IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 1, Issue 1, March, 2013
ISSN: 2320 - 8791
www.ijreat.org

Support Vector Machine For Heart Failure Detection And Severityassessment

N.Nathiya[1] S.Kuttala Kumar[2] and V.Jothimurugan[3]

[1]Assistant Professor,Computer Science and Engineering, V.S.B.Engineering College, Karur, Tamil nadu.

[2,3]Computer Science and Engineering V.S.B.Engineering College, Karur, Tamil nadu.

Abstract

Heart failure (HF) is one of the major disorder in which the damage to the heart causes weakening of cardiovascular system. Congestive heart failure (CHF) is generally the inability of the heart to supply sufficient blood flow to meet the needs of the body. Data mining has become a fundamental methodology for computing applications in medical informatics. Various algorithms associated with data mining have significantly helped to understand medical data more clearly, by distinguishing pathological data from normal data, for supporting decision-making as well as visualization and identification of hidden complex relationships between diagnostic features of different patient groups. Heart failure detection via classification and regression tree method gives lower sensitivity value. The proposed methodology uses Support Vector Machine (SVM), a classification method to improve the sensitivity of HF prediction.SVM is a new generation learning system based on recent advances in statistical learning theory.

Keywords - Data Mining, Heart Failure (HF), Support Vector Machine (SVM), Congestive heart failure (CHF).

1. INTRODUCTION

Heart failure (HF) means weakening of the cardiovascular system. Congestive heart failure (CHF) is generally the inability of the heart to supply sufficient blood flow to meet the needs of the body. Heart failure is a common, costly, disabling, and potentially deadly condition. In developed countries, around 2% of adults suffer from heart failure, but in those over the age of 65, this increases to 6–10% Heart failure symptoms are traditionally and somewhat arbitrarily divided into "left" and "right" sided, recognizing that the left and right ventricles of the heart supply different portions of the circulation. Heart failure can cause a number of symptoms including shortness of breath, leg swelling, and exercise intolerance. The condition is diagnosed with echocardiography and blood tests. There are several data mining methodology available for the diagnosis of HF stages.

Due to rapidly growing aging population, the increased burden of chronic diseases, and the increasing healthcare costs, there is an urgent need for the development, implementation, and deployment, in everyday medical practice, of new models of healthcare services. In this scenario, ICT, and especially home monitoring (HM) [13] and data mining (DM) [14], play an important role. DM is the computer-assisted process of digging through and analyzing a large quantity of data in order to extract meaningful knowledge and to identify phenomena faster and better than human experts. As regards HM, although a wide literature

describes technical solutions, the evidence of ICT cost-effectiveness is limited [15] and only a few studies compare

HM with other models of disease management programs (DMPs) [16]. DMPs are more cost-effective than ambulatory follow-up, which is the gold standard [8], without using costly technologies, which are not familiar to the elderly.

Also, HM is reported to be more effective [9] than follow-up. Nonetheless, HM is equally effective as DMPs, but less efficient because it is about five times more costly than DMPs and about 20 times more costly than ambulatory follow-up [10]. This led us to search for new models of HM, which incorporate further intelligent and automatic systems/services to exceed DMPs in effectiveness, offering advanced functionalities for early detection of any worsening in patient's condition, which could otherwise require more complex and expensive care. Among cardiovascular pathologies, heart failure (HF) is one of the most studied both for HM and for DM, perhaps because it has a considerable impact on healthcare costs, being chronic, degenerative, age related [12], and a leading cause of the elderly hospitalization [13].

One of the most promising methods to study HF is the heart rate variability (HRV), a noninvasive measure, which reflects the variation over time of the interval between consecutive heartbeats. The proposed methodology uses Support Vector Machine (SVM), a classification method to improve the sensitivity of HF prediction.SVM is a new generation learning system based on recent advances in statistical learning theory.

