

Estimation of heart failure patients medication adherence through the utilization of saliva and breath biomarkers and data mining techniques

Evanthia E. Tripoliti, Theofilos G. Papadopoulou, Georgia S. Karanasiou, Fanis G. Kalatzis, Dimitrios I. Fotiadis
Department of Biomedical Research
Institute of Molecular Biology and Biotechnology
Foundation for Research and Technology-Hellas
Ioannina, Greece
etripoliti@gmail.com,
tpapado2011@gmail.com,
g.karanasiou@gmail.com,
tkalatz@gmail.com, fotiadis@cc.uoi.gr

Yorgos Goletsis
Department of Economics
University of Ioannina
Ioannina, Greece
goletsis@cc.uoi.gr

Katerina K. Naka, Aris Bechlioulis
Michaelidion Cardiac Center,
2nd Department of Cardiology,
University of Ioannina
Ioannina, Greece
drkknaka@gmail.com,
md02798@yahoo.gr

Silvia Ghimenti, Tommaso Lomonaco, Francesca Bellagambi, Roger Fuoco
Department of Chemistry and Industrial Chemistry, University of Pisa
Pisa, Italy
silvia.ghimenti@ifc.cnr.it,
tlomonaco@ifc.cnr.it,
fbellagambi@ifc.cnr.it, fuoco@dcci.unipi.it

Mario Marzilli, Maria Chiara Scali
Azienda Ospedaliera-Universitaria
Pisana
Cardiothoracic and Vascular
Department Pisa, Italy
mario.marzilli@med.unipi.it,
chiara_scali@yahoo.it

Abdelhamid Errachid
Université de Lyon
Institut de Sciences Analytiques
(ISA) – UMR
Villeurbanne France
abdelhamid.errachid@univ-lyon1.fr

Abstract—The aim of this work is to estimate the medication adherence of patients with heart failure through the application of a data mining approach on a dataset including information from saliva and breath biomarkers. The method consists of two stages. In the first stage, a model for the estimation of adherence risk of a patient, exploiting anamnestic and instrumental data, is applied. In the second stage, the output of the model, accompanied with data from saliva and breath biomarkers, is given as input to a classification model for determining if the patient is adherent, in terms of medication. The method is evaluated on a dataset of 29 patients and the achieved accuracy is 96%.

Keywords—medication adherence; saliva biomarkers; breath biomarkers; data mining; heart failure

I. INTRODUCTION

Heart failure (HF) is a common diagnosed chronic disease among individuals 65 year and older. It is characterized by recurrent hospitalizations, high mortality, poor quality of life and increased healthcare costs. A variety of reasons contribute to the high prevalence of this disease. Non-adherence of HF patients to treatment suggestions (suggestions regarding medication, nutrition and physical activity exercising) has been proven a significant contributor to the presence of HF adverse events (destabilizations, re-hospitalizations, mortality) [1-4].

Direct and indirect methods have been implemented for the estimation of medication adherence, with the Medication Event Monitoring System (MEMS) and measurement of drug levels in blood and urine to be the golden standard of indirect and direct methods, respectively [5]. Estimation of medication adherence has also gained the interest of researchers who performed studies in order to identify the modifiable factors associated with medication adherence and developed models for predicting adherence in adults with HF. The prediction models reported in the literature exploit information concerning sociodemographic characteristics of the patients, frequency of medication intake, medical condition, biological data and results of clinical examinations [3, 6-13].

The last years' studies have revealed the strong correlation of saliva and breath biomarkers with HF severity, progression and mortality through statistical analysis methods [14-28]. Furthermore, in [29] a computational method for the estimation of HF patient status, using saliva biomarkers, is presented.

