Article

An Associative Memory Approach to Healthcare Monitoring and Decision Making

Mario Aldape-Pérez [1,*] 📷, **Antonio Alarcón-Paredes** [2] 📷, **Cornelio Yáñez-Márquez** [3] 📷, **Itzamá López-Yáñez** [1] 📷 **and Oscar Camacho-Nieto** [1] 📷

[1] Instituto Politécnico Nacional, Computational Intelligence Laboratory at CIDETEC, Ciudad de Mexico 07700, Mexico; ilopezy@ipn.mx (I.L.-Y.); ocamacho@ipn.mx (O.C.-N.)
[2] Universidad Autónoma de Guerrero, Engineering Department, Guerrero 39079, Mexico; aalarcon@uagro.mx
[3] Instituto Politécnico Nacional, Computational Intelligence Laboratory at CIC, Ciudad de Mexico 07738, Mexico; cyanez@cic.ipn.mx
***** Correspondence: maldape@ipn.mx; Tel.: +52-555-729-6000

check for updates

Abstract: The rapid proliferation of connectivity, availability of ubiquitous computing, miniaturization of sensors and communication technology, have changed healthcare in all its areas, creating the well-known healthcare paradigm of e-Health. In this paper, an embedded system capable of monitoring, learning and classifying biometric signals is presented. The machine learning model is based on associative memories to predict the presence or absence of coronary artery disease in patients. Classification accuracy, sensitivity and specificity results show that the performance of our proposal exceeds the performance achieved by each of the fifty widely known algorithms against which it was compared.

Keywords: associative memories; decision support systems; e-Health; Internet of Things; pattern classification

1. Introduction

In a large number of countries, healthcare systems and services have become an essential human right; consequently, individuals' access to the public health systems has become an indicator of the well-being and the development of nations [1,2]. Public health systems were designed to attend and provide health services to specific individuals. However, the constant growth of population and the increasing costs of health services imply that the public health system will face new challenges [3]. Some important aspects to consider are to develop and evaluate innovative approaches for improving the quality of healthcare using sensors applications in medical monitoring. Two decades ago, technology researchers took an important role towards improving the medical care of patients through the evolution of the concept of a network of smart devices, which would be known as Wireless Sensor Networks (WSNs) [4]. In the same decade, the concept of moving small amounts of data to a large set of nodes evolved to what today is known as the Internet of Things (IoT) [5,6]. The IoT paradigm represents one of the most disruptive technologies, enabling ubiquitous computing scenarios for medical monitoring, and decision making [7,8]; creating the well-known healthcare paradigm of e-Health [9]. This paradigm arises as a result of the combination of emerging technologies, such as IoT, ubiquitous computing, WSNs, high-speed communications infrastructure, and the social need for more effective health services with better accessibility and availability [10]. The e-Health paradigm has changed the traditional way in which healthcare services are provided. With this paradigm the user of healthcare services does not need to move to medical facilities to carry out a routine follow-up [11,12]; on the contrary, medical monitoring can be carried out from where the patient is located [13,14].

In addition, all acquired data can be transmitted, processed and stored for data mining and decision making by medical specialists [15,16]. Nowadays, it is increasingly common to use applications based on artificial intelligence techniques to support medical specialists in decision making. For more than a decade, statistical techniques [17], expert systems [18], neural networks [19], decision trees [20] and associative memories [21,22] have been widely used for pattern recognition, feature selection, data mining and decision making in the medical field [23–25].

In this paper, an embedded system capable of monitoring, learning and classifying biometric signals is presented. The machine learning model is based on associative memories to predict the presence or absence of coronary artery disease in patients. Classification accuracy, sensitivity and specificity results show that the performance of our proposal exceeds the performance achieved by each of the fifty widely known algorithms against which it was compared.

