34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Heart Failure Analysis Dashboard for Patient's Remote Monitoring Combining Multiple Artificial Intelligence Technologies

G. Guidi, M. C. Pettenati, R. Miniati and E. Iadanza, *Member IEEE*

Abstract— **In this paper we describe an Heart Failure analysis Dashboard that, combined with a handy device for the automatic acquisition of a set of patient's clinical parameters, allows to support telemonitoring functions. The Dashboard's intelligent core is a Computer Decision Support System designed to assist the clinical decision of non-specialist caring personnel, and it is based on three functional parts: Diagnosis, Prognosis, and Follow-up management. Four Artificial Intelligence-based techniques are compared for providing diagnosis function: a Neural Network, a Support Vector Machine, a decision tree and a Fuzzy Expert System whose rules are produced by a Genetic Algorithm. State of the art algorithms are used to support a score-based prognosis function. The patient's Follow-up is used to refine the diagnosis.**

I. Introduction

In this paper we present an Heart Failure (HF) analysis dashboard. We mainly focus on explaining the system's smart core, a Computer Decision Support System (CDSS) [1]. The main objective of the dashboard is to make feasible the telemonitoring by non-specialist staff equipped with a device for automatic detection of patient's vital signs at home or other point of care, also allowing the collaboration with specialist staff. To do this a CDSS is necessary to help the clinical decision of non-specialist staff such as General Practitioners (GP) or Nurses in order to involve the specialist only in severe cases. The CDSS receives as input anamnestic and instrumental data and provides diagnosis and prognosis output related to the current state of the patient and a comparison with respect to the patient's clinical history.

The Dashboard encompasses a CDSS, for assisting decision of non-specialist staff, and a Web portal, to allow collaboration between specialist and non-specialist staff in case of need. The CDSS uses Artificial Intelligence (AI) technology to perform a first round of diagnosis, focusing mainly on establishing the patient's current status in terms of HF severity (Mild, Moderate, Severe), that is the main objective of this paper. State of the art algorithms are used to support a score based prognosis function. Additional information such as HF worsening or HF type (Chronic stable, Acute, Chronic with frequent exacerbations, etc.) is provided by comparing patient's current status with previous system diagnosis in the database. The system also gives an outcome prediction using time series technology and comparing the patient's parameters trend with known outcomes (still in progress). The Web application provides interfaces both for non-specialist caring personnel as well as for heart specialists. The non-specialist client interface allows the input of patient anamnestic and instrumental data, the management of patient follow-up and the displaying of the CDSS response. The heart specialist client interface allows the cardiologist to consult the patients' database and to display each patient alarm produced by the CDSS, and any reports added by the non-specialist operator. Some reviews have been published to demonstrate the usefulness of CDSS in medicine: A.Garg et Al. [2] show that the CDSS improves practitioner performance; K.Kawamoto et Al. [3] assert that CDSS significantly improved clinical practice in 68% of trials. The Cochrane Collaboration published a systematic review regarding HF telemonitoring showing that the telemonitoring for HF patients reduces hospitalization and mortality [4]. In this work we combine the proven advantages of the HF telemonitoring, with the advantages of a CDSS.

II. Previous Works

In order to design our CDSS we have planned to integrate and compare the most used technologies in the field of diagnosis and classification of HF.

A. *Neural Networks*

Elfadil et al. classify HF patients in four groups by using supervised and unsupervised Neural Networks (NN). [5] (supervised NN: 83.65% Accuracy, Unsupervised NN: 91.43% Accuracy). Gharehchopoghi et al. use NN to detect presence or absence of HF obtaining a 95% learning ability on the training set and a 85% of correctly classified patients in the test set [6].