2. RELATED WORK

M. H. Asyali et.al [1] have proposed the idea of Heart Rate Variability (HRV)by time- or frequency-domain methods. The time-domain HRV measures are based on beat-to-beat intervals whereas frequency-domain analysis expresses HRV in terms of its constituent frequency components. HRV analysis has emerged as a diagnostic tool that quantifies the functioning of the autonomic regulation of the heart and heart's ability to respond. The long-term HRV measures indicate the cardiac condition with higher sensitivity and specificity.

The SDNN, standard deviation of all normal-to-normal beat intervals, has the highest class discrimination power, has consequently designed a Bayesian classifier that produced an optimal threshold for SDNN. The observed sensitivity (true positive) and specificity (true negative) rates of 81.82% and 98.08% respectively. The results obtained were not good with accuracy and sensitivity values.

J. T. Bigger, J. L. Fleiss, R. C. Steinman et.al [2] have proposed the normal values of RR variability for middle-aged persons and compare them with values found in patients early and late after myocardial infarction. The presence or absence of heart disease, age, and sex(in this order of importance) are all correlated with RR variability. The spectral analyses on continuous 24-hour ECG recordings to quantify the total power, high and low frequency. Thus, measures of RR variability are used to screen groups of middle-aged persons to identify individuals who have substantial risk of coronary deaths or arrhythmic events, misclassification of healthy middle-aged persons should be rare.

L.Breiman, J.H.Friedman, C.J.Stone et.al [4] have proposed the idea on prediction tree varieties, regression trees and classification trees.The CART decision tree is a binary recursive partitioning procedure capable of processing continuous and nominal attributes both as targets and predictors. Data are handled in their raw form; no binning is required or recommended. Trees are grown to a maximal size without the use of a stopping rule and then pruned back (essentially split by split) to the root via cost-complexity pruning. The CART mechanism is intended to produce not one, but a sequence of nested pruned trees, all of which are candidate optimal trees. The “right sized” or “honest” tree is identified by evaluating the predictive performance of every tree in the pruning sequence.

L. Pecchia, P. Melillo,M. Sansone, and M. Bracale et.al [6] proposed the idea of Heart Rate Variability (HRV) as aquantitative marker of autonomic nervous system activity and as a powerful method of analysis to diagnose andprevent critical events. The results suggest that differences exist between HRV features of normal subjects and a patient suffering from CHF.These differences seems to be related to the severity of the pathology. HRV reflects the actions of sympathetic and parasympathetic branches of the autonomic nervous system on the regulation of the sinus node, which is the natural pacemaker of the heart. They have concluded with the problem in signal processing and extractions based on neural networks were not good.

3. PROBLEM DEFINITION

The problem of sensitivity is solved by the classification algorithm (Support Vector Machine) .Sensitivity deals with the percentage of sick people who are identified as having the condition. Support Vector Machine is a state of art methods which classifies the dataset, train and tests them in hyper plane to meet the sensitivity problem effectively.

4. DATA MINING

This section illustrates the preprocessing of ECG signals using standard algorithm, HRV features extraction depends on toolkit and patient classification based on HF detection and HF severity assessment.

A. Preprocessing

The ECGs were processed following the international guidelines on HRV analyses [8]. After filtering, QRS complexes are detected using a standard algorithm [28]. Although this algorithm could be improved in future, we are first interested in comparing our results with those obtained using other available tools, during clinical trial.

B. HRV Features Extraction

We performed standard short-term HRV analysis, according to international guidelines [8]. We developed the web services using the algorithms and the code of PhysioNet’s HRV Toolkit [11], since it is rigorously validated and because the tool will be used as a valuable benchmark during the clinical trial. This toolkit enables calculation of basic time- and frequency-domain HRV features widely used in the literature.

C. Patient Classification

The platform supported a strategy of automatic classifications consisting of two steps: “HF detection” and “HF severity assessment.” The former, discussed in detail elsewhere [9], was used in the platform to prescreen patients before they underwent the latter. The whole classification aimed to early detection of any worsening, assuming that during worsening, patients will gradually show characteristic of a more severe HF. Both classifiers were based on the CART methods. We pruned the trees according to a tradeoff is classification probability and tree complexity, defined as its number of nodes. This reduced the risk of over fitting as further detailed in [10]. The performances of both classifiers were assessed using a cross-validation technique [12].

