

# Predicting adherence of patients with HF through machine learning techniques

Georgia Spiridon Karanasiou<sup>1</sup>, Evanthia Eleftherios Tripoliti<sup>1</sup>, Theofilos Grigorios Papadopoulos<sup>2</sup>, Fanis Georgios Kalatzis<sup>1</sup>, Yorgos Goletsis<sup>3</sup>, Katerina Kyriakos Naka<sup>4,5</sup>, Aris Bechlioulis<sup>4,5</sup>, Abdelhamid Errachid<sup>6</sup>, Dimitrios Ioannis Fotiadis<sup>1,2</sup> ✉

<sup>1</sup>Department of Biomedical Research, Institute of Molecular Biology and Biotechnology, FORTH, GR 45110 Ioannina, Greece

<sup>2</sup>Unit of Medical Technology and Intelligent Information Systems, University of Ioannina, GR 45110 Ioannina, Greece

<sup>3</sup>Department of Economics, University of Ioannina, GR 45110 Ioannina, Greece

<sup>4</sup>Michaelidion Cardiac Center, University of Ioannina, GR 45110 Ioannina, Greece

<sup>5</sup>Department of Cardiology, University of Ioannina, GR 45110 Ioannina, Greece

<sup>6</sup>Université de Lyon, Institut de Sciences Analytiques, ISA, FR 69100 Villeurbanne, France

✉ E-mail: fotiadis@cc.uoi.gr

Published in Healthcare Technology Letters; Received on 13th May 2016; Revised on 26th August 2016; Accepted on 7th September 2016

Heart failure (HF) is a chronic disease characterised by poor quality of life, recurrent hospitalisation and high mortality. Adherence of patient to treatment suggested by the experts has been proven a significant deterrent of the above-mentioned serious consequences. However, the non-adherence rates are significantly high; a fact that highlights the importance of predicting the adherence of the patient and enabling experts to adjust accordingly patient monitoring and management. The aim of this work is to predict the adherence of patients with HF, through the application of machine learning techniques. Specifically, it aims to classify a patient not only as medication adherent or not, but also as adherent or not in terms of medication, nutrition and physical activity (global adherent). Two classification problems are addressed: (i) if the patient is global adherent or not and (ii) if the patient is medication adherent or not. About 11 classification algorithms are employed and combined with feature selection and resampling techniques. The classifiers are evaluated on a dataset of 90 patients. The patients are characterised as medication and global adherent, based on clinician estimation. The highest detection accuracy is 82 and 91% for the first and the second classification problem, respectively.

**1. Introduction:** Heart failure (HF) is a chronic, progressive disease that affects 26 million people globally. The number increases by 3.6 million every year [1]. HF affects mainly older adults that are frequently hospitalised. The increasing incidence of HF is a considerable economic burden, taking into account that the annual cost of HF is ~6000€ per person per year, whereas the hospitalisation cost is more than 9000€ [1].

HF hospitalisation can result from a variety of reasons, among which medication non-adherence is the most important [2]. A lot of studies have focused on medication non-adherence and investigated this factor. Medication non-adherence can be utilised as the best predictor of hospitalisation in HF patients [3]. Medication adherence can be a result of a variety of self-care behaviours including adherence to suggested treatment regimens, symptom monitoring and management that can prevent HF events and improve the patient's condition. Recent studies showed that the medication non-adherence rates are between 40 and 60% in HF patients [3, 4].

However, the optimal management of HF patients requires the patient to be adherent, not only to medication treatment, but also to the guidelines/suggestions, provided by the health care professional, related to nutrition and physical activity exercising. Patient adherence, in general, is a multifactorial problem and there are several factors that adherence can be influenced by such as: (i) social and economic dimensions, (ii) health care system, (iii) patient health condition including New York Heart Association (NYHA) class and the presence of other comorbidities, (iv) suggested therapy, (v) knowledge about disease and (vi) patient-related factors. Specifically, education showed a positive influence on adherence, since patients with low-educational level face difficulties in understanding the medication instructions [5, 6]. The duration of the therapy, the frequency and the intake of different medications have a negative influence in patients with chronic diseases [6, 7]. Patients of higher or middle age seem to be less adherent [5, 8].

In addition, gender, personality and mental comorbidities influence adherence rates [9].

For the measurement of adherence, direct and indirect methods have been implemented. Measurements of drug levels in plasma and urine belong to direct methods, whereas pill counts and self-reported questionnaires for estimating medication adherence [10] are indirect methods. The Medication Event Monitoring System (MEMS) is currently considered as the gold standard for medication adherence. It is an electronic 'umbilical cord' that tracks the dates and times of bottle cap openings of patient's medication. MEMS is a mean of checking and ensuring that the patient perceives the prescribed drugs.