The paper is organized as follows. Section 2 presents previous works related to machine learning based systems that have been applied to predict the presence or absence of coronary artery disease in patients. In Section 3, a succinct description of associative memories fundamentals is presented. In Section 4, an improvement to Delta Associative Memory original model, called IDAM, is proposed. Section 5 presents the three main performance indicators of a binary classification test. The experimental phase is described in Section 6. In Section 7, sensitivity, specificity and classification accuracy results achieved by each of the compared algorithms in two datasets related to coronary artery disease diagnosis, are presented. Finally, our proposal's advantages, as well as some conclusions are discussed in Section 8.

2. Previous Works

For more than a decade, machine learning based systems have been tested in patients with cardiac disease to predict outcome, or in the general population to detect cardiac diseases. In 2008, Kahramanli and Allahverdi [26] proposed a hybrid neural system for heart diseases that includes artificial neural network (ANN) and fuzzy neural network (FNN); the dataset was obtained from the University of California at Irvine (UCI) machine learning repository [27]. In 2009, Polat and Güneş [28] proposed a feature selection method on classification of medical datasets called Kernel F-score feature selection (KFFS); experimental results showed that KFSS achieved better results compared to F-score feature selection (FFS). In 2011, McSherry [29] proposed an approach to conversational case-based reasoning (CCBR) in medical classification and diagnosis that aims to increase transparency while also providing high levels of accuracy and efficiency; two datasets from the UCI machine learning repository were used in the experimental phase. In 2012, Aldape-Pérez et al. [23] proposed an associative memory approach to medical decision support systems. This work focuses on the use of classical associative memories for medical patterns classification. This approach incorporates a learning reinforcement stage, which increases the classification performance of classical models of associative memories. The performance was validated on medical datasets collected from the UCI machine learning repository. In 2012, Anooj [30] proposed a weighted fuzzy rule-based clinical decision support system (CDSS) for the diagnosis of coronary artery disease; the experimentation was carried out on the proposed system using the datasets obtained from the UCI machine learning repository and the performance of the system was compared with a neural network-based system utilizing accuracy, sensitivity and specificity. In 2013, Nahar et al. [31] proposed a computational intelligence approach for association rule mining to investigate the sick and healthy factors which contribute to coronary artery disease for males and females; the dataset that was used in the experimental phase was the UCI Cleveland dataset. In 2014, Biswas et al. [32] proposed a method to extract symbolic weights from a trained neural network by observing the whole trained neural network as an AND/OR graph and then finding a solution for each node that becomes the weight of a corresponding node. The performance was validated on coronary artery disease dataset collected from the UCI machine learning repository. In 2015, Aldape-Pérez et al. [24] proposed a collaborative learning approach based on associative models to pattern classification in medical datasets. In this work, Delta Associative Memory was

presented. The operation of this model is based on the differences that exist between patterns of different classes and a dynamic threshold that is calculated for each unknown pattern to be classified. The experimental results were competitive, when compared against algorithms in the current literature. In 2015, Nguyen et al. [33] proposed an integration of fuzzy standard additive model (SAM) with genetic algorithm (GA), called GSAM, to deal with uncertainty. The proposed method was evaluated using Cleveland coronary artery disease dataset from the UCI machine learning repository. In 2016, Leema et al. [34] proposed a Computer-Aided Diagnostic (CAD) system that uses an Artificial Neural Network (ANN) trained by Differential Evolution (DE), Particle Swarm Optimization (PSO) and gradient descent based backpropagation (BP) for classifying clinical datasets, obtained from the UCI machine learning repository. In 2016, Nahato et al. [35] proposed a classifier that combines the fuzzy sets and extreme learning machine (FELM) for clinical datasets. The three major subsystems in the FELM framework are preprocessing subsystem, fuzzification subsystem and classification subsystem. Missing value imputation and outlier elimination are handled by the preprocessing subsystem. Cleveland coronary artery disease dataset from the UCI machine learning repository was used for experimentation. In 2017, Ramírez-Rubio et al. [25] proposed an associative model called Normalized Difference Associative Memory. This associative model overcome the limitations of the original Alpha-Beta Associative Memories [36]. In 2017, Shah et al. [37] proposed a methodology which uses the results of medical tests as input, extracts a reduced dimensional feature subset and provides diagnosis of coronary artery disease. The proposed methodology extracts high impact features in a new projection by using Probabilistic Principal Component Analysis (PPCA). The feature subset with the reduced dimension is provided to radial basis function (RBF) kernel based Support Vector Machines (SVM). Methodology performance was evaluated through accuracy, specificity and sensitivity over the three datasets of the UCI machine learning repository.