B. *Support Vector Machines*

Guiqiu Yang et al. combined two Support Vector Machines (SVM) to classify HF patients in three groups. (74.4% Global Accuracy, 78.8% - 87.5% - 65.6% Accuracy to classify Healthy - HF prone - HF respectively) [7]. Wang et al. combined SVM with other signal analysis techniques to distinguish healthy persons from HF patients, obtaining an accuracy of 89%. [8]

C. *Fuzzy Expert System*

Akinyokun et al. used a neuro-fuzzy system to classify HF patients in three categories (Mild HF, Moderate HF and

G.Guidi is with the Department of Electronics and Telecommunications – University of Florence, Via Santa Marta 3 50039 - Florence, Italy (e-mail: gabriele.guidi@unifi.it).
M.C. Pettenati is with ICON (International Center of Computational Neurophotonics) Foundation, 50019 Florence, Italy
R. Miniati is with the Department of Electronics and Telecommunications – University of Florence, Via Santa Marta 3 50039 - Florence, Italy
E. Iadanza is with the Department of Electronics and Telecommunications – University of Florence, Via Santa Marta 3 50039 - Florence, Italy (corresponding author: 0039-3475922874; e-mail: ernesto.iadanza@unifi.it).

Severe HF). In particular they trained a NN and extracted fuzzy rules from the trained dataset. For the NN they obtained an average training Normalized Mean Square Error of 0.026 and a strong correlation between the orthodox results and the neuro-fuzzy results was observed. [9]
Adeli et al. built a Mandami Fuzzy Expert System for classifying patients with heart disease (no specific HF) in five groups. With 44 manual entered rules they obtained a 94% of coherence with an expert human decision [10]. Chiarugi et al. implemented a CDSS for HF that analyzed electro- and echocardiograms. The rules are input in the knowledge base using guidelines and experts' interviews [11]. This scenario is quite different from our telemonitoring environment because we cannot use echocardiograms images and our primary goal is to automatically discover rules from our dataset.

D. Decision Tree

Candelieri et al. developed a decision tree (coming from data mining techniques) to detect patient's destabilizations [12] (Decision tree: 88% Accuracy - SVM: 82% Accuracy, - SVM+GA: 87% Accuracy). Pechenizkiy et al. used decision trees to predict HF patients hospitalizations [13] . In addition Pecchia et al. used decision tree techniques to classify patients in three groups of severity (Healthy, Moderate, Severe) using Heart Rate Variability measurements. (HF vs Normal Subject: 96% Accuracy - Severe vs Moderate: 79.3% Accuracy). [14]

E. Prognosis

For HF prognosis various models can be found, that perform regression techniques on large databases of patients. A 2008 review [15] analyzed four models: - *The Seattle Heart Failure Model (SHFM),* for outpatients, the output provides the probability of death within 5 years [16]; - *CHARM Model,* (The Candesartan in Heart Failure: Assessment of Reduction in Mortality and morbidity) it is a model derived on a large database of outpatients that are part of an investigation on the reduction of mortality due to treatment with Candesartan anti-hypertensive [17]; - EFFECT, model based on inpatients, the output provides the probability of death within 30 days or within a year [18]; - *ADHERE,* model based on inpatients; it provides the probability of death in hospital [19].