Taking into account the above-mentioned facts, numerous studies approached medication non-adherence, through the identification of modifiable factors associated with medication non-adherence and the development of models aiming to predict adherence in adults with HF.

Riegel and Knafl [3] utilised adaptive modelling methods and examined patterns of medication adherence in HF patients, assuming that poor medication adherence is related to 6 months hospitalisation. Knafl and Riegel [11] used the MEMS to assess a wider variety of condition, as well as patient factors, to improve the ability of identifying those that contribute to poor medication adherence, and thus increase the risk for hospitalisation. They showed that a high number of comorbidities, daily medications and bad sleep quality are most probable for poor medication adherence. Hajduk *et al.* [12] focused on hospitalised HF patients and studied the association between cognitive impairment and adherence. The analysis showed that screening for memory impairment in HF patients could assist in identifying the high-risk non-adherent patients. Mathes *et al.* [13] showed that some consistent factors, that have a negative influence on adherence, are the ethnic minority, the unemployment and the medication cost. Dickson *et al.* [14]

analysed data from a cohort of HF patients to recognise differences in predictors of medication non-adherence by racial group and showed that interventions for addressing specific risk factors among black HF patients are required. Aggarwal *et al.* [15] performed a pilot study with the aim to receive feedback for the design of educational interventions for HF patients to improve medication adherence. Juarez *et al.* [16] conducted an analysis of data from the Practice Variation and Care Outcomes study, investigating if demographic or behavioural factors are related to the probability of categorising the HF patients in different adherence groups. Son *et al.* [17] employed two support vector machines (SVMs) to predict medication adherence. The input of the two SVMs is a set of five and seven predictor variables selected through the application of a feature selection (FS) process to a set of 11 features, which according to the literature affect patients' medication adherence. In HEARTEN project [18], adherence is addressed by two different modules of the HEARTEN knowledge management system, *adherence risk module* and *treatment adherence module*. *Adherence risk module* provides an estimation of the adherence of the patient allowing the experts to focus to this specific patient. Specifically, the expert is informed whether a new patient is likely to be adherent or not and in which adherence group this patient belongs to (e.g. low adherence, medium adherence and high adherence). *Treatment adherence module* checks if the patient is adherent or not to the guidelines provided by the experts regarding the medication, the nutrition and the physical activity.

The already available approaches for patient adherence in HF focus mainly on examining and correlating the factors that influence the levels of patient adherence. From these studies, only Son *et al.* [17] examined which of these factors can act as predictors of medication adherence in the HF patients through the utilisation of SVM classifier. HEARTEN *adherence risk module* predicts not only medication adherence of the patient, but also of the global adherence of the patient.

**2. Materials and methods:** Two classification problems are addressed: (i) if the patient is global adherent or not and (ii) if the patient is medication adherent or not. To achieve this, a three stages method is proposed, as depicted in Fig. 1. A detailed description of these stages follows.

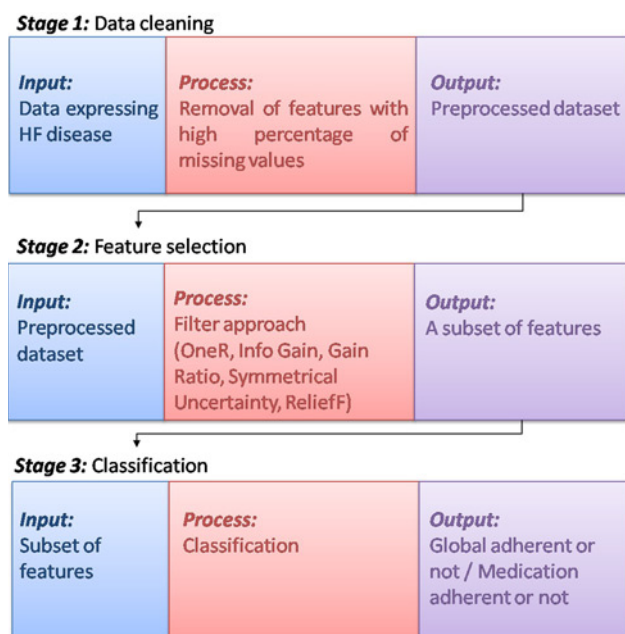


Fig. 1 Flowchart for the prediction of adherence of the patient

The first stage focuses on the removal of features that present high percentage of missing values. The second stage aims to identify predictors of adherence/non-adherence of the patients. The third stage is the core stage of the proposed method, since it informs the medical experts if the patient is adherent or not.