3. Associative Memories

The first models of Associative Memories arise with the scientific findings of Steinbuch in the 1960s [38–40], which over time would be known as Learning Matrices. In any learning matrix, there are two phases that determine the performance of each model, namely learning phase and classification phase. Learning matrices are structures formed by rows and columns whose intersection points are formed by connecting elements [41]. The characteristics of an object are presented during the learning phase to the columns as binary signals via a suitable transducer. Simultaneously, a meaning of an object associated with this set of characteristics is applied in the form of a signal to one of the rows. Therefore, so-called conditioned connections are effected in the connective elements of the row selected by the meaning [42]. Generalizing, a conditioned connection is a functional connection between a row and a column. In this way, during the learning phase each input vector $\mathbf{x}^\mu \in A^n$, $A = \{0, 1\}$ (characteristics of an object) forms an association with its corresponding output vector $\mathbf{y}^\mu \in A^m$ (meaning of an object associated with this set of characteristics), so for each γ integer and positive, the corresponding association is denoted as: $(\mathbf{x}^\gamma, \mathbf{y}^\gamma)$. Thus, an associative memory $\mathbf{M}$ is generated from an a priori finite set of known associations, called the fundamental set of associations. If μ is an integer and positive value, the fundamental set is represented as: $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, ..., p\}$ with p as the cardinality of the set. A distorted version of a pattern $\mathbf{x}^\gamma$ to be recalled is denoted as $\tilde{\mathbf{x}}^\gamma$. An unknown input pattern to be recalled is denoted as $\mathbf{x}^\omega$. If when an unknown input pattern $\mathbf{x}^\omega$ is fed to an associative memory $\mathbf{M}$, it happens that the output corresponds exactly to the associated pattern $\mathbf{y}^\omega$, it is said that recalling is correct.

Associative memories have been widely used to perform pattern recognition tasks effectively, however, they present a limitation known as *cross-talk*. The influence of *cross-talk* causes the associative memory to become saturated and consequently the classification performance is negatively affected.

4. Our Proposal

Negative effects of *cross-talk* are due to an order relation between patterns that constitute the fundamental set $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, ..., p\}$ with p as the cardinality of the set [36]. To improve classification performance of Delta Associative Memory [24], as well as to eliminate the negative effects of *cross-talk*, an improvement to Delta Associative Memory model is proposed, called Improved Delta Associative Memory (IDAM). This modification consists of adding a data preprocessing stage before the Delta Associative Memory learning phase. This additional stage is based on information quality estimation concepts that were proposed by Aldape-Pérez et al. [23] to reinforce learning in an associative memory. In the present paper, those concepts are used but with very different purposes, namely: to obtain a transformed fundamental set of patterns. The details of Delta Associative Memory model can be reviewed in Reference [24]. It should be noted that both the learning phase and the classification phase of Delta Associative Memory model remained unchanged. For clarity purposes, in the present paper, the same symbology is used.