III. Materials and Methods

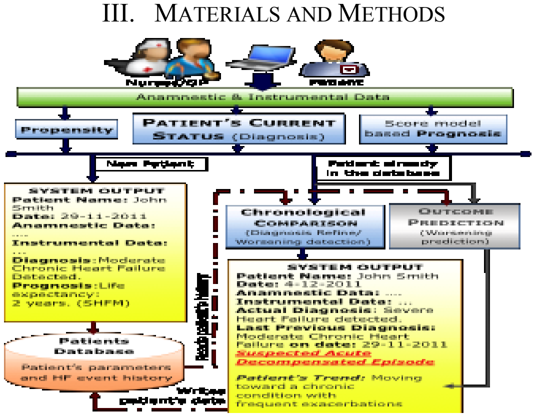

Figure 1. Schematic overview of CDSS operation

The users (nurses or GP) enter the patient's medical history and instrumental parameters that are processed by the blocks of Patient's Current Status (PCS) to perform a diagnosis of current HF severity. Using Score Model based Prognosis block the user can choose whether to evaluate the patient's prognosis by selecting one of the above mentioned models. Propensity block calculates an HF risk score obtained by independent HF predictors identified in the Framingham Study [20]. In case of new patient no other follow-up information is available, so the system provides a simple output as shown in Fig. 1. However if the patient is already in the database, once the PCS block has processed the parameters, the system activates Chronological Comparison and Outcome Prediction blocks (under construction) to perform a diagnosis refinement and a trend-based prognosis respectively. By comparing current PCS output with the previous ones the system is able to establish the type of HF (chronic stable HF, chronic HF with frequent or rare exacerbations, acute episode etc.) and if the patient status is worsened. For implementing PCS block we adopted some proven Artificial Intelligence (AI) techniques as explained in the previous work section: Neural Network, SVM, Decision Tree and Fuzzy Expert System, trained with our database and integrated into a single diagnostic dashboard. As for the Fuzzy Expert System, unlike the previous studies, the rules will be found using a genetic algorithm using Pittsburgh approach [21]. This is obtained by optimizing the algorithm in [22] to receive as input an echocardiography image, in order to work with our data. The four AI technologies are compared and the PCS block provides a three level output: Mild HF, Moderate HF, Severe HF. At the current stage of our work, all the AI algorithms are prototypes developed in Matlab and objective of this study is also determining the most performing AI technique for our goals. The AI functional block is trained using an anonymised database of patients affected by HF with varying severity degree and treated by the Cardiology Department of the Hospital Santa Maria Nuova in Florence, Italy. In this database currently there are 136 patients, 50 of which in a Mild HF status, 39 in the Moderate HF status and 47 in Severe HF status. 92 out of 136 patients are male and 44 are female. We used hold-out technique so the patients dataset was randomly divided in a training set of 100 patients and a test set of 36 patients as shown in Table I.

TABLE I. Train-Test set

Train Test	Dataset Structure			
	Mild HF	*Moderate HF*	*Severe HF*	*Total*
Train set	35	31	34	100
Test set	15	8	13	36

CDSS has been trained with the following 12 parameters: age, sex, weight, systolic and diastolic blood pressure, heart rate, NYHA, ejection fraction (EF), Brain Natriuretic Peptide (BNP), ECG parameters (atrial fibrillation, left bundle branch block, ventricular tachycardia). BNP and EF will not be assessed at each follow-up as they have a slower variation related to the change of the patient state. Furthermore, while some portable devices exist to assess BNP, the ejection fraction can be evaluated only with the ultrasound examination that is performed at point of care. In order to provide an exhaustive management dashboard the system also memorize etiology, comorbidity and therapy. Web-based dashboard including cardiologist and non-specialist clients are developed using ASP.NET framework.

A. PCS block

We adopted two feed forward - back propagation NN, one with 10 neurons in the hidden layer (NN_10) and one with 4 neurons (NN_4). Both NN have 12 input neurons and 3 output neurons. SVM is a binary classifier, so two SVMs are combined to obtain the desired three level output. Regarding decision tree techniques we implemented a Classification and Regression Tree (CART), using the appropriate Matlab function, and Gini split criterion. With reference to fig. 3, the genetic algorithm for producing fuzzy rules works as follows. An initialization block randomly generates a population of N-Rules Sets each composed of M rules. The Desired Output block (DO) consists in anamnestic and instrumental data of patients with corresponding outcome provided by the doctor and categorized in Mild, Moderate, Severe HF. The Input block consists only in patient data. The Mandami FIS (Fuzzy Inference System) generates outputs using Input block patient's data (after fuzzyfication) and using the rules generated by the initialization block.