**2.1. Dataset:** For the evaluation of the classifiers, retrospective data from 90 patients are utilised. The data were provided by the second Department of Cardiology, University Hospital of Ioannina. Patients: (i) who have been diagnosed with HF (Framingham criteria) with continuous symptoms and frequent recurrence, (ii) who have HF in the functional NYHA II–IV classes with optimal treatment (angiotensin converting enzyme inhibitors, sartanes, beta-blockers, furosemide, spironolactone and/or digoxin), (iii) with a recent hospitalisation, emergency admission or specialised consultation (at least one in the last six months) for decompensated HF, disregarding the fraction of ejection and any etiology, (iv) who, at least, have undergone in the last 12 months one electrocardiogram and HF compatible symptoms, are included in the Letter. Patients under 18 years old, with very severe HF or in situation of biological termination, with morbid obesity (body mass index. larger than 40), with advanced chronic kidney failure (glomerular filtration rate under 30 ml/min) are excluded.

Patients are characterised as adherent or not, based on clinician estimations. They are characterised in terms of adherence: (i) to the guidelines provided by the clinician (global adherence) and (ii) to the suggested medication treatment. More specifically, from the 90 patients: (i) 61 are characterised as adherent and 29 patients as non-adherent, whereas (ii) 74 patients are characterised as medication adherent and 16 patients as non-medication adherent. The data are collected during the retrospective phase of the HEARTEN project.

The features that are recorded for each patient can be grouped to the following six categories: (i) general information (age, gender and caregiver), (ii) allergies and (iii) medical condition, which includes information regarding the KILLIP classification, the NYHA class, the smoking habit, the alcoholism habit, of the patient, as well as, the presence or not of comorbidities, (iv) drugs (the active substances, the dose and the frequency of intake), (v) biological data related with HF disease and (vi) clinical examinations (left bundle branch block or intraventricular delay, left ventricular ejection fraction). Totally, 100 features are recorded for each patient that according to the literature are correlated, positively or negatively, with patient adherence.

These six categories of features are related to the third (patient health condition including NYHA class and the presence of other comorbidities), the fourth (suggested therapy) and the sixth (patient-related factors) group of factors affecting adherence, as mentioned in Section 1. Data about social and economic dimensions, health care system and knowledge of patient about the disease will be collected during the prospective phase of the HEARTEN project.

**2.2. Data cleaning:** Features with more than 60% of missing values are removed since imputation of missing values cannot be performed due to the nature of data. A set of 80 features are retained.

**2.3. Feature selection:** To ensure the independence of the selected features from the learning algorithm and to evaluate the worth of the extracted features, a filter approach is employed. More specifically, Info Gain, Gain Ratio, Symmetrical Uncertainty, Relief-F, One-R and Chi-squared FS measures are employed [19–24].

**2.4. Classification:** Adherence estimation is expressed as a two-class classification problem. About 11 classifiers are employed: (i) random forests (RFs) [25], (ii) random tree (RT)

[25], (iii) logistic model trees (LMTs) [26, 27], (iv) J48, [28], (v) rotation forest [29], (vi) SVM [30–32], (vii) radial basis function network (RBF network) [33], (viii) Bayesian network (Bayesnet) [34], (ix) Naive Bayes [35, 36], (x) multiple layer perceptron (MLP) [37, 38] and (xi) simple classification and regression tree (simple CART) [39].

*RFs* belong to the category of ensemble classifiers since it is composed of many decision trees. Each tree votes for one of the classes. The final classification of a sample derives from a majority voting procedure. More specifically, the class that collects the majority of tree votes is the predicted class of the classified sample. Each tree of the forest is built to the maximum size without pruning, using a new subset of samples selected from the dataset. At each node of the tree, a subset of features is employed. The square root of the number of features of the dataset equals to the cardinality of this subset of features. The number of selected features remains constant throughout the construction of the forest. To determine the best split attribute for each node of the tree, the Gini index is employed. Each tree of the forest is tested on the samples that are not selected for its construction. These samples constitute the so-called out-of-bag (OOB) samples while OOB error is the error of the tree using these samples.

*RT* follows the procedure of the construction of a single tree of the *RFs* described previously.

*LMTs* are classification trees with logistic regression functions at the leaves. Linear logistic regression and tree induction classification schemes are combined. The LogitBoost algorithm consists of two major steps. First, it produces a logistic regression model for each node of the tree. Second, the node is split using the C4.5 criterion. For each LogitBoost invocation, the results of the parent node are employed. Finally, the tree is pruned using CART-based pruning.

*J48* constructs a decision tree using C4.5 tree inductive algorithm. It splits each node of the tree using the attribute with the highest normalised information gain and then it recurs on the smaller sub-lists. If all samples belong to the same class, it creates a leaf node that selects the specific class. In case the features do not fulfil the splitting criterion, it creates a decision node higher up the tree, using the expected values of the class. This procedure is performed also in case an instance of previously unseen class is encountered.