5. HEART FAILURE DETECTION AND ASSESSMENT

Heart Rate Variability is measured by time- or frequency-domain methods. The time-domain HRV measures are based on beat-to-beat intervals where as frequency-domain analysis expresses HRV in terms of its constituent frequency components. HRV analysis has emerged as a diagnostic tool that quantifies the functioning of the autonomic regulation of the heart and heart’s ability to respond. RR variability measure provides spectral analyses on the ECG recordings to quantify the features. Decision tree method is capable of processing continuous and nominal attributes. The short terms HRV are assessed by Bayesian classifier. The common problems with existing methodologies are its accuracy, specificity values with high percent but with lower sensitivity.

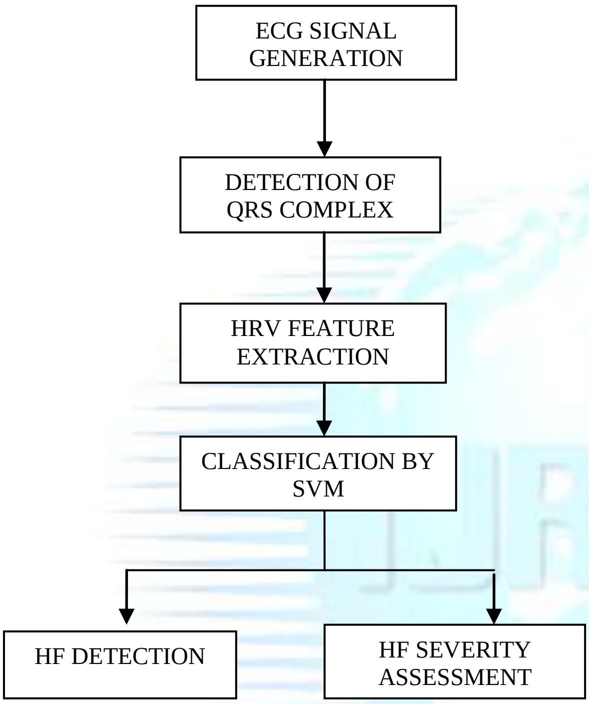

Fig 5.1 Overview of system architecture

The ECG signals are simulated from the system is preprocessed to extract the QRS complexes. The long term heart rate variability features such as total power (TP), ratio between low and high frequencies (LF/HF), root mean square value (RMS), and energy of the signal (E), standard deviation (SD) for normal-normal intervals. Classification algorithm classifies features based on SVM in multidimensional plane for prognosis of HF patients. HF severity can be assessed by the same methodology.

The advantages of using SVM are,

- Sensitivity analysis can be improved
-

recision and specificity values are expected to be good

SVM offers number of advantages over other types of multivariate classifiers.SVM have gained wide acceptance due its solid theoretical basis and high generalization ability. The two key features of SVM are generalization theory and kernel functions. The support vector machine has been developed as a robust tool for classification and regression in noisy and complex domains.

6. DATA SOURCE CALCULATION

Data source has been computed manually because of the lack of availability HRV toolkit. The dataset contains nearly 200 tuples with six attributes. Those attributes are Standard

Deviation, RMS value, low and high frequencies of the signal and its ratio, power and energy of the signal.

Root Mean Square value can be calculated as,

$$RMS\ (f(t))=\sqrt{1/T\int_{t-T}^{t} f(t)^2} \qquad (4.1)$$

Where f (t) =input signal

T =1/fundamental frequency

Measure	Description	Unit
SD	Standard Deviation	ms
RMS	Square root of the Mean of the Sum	ms
FR	Frequency of the Signal	ms^2
E	Total Energy	ms^2
P	Total Spectral Power	ms^2

Table 1. Selected HRV Features

7. PERFORMANCE EVALUATION

Below screen illustrates training and testing graph of SVM simulated by MATLAB tool. The accuracy rate achieved is 98% which is effective than the accuracy achieved by previous methods.