The current work, taking into account the valuable information that saliva and breath biomarkers carry regarding HF patient status and the dependence of HF patient status on treatment adherence, aims to examine and identify the potentially significant correlates of HF patient medication adherence to HF-related saliva and breath

biomarkers. In order this to be achieved a two stage model is built by employing data mining techniques and utilizing saliva and breath biomarkers data in combination with sociodemographic, medical and clinical information. More specifically, in the first stage, the estimation of adherence risk of HF patients in terms of medication and overall adherence (medication, nutrition and physical activity exercise) is applied, while in the second stage, the HF patient is classified as medication adherent or not by combining the output of the first stage with saliva biomarkers data, breath biomarkers data, New York Heart Association (NYHA) class and the status (acute, progressive, stable) of the HF patient.

The NYHA class, and the status of the patient, at the current phase, are provided by the experts. However, in the future they will be derived from the NYHA class detection module and Event Prediction Module, respectively, of the HEARTEN Knowledge Management System (KMS) developed within the HEARTEN project [30].

The HEARTEN project [30] creates an mHealth ecosystem for empowering HF patients, optimizing disease management and improving patient adherence. The last is accomplished by two modules of the HEARTEN KMS, Adherence risk module and Treatment adherence module. The main innovation of HEARTEN, from the clinical point of view, is the integration of different data deriving from a variety of sensors, as well as the development of non-invasive breath/saliva biosensors. The correlation of these different sensor/biosensor outputs has not previously been produced and the extraction of meaningful knowledge through an automated, quick and reliable process is currently lacking.

II. MATERIALS AND METHODS

A. Dataset

The proposed method is evaluated using a dataset of 29 patients collected by the clinical center of the Università Di Pisa (UNIP), Italy within the HEARTEN project [30]. The dataset consists of patients: (i) diagnosed with HF (Framingham criteria) who have continuous symptoms with frequent recurrence, (ii) belonging to the functional NYHA I-IV class followed by an optimal treatment, (iii) who have been recently hospitalized, (at least one in the last six months), (iv) who have undergone one electrocardiogram (in the last 12 months) and have HF symptoms. Patients who are underage, with very severe HF, with obesity and advanced chronic kidney failure are not included.

The features recorded for each patient can be grouped into the following categories: (i) General Information, (ii) Allergies, (iii) Medical Condition, (iv) Drugs, (v) Biological data related to HF disease, (vi) Clinical Examinations, (vii) Adherence, (viii) Biomarkers. Uric Acid, Tumor Necrosis Factor α , Cortisol and 8-iso-prostaglandin F2a, Isoprene and Acetone are measured.

These features are recorded from the first time of patient's hospitalization (*Hosp*) until discharge (*Dis*) every

second day. Thus, a set of 57 instances are collected (21 instances at *Dis* phase and 8 instances at *Hosp* and *Dis* phase). Each instance is characterized by the experts as high, medium, low adherent. The dataset includes: 42 instances characterized as high adherent and 15 instances characterized as medium adherent. Taking into consideration that the patients are in the hospital, the case of low adherence is not included. Since the aim of the study is estimation of medication adherence, a control group is not included.

B. The proposed method

The proposed method consists of two stages: i) estimation of adherence risk of the patient, ii) classification of the patient as medication adherent or not. A schematic representation of the proposed method is shown in Fig. 1 and a detailed description of each stage is provided below.

Stage 1 – Estimation of adherence risk: Estimation of medication and overall adherence risk is performed. The model presented in [13] is applied. The model takes as input features belonging to groups (i)-(vii) (Section II-A) and applies feature selection and classification techniques (Fig. 2). The Stage 1 gives two outputs. The first output is, if the patient is medication adherent or not and the second, if the patient is overall adherent or not.