*Rotation forest*, a modification of the *RFs* algorithm, employs a linear combination of features in each splitting node. More specifically, principal component analysis is used in each one of the  $r$  subsets in which the feature set is randomly split. All the principal components are retained and the new extracted feature set is created. The data is transformed linearly into the new features. Through this new dataset, a tree classifier is trained. For the construction of the trees, the *J48* algorithm is used.

*SVM* is a supervised learning method which is utilised for classification. *SVM* is based on the definition of an optimal hyperplane, which separates the training data between the classes. The hyperplane is defined so that simultaneously the expected risk is minimised and the distance of the data points from the linear decision boundary is maximised. This to be achieved, support vectors are used. The computations in the feature space are simplified through the kernel function that is used. The kernel function must satisfy the Mercer's conditions.

*RBF network* constructs a normalised Gaussian RBF network. *RBF network* consists of: (i) an input vector, (ii) a layer of RBF neurons and (iii) an output layer with one node per category or class of data. A measure of similarity between the input and its prototype vector is computed in each RBF neuron. There are a variety of similarity functions that can be chosen, but the most popular is based on the Gaussian.

*Bayesian networks* are utilised for automatically representing domain knowledge along with data-driven probabilistic dependencies for the variables of interest. A Bayesian network is a

directed acyclic graph. The structure of the graph is defined by two sets: (i) set of nodes and (ii) set of directed edges. In each node of the graph a set of random variables is assigned, while the edges of the graph represent direct dependence among the variables.

*Naive Bayes* is a supervised learning approach that uses the probabilities of each attribute belonging to each class to make a prediction. The calculation of probabilities is simplified under the assumption, made by the Naive Bayes algorithm. According to this assumption, the probability of each attribute belonging to a given class value is independent of all other attributes.

*MLP* is a network that uses backpropagation to classify instances. It is organised in the following layers: (i) input layer, (ii) hidden layers and (iii) output layer. Each layer is constructed by a number of nodes, which have an activation function. The samples of the dataset are presented to the network through the input layer, the sample is processed to the hidden layers and the decision is given to the output layer.

*Simple CART* is a non-parametric binary decision tree learning technique. Measure of impurity depends on the nature of the features (categorical or continuous). Stopping criteria may concern the structure of the node (size, purity, values of predictors etc.) and the structure of the tree (depth). Once the growing process is completed, pruning process starts. *CART* algorithm handles missing values by utilising 'fractional instances' method or surrogate split method.

The implementation of the classifiers, provided by the Weka software (<http://www.cs.waikato.ac.nz/ml/weka/index.html>), is employed. More specifically, version Weka 3.6.13 of the software is utilised. The default values of the parameters have been used for every one of the above-mentioned classifiers and no optimisation procedure has been applied.

**3. Results:** The proposed method is evaluated using a dataset of 90 patients. For each patient 100 features, described in Section 2.1, are recorded. The feature vector of each patient is augmented by two features expressing the annotation of the patient as medication adherent (*med\_adh*) and global adherent (*adh*), provided by the experts. The proposed method is applied two times in order the two classification problems to be addressed. The first problem aims to classify a patient as global adherent or not, whereas the second aims to classify a patient as medication adherent or not. For each classification problem, two sub-datasets  $N \times M$  are created. For the first classification problem, where the target output  $c$  of the classification is global adherent or not ( $c = adh$ ), the dimensions of the two sub-datasets are  $N = 90$  instances and  $M = 81$  features (80 features plus *med\_adh*) (**1\_a**) and  $N = 90$  instances  $M = 80$  features (**1\_b**). Accordingly, in the second classification problem, where  $c = med\_adh$ , the two datasets include  $N = 90$  instances and  $M = 81$  (80 features plus the *adh*) (**2\_a**) and  $N = 90$  instances and  $M = 80$  features (**2\_b**). The motivation for the division of the dataset to **1\_a** and **1\_b** sub-datasets is to study if the inclusion of information regarding medication (non)adherence of the patient affects the prediction of global adherence, while through the creation of the **2\_a** and **2\_b** sub-datasets we aim to examine if the fact that a patient is global adherent or not affects the prediction of medication adherence. Each one of the four sub-datasets are given as input to the proposed method, described in Section 2, and the results are reported in Table 1. The combination of FS measure and classifier that provide the best accuracy (Acc) are reported only. For the evaluation of the classifiers, ten-fold stratified cross-validation procedure is applied.