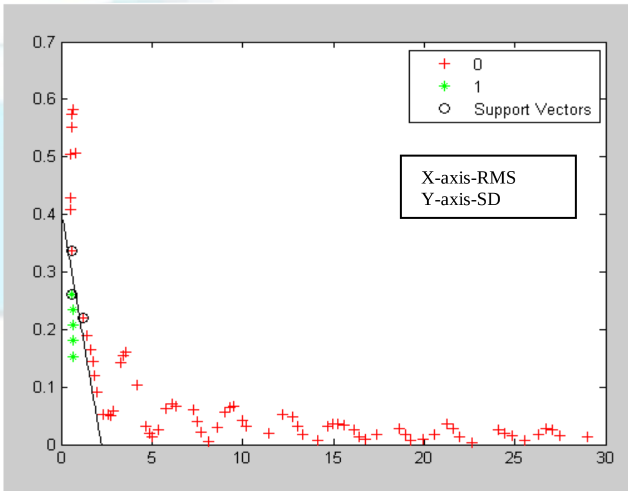

Fig 6.1. Trained graph of SVM

In fig 5.1 shows the proposed approach, support vector machine is used to find out heart failure detection and severity assessment between the root mean square value and standard deviation are calculated from ECG signals. The trained performance of SVM is shown in fig 5.1; it yields some sensible and good accuracy for signals which are simulated in the assessment.

The samples are taken in the form of signals for testing the machine with the heart rate variability features such as root mean square value (RMS), standard deviation (SD). In fig 5.2 shows the testing the signals such as 0(training), 0(classified), 1(training), 1(classified), which is simulated from HRV feature extraction method to determine QRS complexes. The QRS complexes are detected and it is again classified by Support Vector Machine (SVM).

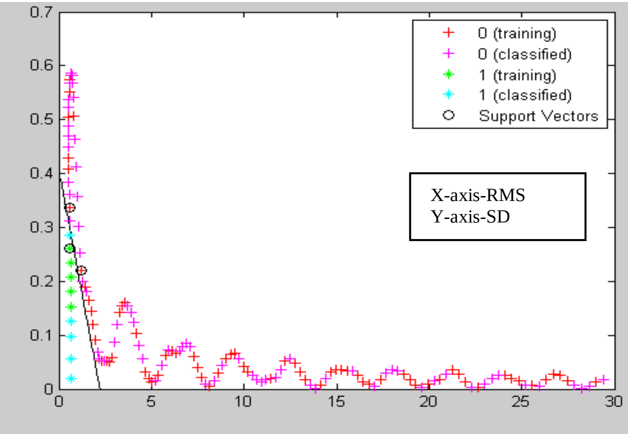

Fig 6.2 Testing by SVM

The failure detection rate is reduced compared to the CART method in Support Vector Machine (SVM) and severity also assessed from the above analysis. In fig 5.4 shows the signal duration is given in terms of minutes, hours and seconds and heart bit rate (HBR) is given as measured previously. Heart Rate Variability is measured by time- or frequency-domain methods. The time-domain HRV measures are based on beat-to-beat intervals where as frequency-domain analysis expresses HRV in terms of its constituent frequency components. HRV analysis has emerged as a diagnostic tool that quantifies the functioning of the autonomic regulation of the heart and heart's ability to respond.