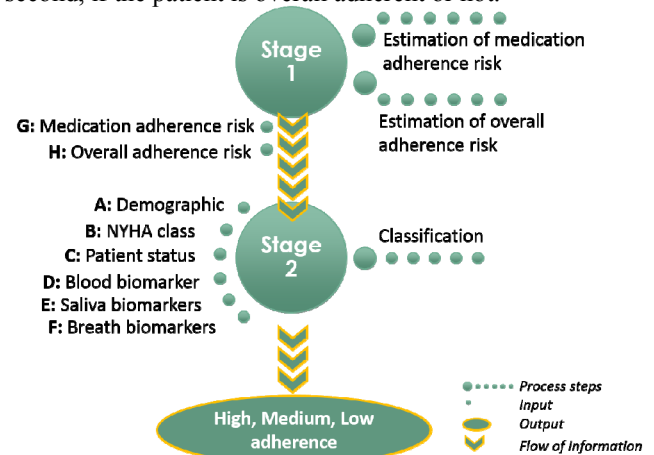


Figure 1. Flowchart of the proposed method.

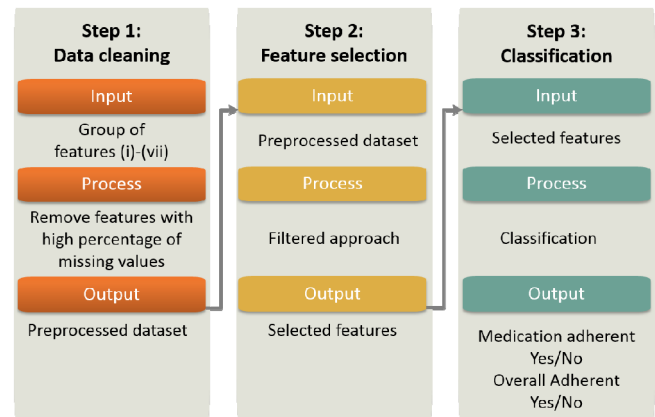


Figure 2. Stage 1 of the proposed method.

TABLE I: ABBREVIATIONS FOR THE TYPE OF INPUT OF STAGE 2 OF THE PROPOSED METHOD.

Abbreviation	Type of input
A	Demographic (Age, Gender)
B	NYHA class
C	Patient status
D	Blood biomarker (NT-proBNP)*
E	Saliva biomarkers (UA, TNF-a, 8-iso-F2a)*
F	Breath biomarkers (Isoprene, Acetone)
G	Medication adherence risk
H	Overall adherence risk

* NT-proBNP: N-terminal pro b-type natriuretic peptide, UA: Uric Acid, TNF-a: Tumor Necrosis Factor-a, 8-iso-F2a: 8-iso-prostaglandin F2a

Stage 2- Classification: It takes as input the output of the Stage 1, as well as demographic information (gender and age), NYHA class of the patient, patient status (acute, progressive, stable), blood biomarkers (N-terminal pro b-type natriuretic peptide), saliva biomarkers (Uric Acid, Tumor Necrosis Factor –a, Cortisol and 8-iso-prostaglandin F2a) and breath biomarkers (Isoprene, Acetone). An abbreviation is assigned to each of the above mentioned inputs (presented in Table I) that will be used throughout the rest of this paper. The number of instances (57) remains constant for all the datasets, while the number of features is differentiated depending on the type of input that is removed.

For the classification nine classifiers are employed. Five of them belong to the category of tree-based classifiers (Random Forests-RF [31], Logistic Model Trees-LMT [32], J48 [33], Simple Classification and regression tree-CART [34], Rotation Forest-ROT [35]), two of them to the category of kernel-based classifiers (Radial Basis Function Network-RBF [36], Support Vector Machine-SVM [37]) and two of them to the category of Bayesian classifier (Bayesian Network-BN [38] and Naive Bayes-NB [39]). The most representative classifiers from each category are selected.

III. RESULTS

The proposed method is applied for each of the datasets presented in Table II. The difference between the datasets is the combination of features that compose each dataset. The results of the proposed method in terms of accuracy, as well as the classifier that provides the best results are presented in Table III.

In order to examine if the participation of medication adherence risk (**G**) affects the performance of the proposed method the type of input **G** is removed from the datasets and the experiments are repeated (Table IV).