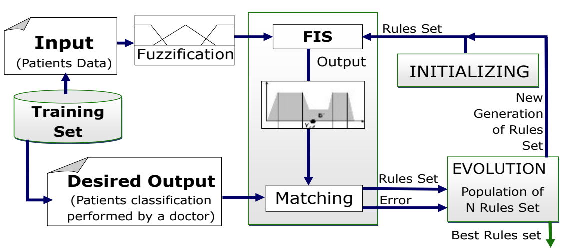

Figure 2. Genetic Algorithm operation scheme

Matching block compares, for each patient, the input-output couple produced by the FIS with the corresponding couple in the DO. If a rules set produces the same output contained in the Training Set, it is considered a good set. In this way the goodness of each rules set of population is assessed and number of correctly classified patients is provided for each rules set (Fitness). The population now evolves genetically. At the end of the evolution the Best Rules Set will be obtained. This is the set of rules that produced the outcome that less deviates from those contained in the DO. The input membership functions are decided using expert advice. Binary input such presence/absence of branch block or atrial fibrillation are not fuzzed. Some of the algorithms we used will have problems analyzing missing data sets. Hence, during training we used complete data, and during final user use complete data are required at least for the first follow-up, while for the subsequent follow-ups the system reloads any missing data from previous patient check. To prevent over-fitting, we limited the neurons in the NN (NN_4), we limited to 45 the number of rules for the fuzzy system, and we tried to prune the CART, but best result are obtained the whole tree, containing 26 nodes.

IV. Results

Currently the CDSS is implemented as a prototype. The AI parameters refinement is still underway so as to validate and improve our preliminary results about PCS block. Table II shows the system performance. Using 10 neurons in the hidden layer the NN correctly reclassified 98 out of 100 patients in the training set, unfortunately not maintaining a good generalization ability as it correctly classified 24 out of 36 patients included in the test set (66.6% Accuracy). Using only 4 neurons in the hidden layer instead the performance in the training set is worse (78 out of 100 training set patients) but we obtained a better generalization capability (27 out of 36 test set patients, 75% accuracy). The two combined SVM blocks correctly reclassified 74 out of 100 patients in the training set, and 25 out of 37 patients in the test set (69.4% Accuracy). Genetic algorithm has been tested with different evolutionary parameters. The best results are achieved with a population composed of 30 individuals each composed by 45 rules. Algorithm evolves for 1000 generations. The results still show a 26% of residual error in the training set. This means that the 45 generated rules correctly classified 74% of the training set (same of SVM). These rules also correctly classified 26 patients of the test set (72.2% Accuracy). The CART seems to perform slightly better with 84% of correctly classified patients in the training set and 28 correctly classified patients in the test set (77.8% Accuracy).

TABLE II. Performance (N° of correctly classified patients)

AI	Correctly classified		*Test set Accuracy%*	*Training set Accuracy%*
	Training	*Test*		
NN_10	98	24	66.2%	98%
NN_4	78	27	75%	78%
SVM	74	25	69.4%	74%
Fuzzy-Genetic	74	26	72.2%	74%
CART	84	28	77.8%	84%

TABLE III. Confusion Matrix of each Algorithm

NN_4	***True Mild***	***True Moderate***	***True Severe***
Classified Mild	10	2	0
Classified Moderate	5	5	1
Classified Severe	1	0	12
SVM	***True Mild***	***True Moderate***	***True Severe***
Classified Mild	10	3	0
Classified Moderate	3	4	2
Classified Severe	1	0	11
Fuzzy-Genetic	***True Mild***	***True Moderate***	***True Severe***
Classified Mild	13	4	0
Classified Moderate	3	1	1
Classified Severe	0	2	12
CART	***True Mild***	***True Moderate***	***True Severe***
Classified Mild	11	3	0
Classified Moderate	0	4	0
Classified Severe	5	0	13

