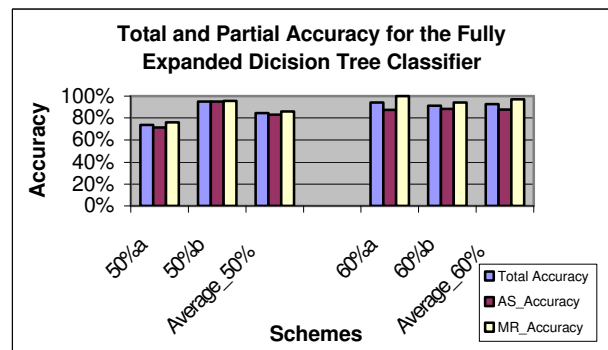

Figure 2: Total accuracy, partial AS and MR accuracy for the Fully Expanded Decision Tree Classifier classification.

In the last part of our work we tried to investigate if we could improve the generalization capabilities of the Decision Trees with pruning. The least node supports that were tried were: 5%, 10% and 15%. In Figure 3 are shown the total classification accuracies achieved with pruned decision trees compared with the previous results from the fully expanded decision tree (0% minimum support at leaf nodes). It is concluded that for our specific differentiation problem pruning does not significantly change the classification accuracy of the decision trees.
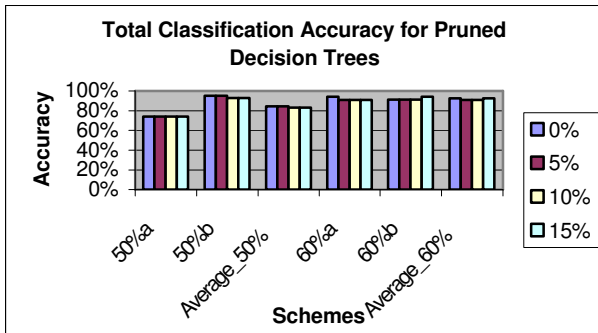
Figure 3: Total Classification Accuracy Results for all data schemes for Pruned Decision Trees.

Finally using pruned decision trees the classification accuracy for the training data sets was not 100% for all schemes; for some schemes reduced at the level of 96%.

4. Conclusions

In this paper has been investigated the applicability and the suitability of a number of decision tree structures for the differential diagnosis of AS from MR. As criterion for this evaluation was considered the classification accuracy (both the total one and the partial ones) for the training set and the test set. The main conclusions from this work are the following:

- The Decision Trees can be used with high levels of success for the differentiation of AS from MR. The generalization capabilities of Decision Trees were found to be satisfactory. However having more training data the decision trees can produce more trustworthy predictions i.e. increasing the size of the training data sets improves the classification accuracy and the general reliability of the system.

- There is a small subset of the initial features that contains most of the information required for the differentiation and this set can be determined using Relevance Analysis.

- The Fully Expanded Decision Tree structures have similar levels of generalization and classification accuracy for new data in comparison with the Pruned Decision Tree structures.

-The decision tree model is simple and the clinicians are familiar with it, because they use a similar model when they make differential diagnosis.

The general heart sound diagnosis problem can be divided into a number of simpler problems. All these simpler problems can be solved by separate specialized Decision Support Systems, which can be based on different methods, algorithms and features. The partial diagnosis given by these Decision Support Systems can then be combined to give a final diagnosis ([2]). The combination of all these Decision Support Systems can lead to an integrated Decision Support System architecture for Heart Sound Diagnosis.

Further research is required for the development of a methodology for the selection of the most appropriate Decision Tree structure and for improving the classification accuracy. Also further research is required in order to investigate the applicability and suitability of Decision Tree-based methods for other significant problems in the area of heart sound diagnosis (e.g. [10]).

Acknowledgements

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;vol 27;pp 591-594.

[2] Stasis A, Loukis E, Pavlopoulos S, 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 354-357.

[3] Myint WW, Dillard B. An Electronic Stethoscope with Diagnosis Capability. Proc of the 33rd IEEE South-eastern Symposium on System Theory; 2001.

[4] 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.

[5] Hebden JE, Torry JN. Identification of Aortic Stenosis and Mitral Regurgitation by Heart Sound Analysis. Computers in Cardiology; 1997; Vol24; pp 109-112.

[6] Kamber M, Winstone L, Gong W, Cheng S, and 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, UK; April 1997; pp 111-120.

[7] Koundourakis G, EnVisioner: A Data Mining Framework Based On Decision Trees, Phd Thesis; University of Manchester Institute of Science and Technology; October 2001.

[8] J. Han, M. Kamber, 'Data Mining: Concepts and Techniques', Morgan Kaufman Publisher 2001

[9] S.M. Weiss, N. Indurkhya, "Small sample decision tree pruning", Proceedings of the 11th International Conference on Machine Learning, Morgan Kaufmann, pp 335-342, 1994.

[10] A. Stasis, D. Skarpalezos, S. Pavlopoulos, D. Koutsouris, Differentiation of opening snap, second heart sound split and third heart sound, using a multiple Decision Tree Architecture, Computational Management Science Conference Crete, Journal of the CMS, May 2003

Address for correspondence.
Name: Antonis Ch. Stasis
Full postal address: Riminiton 42 Voula, Athens, Greece, 16673
E-mail address: astasis@biomed.ntua.gr