Taking into account that the characterization of patient status as acute, progressive and stable is annotated by the clinical experts without the utilization of an established scale, a fact that may introduce subjectivity to the datasets and influence the results of the proposed method, the specific feature (type of input **C**) is removed and the Stage 2 of the proposed method is reapplied. The results are presented in Table V.

In addition, both types of input, **C** and **G**, are removed and the second stage of the proposed method is evaluated again. Table VI presents the results.

In order to examine the performance of the proposed method, when only the saliva or breath biomarkers are available two versions of the Dataset #1 are created. The first version includes only saliva biomarkers, while the second includes only breath biomarkers. The accuracy for the two cases is presented in Table VII.

TABLE II: DESCRIPTION OF THE DATASET.

Dataset	A	B	C	D	E	F	G	H
#1	✓	✓	✓	✓	✓	✓	✓	✓
#2		✓	✓	✓	✓	✓	✓	✓
#3	✓		✓	✓	✓	✓	✓	✓
#4	✓	✓	✓		✓	✓	✓	✓
#5			✓	✓	✓	✓	✓	✓
#6		✓	✓		✓	✓	✓	✓
#7			✓		✓	✓	✓	✓
#8			✓		✓	✓		

Grey cells represent the type of input that is removed from each dataset

TABLE III: RESULTS OF THE PROPOSED METHOD.

Datasets							
#1	#2	#3	#4	#5	#6	#7	#8
Accuracy							
95%	91%	93%	91%	88%	89%	88%	86%
Classifier							
RF	BN ROT	BN	RF	BN ROT LMT RF	ROT	ROT	RF CART

TABLE IV: RESULTS OF THE PROPOSED METHOD BY REMOVING TYPE OF INPUT G FROM THE DATASETS DESCRIBED IN TABLE II.

Datasets							
#1	#2	#3	#4	#5	#6	#7	#8
Accuracy							
96%	89%	95%	93%	91%	89%	88%	86%
Classifier							
ROT	RF	BN ROT	ROT RF CART	RF	LMT	RF	RF CART

TABLE V: RESULTS OF THE PROPOSED METHOD BY REMOVING TYPE OF INPUT C FROM THE DATASETS DESCRIBED IN TABLE II.

Datasets							
#1	#2	#3	#4	#5	#6	#7	#8
Accuracy							
95%	91%	96%	95%	89%	86%	88%	86%
Classifier							
BN ROT RF	ROT	RF	RF	RBF LMT RF	RF	RBF RF	RF CART

TABLE VI: RESULTS OF THE PROPOSED METHOD BY REMOVING TYPE OF INPUT C AND G FROM THE DATASETS DESCRIBED IN TABLE II.

Datasets							
#1	#2	#3	#4	#5	#6	#7	#8
Accuracy							
96%	89%	95%	96%	89%	88%	86%	86%
Classifier							
RF	ROT	BN LMT	RF	BN	RBF	RBF RF CART	RF CART

TABLE VII: RESULTS OF THE PROPOSED METHOD WHEN ONLY ONE TYPE OF BIOMARKER IS INCLUDED IN THE DATASET.

Dataset	Accuracy	Classifier
#1 with saliva biomarkers	95%	ROT
#1 with breath biomarkers	96%	ROT

IV. DISCUSSION

An automated computational approach for the estimation of medication adherence of the HF patients is presented based on data expressing saliva and breath biomarkers, a feature that differentiate the proposed method from other methods reported in the literature. The proposed method consists of two stages. In the first stage, a model for the estimation of medication and overall adherence risk is applied, while in the second stage a classification model to classify a patient as medication adherent or not is built. The proposed method is applied to eight datasets described in Table II. The best results (95% accuracy) are obtained when all types of input (A to H) participate in the dataset and the RF classifier is applied. The results presented in Table III also indicate that the removal of information expressing NYHA class (Dataset #3) and the removal of blood biomarker NT-proBNP (Dataset #4) decrease slightly, 2% and 4% percent, respectively, the performance of the proposed method.