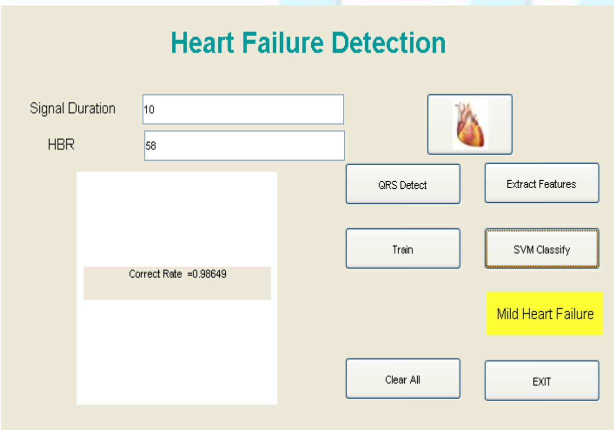

Fig 6.3 Accuracy achieved by SVM

RR variability measure provides spectral analyses on the ECG recordings to quantify the features. HRV are assessed by Bayesian classifier. The ECG signals are simulated from the system is preprocessed to extract the QRS complexes. In the above fig 5.4, the QRS is detected, HRV features are extracted from the given signals, it is trained by the SVM and SVM accuracy is measured. A classification method (SVM) to improve the sensitivity of HF prediction.SVM is a new generation learning system based on recent advances in statistical learning theory.

8. CONCLUSION

The preliminary results of classifiers for HF severity detection, which are innovative in comparison to the others previously published. These results are consistent and confirm that patients suffering from HF present a depressed HRV. Similarly, those patients suffering from severe HF present a more depressed HRV compared to those affected by mild HF. The existing study gives higher precision and specificity values, but lower sensitivity. Moreover, the classifier is fully human understandable. The results produced by the decision tree method provide classification of patients under CART method. Support Vector Machine classification technique is used for sensitivity analysis.SVM offers number of advantages over other types of multivariate classifiers.SVM have gained wide acceptance due its solid theoretical basis and high generalization ability. The two key features of SVM are generalization theory and kernel functions.

References

[1] M. H. Asyali, "Discrimination power of long-term heart rate variability measures," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Bio. Soc.*, Sep.

[2] J. T. Bigger, J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, J. Schneider and, P. K. Stein "RR Variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial-infarction".

[3] L. Liao, J. G. Jollis, K. J. Anstrom, D. J. Whellan, D. W. Kitzman, G. P. Aurigemma, D. B.Mark, K. A. Schulman, and J. S. Gottdiener, "Costs for heart failure with normal vs reduced ejection fraction," *Arch. Int. Med.*, vol. 177, no. 1, pp. 112–118, 2007.

[4] "Uniqueness of medical data mining", artif.intell., med., vol 26,no.1.

[5] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*.

[6] L. Pecchia, P. Melillo,M. Sansone, and M. Bracale, "Heart rate variability in healthy people compared with patients with congestive heart failure,"in *Proc. 9th Int. Conf. Inform. Technol. Appl. Biomed.*, Larnaca, Cyprus,2009, pp. 1–4.

[7] L. Pecchia, P. Melillo, M. Sansone, and M. Bracale, "Discrimination power of short-term heart rate variability

measures for CHF assessment,"*IEEE Trans. Inform. Technol. Biomed*.

[8] M. Jessup, W. T. Abraham, D. E. Casey, A. M. Feldman, G. S. Francis, T. G. Ganiats, M. A. Konstam, D. M. Mancini, P. S. Rahko, M. A. Silver, L. W. Stevenson, C. W. Yancy, S. A. Hunt, M. H. Chin, H. F. W. Comm, and W. C. Members, "2009 focused update: ACCF/AHA guidelines for diagnosis and management of heart failure in adults a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines," *Circulation*, vol. 119, no. 14, pp. 1977–2016, 2009.

[9] L. Pecchia, P. Melillo, M. Sansone, and M. Bracale, "Discrimination power of short-term heart rate variability measures for CHF assessment," *IEEE Trans. Inform. Technol. Biomed.*, to be published. DOI: 10.1109/TITB.2010.2091647.

[10] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Belmont, CA: Wadsworth, 1984.

[11] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a newresearch resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[12] M. W. Browne, "Cross-validation methods," *J. Math. Psychol.*, vol. 44, no. 1, pp. 108–132, 2000.