Since the number of instances that belong to the global adherent and no global adherent classes are 61 and 29, respectively, while the number of instances of the medication adherent and medication non-adherent classes are 74 and 16, respectively, an imbalanced classes handling approach is employed. The imbalanced classes, an issue often found in real world applications, affect the

**Table 1** Classification results (Acc in percentage) of the proposed method

<i>I_a</i> : $c = adh, N = 90, M = 81$	
classifier – FS measure – SMOTE	Acc
RFs and gain ratio and SMOTE	82
<i>I_b</i> : $c = adh, N = 90, M = 80$	
classifier – FS measure – SMOTE	Acc
RFs and symmetrical uncertainty and SMOTE	76
<i>2_a</i> : $c = med\_adh, N = 90, M = 81$	
classifier – FS measure – SMOTE	Acc
SVM and Relief-F and SMOTE	91
<i>2_b</i> : $c = med\_adh, N = 90, M = 80$	
classifier – FS measure – SMOTE	Acc
rotation forest and info gain and SMOTE	78

performance of the classifiers, since they tend to be biased toward the majority class.

In our Letter, the synthetic minority over-sampling technique (SMOTE) [40] is applied. SMOTE is an over-sampling approach, in which the minority class is over-sampled by creating synthetic examples. SMOTE is preferred over an under sampling approach, in order the drawback of disregarding potentially useful data to be avoided. SMOTE is applied during the ten-fold cross-validation procedure and it is applied only to the training set and not to the test set in order the instances which are used for evaluation not to be employed for creating the synthetic instances in SMOTE.

The evaluation of the proposed method is performed stage by stage. The classification results after the application of the first stage is presented in Table 2. The results indicate that in most of the cases (except case *2\_a*), tree-based classifiers characterise the patient as global adherent and medication adherent with high Acc. FS process (Stage 2) described in Section 2.3 is applied and the results, only for the cases where the highest Acc is obtained, are reported in Table 3. The remaining features belong mainly to the medical condition and drugs categories of features, a fact that is expected since the presence of comorbidities leads to complicated therapeutic regimen, increased number of drugs and thus to important barriers to adherence.

Finally, classification in combination with re-sampling, using the SMOTE algorithm, Stage 3 is applied. For the two sub-datasets of the first classification problem (*I\_a* and *I\_b*), the value of the parameter that specifies the percentage of instances that are created by SMOTE is set equal to 100, while for the sub-datasets of the second classification problem (*2\_a* and *2\_b*) is set equal to 300. The values 100 and 300 are selected in order the number of instances that belong to each class to become balanced. For all the sub-datasets,

**Table 2** Classification results (Acc in percentage) without FS and without re-sampling

Classifier	Classification problem			
	<i>1_a</i>	<i>1_b</i>	<i>2_a</i>	<i>2_b</i>
<i>RF</i>	77	<b>68</b>	82	81
<i>RT</i>	72	<b>68</b>	70	68
<i>LMT</i>	<b>86</b>	59	83	79
<i>J48</i>	79	54	77	68
<i>rotation forest</i>	74	62	82	<b>82</b>
<i>SVM</i>	76	62	<b>84</b>	73
<i>RBF network</i>	63	59	79	79
<i>Bayesnet</i>	73	64	81	76
<i>Naive Bayes</i>	68	62	81	79
<i>MLP</i>	71	59	81	72
<i>simple CART</i>	<b>86</b>	<b>68</b>	80	<b>82</b>

Highest Acc in bold.

**Table 3** Classification results (Acc in percentage) with FS

	FS measure					
	1	2	3	4	5	6
<i>I_a</i> : $c = adh, N = 90, M = 81$						
<i>LMT</i>	86	86	86	86	86	<b>89</b>
<i>simple CART</i>	86	86	86	86	86	86
<i>RF</i>	87	87	<b>88</b>	87	86	87
<i>I_b</i> : $c = adh, N = 90, M = 80$						
<i>RF</i>	76	78	77	77	<b>79</b>	73
<i>RT</i>	70	77	70	74	71	72
<i>simple CART</i>	70	<b>73</b>	71	<b>73</b>	71	72
<i>2_a</i> : $c = med\_adh, N = 90, M = 81$						
<i>SVM</i>	87	92	92	92	92	<b>93</b>
<i>2_b</i> : $c = med\_adh, N = 90, M = 80$						
<i>rotation forest</i>	82	<b>87</b>	84	84	84	86
<i>simple CART</i>	82	82	82	82	82	82

Highest Acc in bold. 1: One-R, 2: info gain, 3: gain ratio, 4: Chi square, 5: symmetrical uncertainty, 6: Relief-F.

the number of nearest neighbours that are used is 5. This is a default parameter of the SMOTE algorithm and it is not optimised in the context of the current work. The results in terms of Acc, positive predictive value (PPV), sensitivity (Sens), specificity (Spec), area under curve (AUC) and F-measure are reported in Table 4. The evaluation measures are extracted by the confusion matrix. They are reported per class (except Acc) and the weighted average is computed, by the Weka software, and presented in Table 4. The evaluation procedure was repeated ten times and the average Acc and the standard deviation (std) of the above-mentioned evaluation measures are reported in Table 5.