Additionally, further experiments are presented to evaluate the performance of the proposed method and more specifically of the second stage of the method. The difference between the experiments is the dataset that is given as input to the Stage 2. More specifically, the removal of the type of input **G** from the datasets affects either positively or negatively (Table IV) the results of the proposed method. The small differences (maximum increase 3%, maximum decrease 2%) in five out of eight cases, presented in Table III and Table IV, indicate that the contribution of medication adherence risk depends on the dataset and the classifier that is utilized.

The contribution of **C** type of input can be estimated from the results presented in Table V. The results indicate that the status of patient, as it is estimated by clinical experts, can contribute to the classification of patient as medication adherent or not when information regarding the NYHA class or NT-proBNP of the patient are not available. This specific observation will be further

validated in the future where the patient status will be provided automatically by the Event prediction module of the HEARTEN KMS through the application of a computational approach. Some preliminary results presented in [29] show that the discrimination of acute versus non acute status can be achieved with 85% accuracy, while the discrimination of stable versus progressive status can be achieved with 69% accuracy.

The utilization of breath and saliva biomarkers as the only input of the proposed method (Table V – Dataset #8) results to an estimation of medication adherence with 86% accuracy. The presence in the dataset of only one type of biomarker results to 95 (using only saliva biomarkers) and 96% (using only breath biomarkers) accuracy.

The current study examines the predictive power of saliva and breath biomarkers toward medication adherence taking into consideration the fact that biomarkers significantly correlate with HF severity and progression and provide valuable information about the status of the patient, which strongly depends on the adherence of the patient to the experts' treatment suggestions. The results verify the existence of correlation between the biomarkers and adherence, however observations regarding which of the saliva and breath biomarkers are more indicative will be extracted in the future, when larger datasets will be available. The results of the study can be used as a basis for the development of low cost, easy to use, non-invasive breath and saliva biosensors. The communication of the biosensors with a mHealth application will facilitate the increase and improvement of the medication adherence.

V. CONCLUSIONS

It is widely accepted that existing medications for HF can provide great health benefits, if the treatment by the medical experts' suggestions are followed. Specifically, HF symptoms, emergency department visits and hospitalizations can be reduced and survival can be increased. However, nowadays, medication adherence remains a challenge for patients with HF resulting not only in health-related adverse events but also in high healthcare costs.

The current study addresses the problem of medication adherence of HF patients through the exploitation of saliva and breath biomarkers in combination with demographic information, NYHA class and status of the patient, blood biomarker and medication and overall adherence risk extracted by a model taking into consideration sociodemographic, medical and clinical information.

It is expected that this method can be fully integrated and employed in a mHealth application allowing the estimation of the patient's medication adherence on time, enabling in this way the provision of personalized advices and suggestions, which in turn will result in avoiding adverse events and improving the patient's health status and overall quality of life.

ACKNOWLEDGMENTS

This work is supported by the HEARTEN project that has received funding from the European Union's Horizon 2020

research and innovation program under grant agreement No 643694.