[13] S. Koch, "Home telehealth—Current state and future trends," *Int. J. Med. Inform.*, vol. 75, no. 8, pp. 565–576, 2006.

[14] C. S. Pattichis, C. N. Schizas, M. S. Pattichis, E. Micheli-Tzanakou, E. C. Kyriakou, and D. I. Fotiadis, "Introduction to the special section on computational intelligence in medical systems," *IEEE Trans.Inform.Technol. Biomed.*, vol. 13, no. 5, pp. 667–672, Sep. 2009.

[15] A. Martinez, E. Everss, J. L. Rojo-Alvarez, D. P. Figal, and A. Garcia-Alberola, "A systematic review of the literature on home monitoring for patients with heart failure," *J. Telemed. Telecare*, vol. 12, no. 5, pp. 234–241, 2006.

[16] J. Gonseth, P. Guallar-Castillon, J. R. Banegas, and F. Rodriguez-Artalejo, "The effectiveness of disease management programmes in reducing hospital re-admission in older patients with heart failure: A systematic review and meta-analysis of published reports," *Eur. Heart J.*, vol. 25, no. 18, pp. 1570–1595, 2004.

[17] D. Lloyd-Jones, R. J. Adams, T. M. Brown, M. Carnethon, S. Dai, G. De Simone, T. B. Ferguson, E. Ford, K. Furie, C. Gillespie, A. Go, K. Greenlund, N. Haase, S. Hailpern, P. M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. M. McDermott, J. Meigs, D. Mozaffarian, M. Mussolino, G. Nichol, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, R. Stafford, T. Thom, S. Wasserthiel-Smoller, N. D. Wong, and J. Wylie-Rosett, "Heart disease and stroke statistics—2010 Update. A report from the American Heart Association," *Circulation*, vol. 121, no. 12, pp. e1–e260, 2009.

[18] M. R. Cowie, A. Mosterd, D. A.Wood, J.W. Deckers, P. A. PooleWilson, G. C. Sutton, and D. E. Grobbee, "The epidemiology of heart failure," *Eur. Heart J.*, vol. 18, no. 2, pp. 208–225, 1997.

[19] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A.Malliani ,A. J. Moss, and P. J. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, no. 3, pp. 354–381, 1996.

[20] J. T. Bigger, J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky,W. J. Schneider, and P. K. Stein, "RR Variability in healthy, middle-aged persons compared with patients with chronic coronary heart-disease or recent acute myocardial-infarction," *Circulation*, vol. 91, no. 7, pp. 1936–1943, 1995.

[21] Musialik-Lydka, B. Sredniawa, and S. Pasyk, "Heart rate variability in heart failure," *Kard. Pol*, vol. 58, no. 1, pp. 10–16, 2003.

[22] G. N. Arbolishvili, V. Y. Mareev, Y. A. Orlova, and Y. N. Belenkov, "Heart rate variability in chronic heart failure and its role in prognosis of the disease," *Kardiologiya*, vol. 46, no. 12, pp. 4–11, 2006.

[23] T. D. J. Smilde, D. J.VanVeldhuisen, andM. P.Van Den Berg, "Prognostic value of heart rate variability and ventricular arrhythmias during 13-year follow-up in patients with mild to moderate heart failure," *Clin. Res. Cardio.*, vol. 98, no. 4, pp. 233–239, 2009.

[24] M. H. Asyali, "Discrimination power of long-term heart rate variability measures," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Bio. Soc.*, Sep. 2003, pp. 200–203.

[25] Y. Isler and M. Kuntalp, "Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive.

[26] L. Liao, J. G. Jollis, K. J. Anstrom, D. J. Whellan, D. W. Kitzman, G. P. Aurigemma, D. B.Mark, K. A. Schulman, and J. S. Gottdiener, "Costs for heart failure with normal vs reduced ejection fraction," *Arch. Int. Med.*, vol. 166, no. 1, pp. 112–118, 2006.