V. Discussion

The PCS block slight varies its performance level depending on the adopted AI techniques. NN (using 4 neurons) and CART produces quite good results about Test set Accuracy if compared with other studies that assess HF severity such [4] or [9] (Severe vs. Mild HF). SVM and NN (using 10 neurons) produced worse results, probably because combining two SVM for obtaining three levels output emphasizes the errors generated in the SVM1 or SVM2 and 10 hidden layer-neurons cause a strongly over training in the NN. Finally, the Fuzzy-genetic technique performances are at present not completely measurable, since it would require many more training patients to function properly. Indeed, while the fact of having 12 inputs would force to have a larger number of fuzzy rules, the fact of having few patients in the training set requires keeping down the number of rules in order not to compromise the generalization capability. As shown in Table III the most classification problems occur with the class Mild HF. CART seems to have the tendency to produce false positives for the class Severe HF. Globally we consider the above results as quite good, since we are implementing a three output-levels classification of HF severity and the analyzed patients parameters are not so strongly output-correlated as in distinguishing HF from healthy. However, 36 Test set and 100 Training set elements are insufficient to correctly evaluate algorithms performs. When more data will be available we will perform additional tests.

A. Advantages and disadvantages of the used algorithms

NN pros: quick and easy to train, relatively easy to control over-fitting. NN cons: Do not provide a easily understood representation of the learned knowledge. Fuzzy-Genetic pros: The knowledge is in the form of humanly understandable if-then rules. Fuzzy-Genetic cons: It needs a long and computationally complex genetic evolution. CART pros: humanly understandable threshold type if-then rules are obtained. CART cons: high risk of over-fitting, in our case the pruning resulted in not good Test set performance.

SVM pros: all the advantages of SVM, not least that of being able to control the generalization ability with the number of train patients. SVM cons: They are binary classifiers, and combine two SVM for three levels output can generate errors.

VI. CONCLUSIONS

In this paper we presented the core system grounding of a dashboard for the remote monitoring of HF patients. This system, coupled with a multiparametric device, could constitute a kit for patients telemonitoring performed by a team of non-specialized caregivers. If the patient's status is severe or worsening, the CDSS generates alarms and sends them to cardiologists. In this paper we focused on providing a detailed description of the CDSS design and implementation and some preliminary results as quite good. The dashboard also includes two Web interfaces allowing collaboration between specialist and non-specialist clinical staff. We are evaluating the operability of the CDSS with existing HF telemonitoring scenarios, even supposing a link between the multiparametric devices and the non specialist staff client in order to acquire parameters automatically and speed up the operations to be performed at each follow-up.

VII. REFERENCES

[1] G. Guidi, E. Iadanza, M. C. Pettenati, M. Milli, F. Pavone, and G. Biffi Gentili, "Heart Failure Artificial Intelligence-based Computer Aided Diagnosis Telecare System (study design)," *ICOST 2012, LNCS, In press*.

[2] A. X. Garg, N. K. J. Adhikari, J. Beyene, J. Sam, and R. B. Haynes, "Effects of Computerized Clinical Decision Support Systems on Practitioner Performance," vol. 293, no. 10, pp. 1223-1238, 2005.

[3] K. Kawamoto, C. a Houlihan, E. A. Balas, and D. F. Lobach, "Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success.," *BMJ (Clinical research ed.)*, vol. 330:765, Apr. 2005.

[4] S. Inglis et al., "Structured telephone support or telemonitoring programmes for patients with chronic heart failure (Review)," *The Cochrane Collaboration*, no. 6, 2011.

[5] N. Elfadil and I. Ibrahim, "Self Organizing Neural Network Approach for Identification of Patients with Congestive Heart Failure," in *International Conference on Multimedia Computing and Systems (ICMCS), 2011*, 2011, pp. 1-6.

[6] F. S. Gharehchopoghi and Z. A. Khalifelu, "Neural Network Application in Diagnosis of Patient: A Case Study," *Computer Networks and Information Technology (ICCNIT), 2011 International Conference on*, pp. 245-249, 2011.