REFERENCES

- [1] M. R. Cowie, "The heart failure epidemic," *Medicographia*, 2012.
- [2] A. S. Desai and L. W. Stevenson, "Rehospitalization for heart failure: predict or prevent?," *Circulation*, vol. 126, no. 4, pp. 501–506, Jul. 2012.
- [3] B. Riegel and G. J. Knaf, "Electronically monitored medication adherence predicts hospitalization in heart failure patients," *Patient Prefer Adherence*, vol. 8, pp. 1–13, Dec. 2013.
- [4] J.-R. Wu, D. K. Moser, T. A. Lennie, and P. V. Burkhart, "Medication adherence in patients who have heart failure: a review of the literature," *Nurs. Clin. North Am.*, vol. 43, no. 1, p. 133–153; vii–viii, Mar. 2008.
- [5] K. C. Farmer, "Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice," *Clin Ther*, vol. 21, no. 6, p. 1074–1090; discussion 1073, Jun. 1999.
- [6] G. J. Knaf and B. Riegel, "What puts heart failure patients at risk for poor medication adherence?," *Patient Prefer Adherence*, vol. 8, pp. 1007–1018, 2014.
- [7] A. M. Hajduk, S. C. Lemon, D. D. McManus, D. M. Lessard, J. H. Gurwitz, F. A. Spencer, R. J. Goldberg, and J. S. Saczynski, "Cognitive impairment and self-care in heart failure," *Clin Epidemiol*, vol. 5, pp. 407–416, 2013.
- [8] T. Mathes, T. Jaschinski, and D. Pieper, "Adherence influencing factors – a systematic review of systematic reviews," *Arch Public Health*, vol. 72, no. 1, Oct. 2014.
- [9] V. V. Dickson, G. J. Knaf, and B. Riegel, "Predictors of medication nonadherence differ among black and white patients with heart failure," *Res Nurs Health*, vol. 38, no. 4, pp. 289–300, Aug. 2015.
- [10] B. Aggarwal, A. Pender, L. Mosca, and H. Mochari-Greenberger, "Factors associated with medication adherence among heart failure patients and their caregivers," *J Nurs Educ Pract*, vol. 5, no. 3, pp. 22–27, 2015.
- [11] D. T. Juarez, A. E. Williams, C. Chen, Y. G. Daida, S. K. Tanaka, C. M. Trinacty, and T. M. Vogt, "Factors affecting medication adherence trajectories for patients with heart failure," *Am J Manag Care*, vol. 21, no. 3, pp. e197–205, Mar. 2015.
- [12] Y.-J. Son, H.-G. Kim, E.-H. Kim, S. Choi, and S.-K. Lee, "Application of support vector machine for prediction of medication adherence in heart failure patients," *Healthc Inform Res*, vol. 16, no. 4, pp. 253–259, Dec. 2010.
- [13] G. S. Karanasiou, E. E. Tripoliti, T. G. Papadopoulos, F. G. Kalatzis, Y. Goletsis, K. K. Naka, A. Bechlioulis, A. Errachid and D. I. Fotiadis, "Predicting adherence of patients with heart failure through machine learning techniques," *Healthc Technol Lett.*, vol. 3(3), pp.165-170, 2016.
- [14] W. Doehner and S. D. Anker, "Uric acid in chronic heart failure," *Semin. Nephrol.*, vol. 25, no. 1, pp. 61–66, Jan. 2005.
- [15] D. L. Mann, "Inflammatory Mediators and the Failing Heart," *Circulation Research*, vol. 91, no. 11, pp. 988–998, Nov. 2002.
- [16] L. I. B. Sikkeland et al., "Increased Levels of Inflammatory Cytokines and Endothelin-1 in Alveolar Macrophages from Patients with Chronic Heart Failure," *PLOS ONE*, vol. 7, no. 5, p. e36815, 2012.
- [17] H. K. Gaggin and J. L. Januzzi, "Biomarkers and diagnostics in heart failure," *Biochim. Biophys. Acta*, vol. 1832, no. 12, pp. 2442–2450, Dec. 2013.
- [18] N. Ansari, A. Hasan, and M. Owais, "A study of inflammatory markers and their correlation with severity, in patients with chronic heart failure," *Biomedical Research*, vol. 23, pp. 408–415, 2012.
- [19] A. Suska, U. Alehagen, I. Lundstrom, and U. Dahlstrom, "Salivary Alpha-Amylase Activity, a New Biomarker in Heart Failure?," *Journal of Clinical & Experimental Cardiology*, Jun. 2012.