**4. Discussion:** We propose an automated supervised method for the prediction of adherence of the patients with HF. The method is

**Table 4** Classification results, in terms of Acc, PPV, Sens, Spec, AUC and F-measure of the proposed method

		Evaluation measure, %					
<i>I_a</i> : $c = adh, N = 90, M = 81$							
confusion matrix	Acc	PPV	Sens	Spec	AUC	F-measure	
56	5	82	84	92	62	79	88
11	18		78	62	92	79	69
weighted average		82	82	72	79	82	
<i>I_b</i> : $c = adh, N = 90, M = 80$							
confusion matrix	Acc	PPV	Sens	Spec	AUC	F-measure	
51	10	76	81	84	60	73	82
12	17		63	59	83	73	61
weighted average		75	76	68	75	75	
<i>2_a</i> : $c = med\_adh, N = 90, M = 81$							
confusion matrix	Acc	PPV	Sens	Spec	AUC	F-measure	
71	3	91	93	96	69	82	95
5	11		79	69	96	82	73
weighted average		91	91	74	82	91	
<i>2_b</i> : $c = med\_adh, N = 90, M = 80$							
confusion matrix	Acc	PPV	Sens	Spec	AUC	F-measure	
64	10	78	87	87	38	63	87
10	6		38	38	87	63	38
weighted average		78	78	45	63	78	

Acc: accuracy, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, AUC: area under curve and FS: feature selection.

**Table 5** Average values and std of the evaluation measures after the repetition of evaluation procedure ten times

		Evaluation measure, %				
<i>I_a: c = adh, N = 90, M = 81</i>						
	Acc	PPV	Sens	Spec	AUC	F-measure
average	81	80	81	70	79	80
std	1.76	1.76	1.74	1.78	1.46	1.63
<i>I_b: c = adh, N = 90, M = 80</i>						
	Acc	PPV	Sens	Spec	AUC	F-measure
average	73	72	73	62	72	72
std	2.47	2.64	2.59	3.13	2.37	2.57
<i>2_a: c = med_adh, N = 90, M = 81</i>						
	Acc	PPV	Sens	Spec	AUC	F-measure
average	90	90	90	72	81	90
std	1.26	1.24	1.25	3.42	2.07	1.21
<i>2_b: c = med_adh, N = 90, M = 80</i>						
	Acc	PPV	Sens	Spec	AUC	F-measure
average	79	77	79	41	68	78
std	2.46	2.51	2.43	5.74	4.05	2.39

Acc: accuracy, PPV: positive predictive value, Sens: sensitivity, Spec: specificity, AUC: area under curve, FS: feature selection and std: standard deviation.

based on data collected through the retrospective phase of the HEARTEN project and it consists of three stages: data cleaning, FS and classification, where an imbalanced classes handling approach is incorporated.

Literature review revealed two main categories of methods those that: (i) study how a feature is correlated with medication adherence, (ii) try to predict medication adherence through the utilisation of machine learning techniques (Son *et al.*) [19].

The proposed method belongs to the second group of methods. However, several features differentiate it from other methods reported in the literature. More specifically, it does not only examine the contribution of each feature to the medication

**Table 6** Comparison with the literature

Authors	Method	Number of patients	Number of features	Evaluation measures, %	
Son <i>et al.</i>	FS SVM	76	11	Acc	78
				PPV	78
				Sens	78
				Spec	78
				AUC	–
our work	data cleaning	90	100	Arc	<i>1_a</i> 82
	FS				<i>1_b</i> 76
	classification				<i>2_a</i> 91
	with				<i>2_b</i> 78
	imbalanced			PPV	<i>1_a</i> 82
	classes				<i>1_b</i> 75
	handling				<i>2_a</i> 91
					<i>2_b</i> 78
				Sens	<i>1_a</i> 82
					<i>1_b</i> 76
					<i>2_a</i> 91
					<i>2_b</i> 78
				Spec	<i>1_a</i> 72
					<i>1_b</i> 68
					<i>2_a</i> 74
					<i>2_b</i> 45
				AUC	<i>1_a</i> 79
					<i>1_b</i> 75
					<i>2_a</i> 82
					<i>2_b</i> 63

adherence, but also to the global adherence of the patients with HF. Furthermore, it provides an estimation if the HF patient is going to be adherent or not, regarding all aspects of patient management such as medication, nutrition and physical activity.

The method is evaluated stage by stage on a dataset of 90 patients. This dataset is divided into four sub-datasets (*I\_a*, *I\_b*, *2\_a*, *2\_b*), two for each classification problem, and the obtained Acc is 82, 76, 91, 78%, respectively. The results indicate that the information regarding medication adherence affects the prediction of global adherence (case *I\_a*) and vice versa (case *2\_a*).