[7] G. Yang, Y. Ren, Q. Pan, and G. Ning, "A heart failure diagnosis model based on support vector machine," *IEEE International Conference on Biomedical Engineering and Informatics*, no. Bmei, pp. 1105-1108, 2010.

[8] C.-chih Wang, "SVD and SVM based approach for Congestive Heart Failure Detection from ECG Signal," *Computers and Industrial Engineering (CIE), 2010 40th International Conference on*, pp. 1-5, 2010.

[9] C. O. Akinyokun, O. U. Obot, and F.-michael E. Uzoka, "Application of Neuro-Fuzzy Technology in Medical Diagnosis: Case Study of Heart Failure," *IFMBE Proceedings*, vol. 25/XII, pp. 301-304, 2009.

[10] A. Adeli and M. Neshat, "A Fuzzy Expert System for Heart Disease Diagnosis," *Proceedings of the International MultiConference of Engineers and Computer Scientists.*, vol. I, 2010.

[11] F. Chiarugi, S. Colantonio, D. Emmanouilidou, M. Martinelli, D. Moroni, and O. Salvetti, "Decision support in heart failure through processing of electro- and echocardiograms.," *Artificial intelligence in medicine*, vol. 50, no. 2, pp. 95-104, Oct. 2010.

[12] A. Candelieri, D. Conforti, F. Perticone, and A. Sciacqua, "Early Detection of Decompensation Conditions in Heart Failure Patients by Knowledge Discovery: The HEARTFAID Approaches," *Computers in Cardiology*, pp. 893-896, 2008.

[13] M. Pechenizkiy, E. Vasilyeva, and I. Zliobait, "Heart Failure Hospitalization Prediction in Remote Patient Management Systems," *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*, no. 12–15 Oct. 2010, pp. 44-49, 2010.

[14] L. Pecchia, P. Melillo, and M. Bracale, "Remote health monitoring of heart failure with data mining via CART method on HRV features.," *IEEE transactions on bio-medical engineering*, vol. 58, no. 3, pp. 800-4, Mar. 2011.

[15] W. C. Levy and D. T. Linker, "Prediction of mortality in patients with heart failure and systolic dysfunction.," in *Current cardiology reports, Vol.10 issue 3*, vol. 10, no. 3, 2008, pp. 198-205.

[16] W. C. Levy et al., "The Seattle Heart Failure Model: prediction of survival in heart failure.," *Circulation*, vol. 113, no. 11, pp. 1424-33, Mar. 2006.

[17] S. J. Pocock et al., "Predictors of mortality and morbidity in patients with chronic heart failure.(CHARM)," *European heart journal*, vol. 27, no. 1, pp. 65-75, Jan. 2006.

[18] D. S. Lee, P. C. Austin, J. L. Rouleau, P. P. Liu, and D. Naimark, "Predicting Mortality Among Patients Hospitalized for Heart Failure Derivation and Validation of a Clinical Model (EFFECT)," *Hospitals*, vol. 290, no. 19, pp. 2581-2587, 2003.

[19] G. C. Fonarow, K. F. Adams, W. T. Abraham, C. W. Yancy, and W. J. Boscardin, "Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis.," *JAMA*, vol. 293, no. 5, pp. 572-80, Feb. 2005.

[20] W. B. Kannel, R. B. D'Agostino, H. Silbershatz, a J. Belanger, P. W. Wilson, and D. Levy, "Profile for estimating risk of heart failure.," *Archives of internal medicine*, vol. 159, no. 11, pp. 1197-204, Jun. 1999.

[21] A. A. Freitas, *Data Mining and Knowledge Discovery with Evolutionary Algorithms. Springer.*, no. August. 2002, p. 280.

[22] L. Bocchi and F. Rogai, "A genetic fuzzy rules learning approach for unseeded segmentation in echography," *EvoApplications 2012, LNCS 7248*, pp. 305–314, 2012.