- [20] U. M. Nater and N. Rohleder, "Salivary alpha-amylase as a non-invasive biomarker for the sympathetic nervous system: current state of research," *Psychoneuroendocrinology*, vol. 34, no. 4, pp. 486–496, May 2009.
- [21] M. Yamaji et al., "Serum cortisol as a useful predictor of cardiac events in patients with chronic heart failure: the impact of oxidative stress," *Circ Heart Fail*, vol. 2, no. 6, pp. 608–615, Nov. 2009.
- [22] D. B. Sawyer, "Oxidative Stress in Heart Failure: What are we missing?," *Am J Med Sci*, vol. 342, no. 2, pp. 120–124, Aug. 2011.
- [23] K. Nakamura et al., "Beta-Blockers and Oxidative Stress in Patients with Heart Failure," *Pharmaceuticals (Basel)*, vol. 4, no. 8, pp. 1088–1100, Aug. 2011.
- [24] B. Pypers and J. T. Pierce, "Lactate physiology in health and disease," *Contin Educ Anaesth Crit Care Pain*, vol. 6, no. 3, pp. 128–132, Jun. 2006.
- [25] C. Lazzeri, S. Valente, M. Chiostrri, and G. F. Gensini, "Clinical significance of lactate in acute cardiac patients," *World J Cardiol*, vol. 7, no. 8, pp. 483–489, Aug. 2015.
- [26] T. Doenst, T. D. Nguyen, and E. D. Abel, "Cardiac Metabolism in Heart Failure - Implications beyond ATP production," *Circ Res*, vol. 113, no. 6, pp. 709–724, Aug. 2013.
- [27] L. T. McGrath, R. Patrick, B. Silke, "Breath isoprene in patients with heart failure," *Eur J Heart Fail* vol. 3, pp. 423-427, 2001.
- [28] F. G. Marcondes-Braga, I. G. Gutz, G. L. Batista, P. H. Saldiva, S. M. Ayub-Ferreira, et al., "Exhaled acetone as a new biomarker of heart failure severity," *Chest* vol. 142, pp. 457-466, 2012.
- [29] E. E. Tripoliti, T. G. Papadopoulos, G. S. Karanasiou, F. G. Kalatzis, Y. Goletsis, S. Ghimenti, T. Lomonaco, F. Bellagambi, M. G. Trivella, R. Fuoco, M. Marzilli, M. C. Scali, A. Errachid, D. I. Fotiadis, "A Computational Approach for the Estimation of Heart Failure Patients Status Using Saliva Biomarkers", submitted to 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, July 2017, Korea.
- [30] "HEARTEN: A co-operative mHealth environment targeting adherence and management of patients suffering from Heart Failure." Available: <http://www.hearten.eu/>.
- [31] L. Breiman, "Random Forests," *Machine Learning*, 1997, vol. 45.
- [32] N. Landwehr, M. Hall, and E. Frank, "Logistic Model Trees," *Machine Learning*, 2005.
- [33] S. L. Salzberg, "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993," *Mach Learn*, vol. 16, no. 3, pp. 235–240, Sep. 1994.
- [34] L. Breiman, J. H. Friedman, R. A. Olshcn, and C. J. Stone, "Classification and Regression Trees. Wadsworth International Group, Belmont CA," Wadsworth International Group, Belmont, California., vol. 1, Jan. 1984.
- [35] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, "Rotation forest: A new classifier ensemble method," *IEEE Trans Pattern Anal Mach Intell*, vol. 28, no. 10, pp. 1619–1630, Oct. 2006.
- [36] J. Cristianini and Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Pattern Recognition and Machine Learning," USA, Cambridge University Press, 2000.
- [37] E. Frank, "Fully Supervised Training of Gaussian Radial Basis Function Networks in WEKA." [Online]. Available: http://www.cs.waikato.ac.nz/~ml/publications/2014/rbf_networks_in_weka_description.pdf. [Accessed: 05-May-2016].
- [38] K. P. Murphy, "Dynamic bayesian networks: representation, inference and learning," PhD dissertation, University of California, Berkeley, Citeseer, 2002.
- [39] T. Mitchell, "Machine Learning," in McGraw-Hill Science/Engineering/Math, 1997.