The comparison of the proposed method with the one reported in the literature (Son *et al.* [17]) is presented in Table 6. The current study provides equal or greater results in terms of Acc, PPV, Sens for three out of four cases that are examined, while Spec of the proposed method is lower compared with the one reported by Son *et al.* [17].

The combination of different classifiers with different FS measures revealed that features expressing medical condition of the patient, as well as, the medication treatment can act as predictors for the adherence. It must be mentioned that the features incorporated in the current study do not express all dimensions of adherence, since information regarding the socioeconomic status of the patient, the health care system and the knowledge of the patient about the disease is not included. This will be addressed in the prospective phase of data collection that will be performed within the HEARTEN project, where data expressing all the factors that affect adherence will be collected. The incorporation of the new information will definitely affect the set of adherence predictors, as well as, the classification Acc of the proposed method. Furthermore, the optimisation of the parameters is expected to improve the performance of the proposed method.

**5. Conclusion:** Medication adherence and global adherence (medication, nutrition and physical activity) of HF patients are both addressed in the current study, through the utilisation of machine learning techniques. The prediction of the adherence of the HF patients with satisfactory Acc indicates that the proposed method can contribute to the optimisation of HF patient management, since it will permit all involved actors to propose a therapeutic regimen and a monitoring plan that will improve the adherence of their patients.

**6. Funding and Declaration of Interests:** This work was supported by the HEARTEN project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 643694. Conflict of interest: None declared.

## 7 References

- [1] Cowie M.R.: 'The heart failure epidemic' (Medicographia, Les Laboratoires Servier France, 2012)
- [2] Desai A.S., Stevenson L.W.: 'Rehospitalization for heart failure: predict or prevent?', *Circulation*, 2012, **126**, (4), pp. 501–506
- [3] Riegel B., Knaf G.J.: 'Electronically monitored medication adherence predicts hospitalization in heart failure patients', *Patient Prefer Adherence*, 2013, **8**, pp. 1–13
- [4] Wu J.-R., Moser D.K., Lennie T.A., *ET AL.*: 'Medication adherence in patients who have heart failure: a review of the literature', *Nurs. Clin. North Am.*, 2008, **43**, (1), pp. 133–153o; vii–viii
- [5] Oosterom-Calo R., van Ballegooijen A.J., Terwee C.B., *ET AL.*: 'Determinants of adherence to heart failure medication: a systematic literature review', *Heart Fail. Rev.*, 2013, **18**, (4), pp. 409–427
- [6] Verbrugge M., Verhaeghe S., Lauwaert K., *ET AL.*: 'Determinants and associated factors influencing medication adherence and persistence to oral anticancer drugs: a systematic review', *Cancer Treat. Rev.*, 2013, **39**, (6), pp. 610–621
- [7] Pasma A., van't Spijker A., Hazes J.M.W., *ET AL.*: 'Factors associated with adherence to pharmaceutical treatment for rheumatoid arthritis patients: a systematic review', *Semin. Arthritis Rheum.*, 2013, **43**, (1), pp. 18–28

- [8] Daley D.J., Myint P.K., Gray R.J., *ET AL.*: 'Systematic review on factors associated with medication non-adherence in Parkinson's disease', *Parkinsonism Relat. Disord.*, 2012, **18**, (10), pp. 1053–1061
- [9] Jimmy B., Jose J.: 'Patient medication adherence: measures in daily practice', *Oman Med. J.*, 2011, **26**, (3), pp. 155–159
- [10] Farmer K.C.: 'Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice', *Clin. Ther.*, 1999, **21**, (6), pp. 1074–1090; discussion 1073
- [11] Knafel G.J., Riegel B.: 'What puts heart failure patients at risk for poor medication adherence?', *Patient Prefer Adherence*, 2014, **8**, pp. 1007–1018
- [12] Hajduk A.M., Lemon S.C., McManus D.D., *ET AL.*: 'Cognitive impairment and self-care in heart failure', *Clin. Epidemiol.*, 2013, **5**, pp. 407–416
- [13] Mathes T., Jaschinski T., Pieper D.: 'Adherence influencing factors – a systematic review of systematic reviews', *Arch. Public Health*, 2014, **72**, (1), p. 37
- [14] Dickson V.V., Knafel G.J., Riegel B.: 'Predictors of medication non-adherence differ among black and white patients with heart failure', *Res. Nurs. Health*, 2015, **38**, (4), pp. 289–300
- [15] Aggarwal B., Pender A., Mosca L., *ET AL.*: 'Factors associated with medication adherence among heart failure patients and their caregivers', *J. Nurs. Educ. Pract.*, 2015, **5**, (3), pp. 22–27
- [16] Juarez D.T., Williams A.E., Chen C., *ET AL.*: 'Factors affecting medication adherence trajectories for patients with heart failure', *Am. J. Manage. Care*, 2015, **21**, (3), pp. e197–e205
- [17] Son Y.-J., Kim H.-G., Kim E.-H., *ET AL.*: 'Application of support vector machine for prediction of medication adherence in heart failure patients', *Healthc. Inform. Res.*, 2010, **16**, (4), pp. 253–259
- [18] 'HEARTEN: 'A co-operative mHealth environment targeting adherence and management of patients suffering from heart failure'. Available at <http://www.hearten.eu/>
- [19] Blum A.L., Langley P.: 'Selection of relevant features and examples in machine learning', *Artif. Intell.*, 1997, **97**, (1–2), pp. 245–271
- [20] Kohavi R.: 'Wrappers for performance enhancement and oblivious decision graphs' (Stanford University, 1995)
- [21] Hall M.: 'Correlation-based feature selection for machine learning' Department of Computer Science, University of Waikato
- [22] Tang J., Alelyani S., Liu H.: 'Feature selection for classification: A review'. Available at [http://www.jiliang.xyz/publication/feature\\_selection\\_for\\_classification.pdf](http://www.jiliang.xyz/publication/feature_selection_for_classification.pdf). accessed 21 March 2016
- [23] Yu L., Liu H.: 'Feature selection for high-dimensional data: a fast correlation-based filter solution – ICML03–111.pdf'. Proc. 20th Int. Conf. on Machine Learning, CA, USA, 2003
- [24] Kononenko I.: 'Estimating attributes: analysis and extensions of RELIEF'. in Bergadano F., Raedt L.D. (Eds.): 'Machine learning: ECML-94' (Springer, Berlin Heidelberg, 1994), pp. 171–182
- [25] Breiman L.: 'Random forests', *Mach. Learn.*, 2001, **45**, pp. 5–32
- [26] Landwehr N., Hall M., Frank E.: 'Logistic model trees', *Mach. Learn.*, 2001, **59**, pp. 161–205
- [27] Sumner M., Frank E., Hall M.: 'Speeding up logistic model tree induction'. in Jorge A.M., Torgo L., Brazdil P., Camacho R., Gama J., (Eds.): 'Knowledge discovery in databases: PKDD 2005' (Springer, Berlin Heidelberg, 2005), pp. 675–683
- [28] Salzberg S.L.: 'C4.5: programs for machine learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993', *Mach. Learn.*, 1994, **16**, (3), pp. 235–240
- [29] Rodríguez J.J., Kuncheva L.I., Alonso C.J.: 'Rotation forest: a new classifier ensemble method', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2006, **28**, (10), pp. 1619–1630
- [30] Tripoliti E.E., Fotiadis D.I., Argyropoulou M., *ET AL.*: 'A six stage approach for the diagnosis of the Alzheimer's disease based on fMRI data', *J. Biomed. Inf.*, 2010, **43**, (2), pp. 307–320
- [31] Cristianini J., Shawe-Taylor: 'An introduction to support vector machines and other Kernel-based learning methods, pattern recognition and machine learning' (Cambridge University Press, USA, 2000)
- [32] Cortes C., Vapnik V.: 'Support-vector networks', *Mach. Learn.*, 1995, **20**, (3), pp. 273–297
- [33] Frank E.: 'Fully supervised training of Gaussian radial basis function networks in WEKA'. Available at [http://www.cs.waikato.ac.nz/~ml/publications/2014/rbf\\_networks\\_in\\_weka\\_description.pdf](http://www.cs.waikato.ac.nz/~ml/publications/2014/rbf_networks_in_weka_description.pdf). accessed 05 May 2016
- [34] Murphy K.P.: 'Dynamic Bayesian networks: representation, inference and learning', PhD dissertation, University of California, Berkeley, Citeseer, 2002
- [35] Mitchell T.: 'Machine learning' (McGraw-Hill Science/Engineering/Math, Springer, 1997)
- [36] Tripoliti E.E., Tzallas A.T., Tsiouras M.G., *ET AL.*: 'Automatic detection of freezing of gait events in patients with Parkinson's disease', *Comput. Methods Programs Biomed.*, 2013, **110**, (1), pp. 12–26
- [37] Theodoridis S., Koutroumbas K.: 'Pattern recognition' (Elsevier, Academic Press, 2009)
- [38] Hassoun M.: 'Fundamentals of artificial neural networks' (MIT Press, 2010)
- [39] Breiman L., Friedman J.H., Olshen R.A., *ET AL.*: 'Classification and regression trees'. Wadsworth Int. Group, Belmont CA, Wadsworth Int. Group, Belmont, CA, 1984, **1**
- [40] Chawla N., Bowyer K., Hall L.O., *ET AL.*: 'SMOTE: synthetic minority over-sampling technique', *J. Artif. Intell. Res.*, 2002, **16**, pp. 321–357