Preprocessing Phase

Data preprocessing phase is applied before Delta Associative Memory learning phase. This phase transforms the values of the input patterns of the fundamental set. This transformation of the input patterns is a data translation process that does not affect its representation or its statistical distribution. Furthermore, negative effects of *cross-talk* are eliminated; consequently, classification performance is improved. The proposed algorithm is as follows (Algorithm 1):

Algorithm 1: Preprocessing phase

Data: Fundamental set of associations $\{(\mathbf{x}^\mu, \mathbf{y}^\mu) \mid \mu = 1, 2, \ldots, p\}$ with p as the cardinality of
    the set

Result: Transformed fundamental set of associations $\{(\mathbf{x}^{\mu'}, \mathbf{y}^{\mu'}) \mid \mu = 1, 2, \ldots, p\}$

for $\mu = 1$ *to* p **do**
    | Generate displaced input patterns $\mathbf{x}^{\mu'}$ according to Expression (2) in [23];
    | Generate displaced input patterns $\mathbf{y}^{\mu'}$ according to Expression (3) in [23]
end

---

5. Performance Evaluation Methods

There are three main performance indicators of a binary classification test: sensitivity, specificity and classification accuracy. These indicators are computed from the confusion matrix values. Sensitivity and specificity are used for assessing the results of diagnostic and screening tests [43]. Sensitivity or True Positive Rate (TPR) represents the proportion of truly diseased persons in a screened population who are identified as being diseased by the test. Sensitivity is a measure of the probability of correctly diagnosing a condition. Specificity or True Negative Rate (TNR) is the proportion of truly healthy persons who are identified as so by the screening test. Classification accuracy of any algorithm can be estimated by taking into account the overall number of test patterns that are correctly classified.

6. Experimental Phase

The experimental phase of this paper is divided into two parts. In the first part, the coronary artery disease dataset, taken from the University of California at Irvine (UCI) machine learning repository [27], was used to evaluate the performance of the proposed model. The results of sensitivity, specificity and classification were compared against the performance achieved by fifty widely known algorithms, available in WEKA 3: Data Mining Software in Java [44]. The purpose of this stage is to evaluate the performance of the proposed algorithm using data that are still widely used by the

scientific community. Cross-validation (CV) was used as a technique to assess the generalizability of the proposed model to unknown patterns; specifically, *k*-fold cross-validation with $k = 10$ was used.

In the second part, the Sensor Platform shown in Figure 1 was integrated to a computing device based on the single-board computer paradigm. In this device, the proposed associative memory model was implemented and a database of medical patterns was generated. Once the device was trained, we proceeded with the tests of unknown patterns, generating classification performance results. To ensure the experimental results are reliable and valid on unknown patterns, all experiments were carried out following Kohavi and John recommendations [45]. Classification performance, sensitivity and specificity of the proposed model was compared against fifty widely known models, available in WEKA 3: Data Mining Software in Java [44].

Figure 1. The platform allows us to monitor biometric signals by using different sensors (courtesy of Libelium).

6.1. Heart Disease Dataset

This dataset comes from the Cleveland Clinic Foundation and was supplied by Robert Detrano, M.D., Ph.D. of the V.A. Medical Center, Long Beach, CA, USA. The purpose of the dataset is to predict the presence or absence of coronary artery disease given the results of various medical tests carried out on a patient. This dataset consists of 270 instances belonging to two different classes: presence and absence (of coronary artery disease). Each instance consists of 14 attributes, including the class attribute. This dataset and more information about the attributes are available at the University of California at Irvine (UCI) machine learning repository [27].

6.2. e-Health Sensor Platform Dataset

This dataset was created using the e-Health Sensor Platform, shown in Figure 1. It was built with the approval of the participants of the Research Projects 20130307 and 20140461, registered in the National Polytechnic Institute of Mexico (IPN). The objectives and goals of the projects were explained to each of the participants. Personally identifiable information was removed so that the dataset is anonymized. The purpose of the dataset is to predict the presence or absence of coronary artery disease given the results of various medical tests carried out on a patient. This dataset consists of 135 instances belonging to two different classes: presence and absence (of coronary artery disease). Each instance consists of seven attributes, including the class attribute. Attribute Information is as follows:

1. age
2. sex
3. maximum heart rate achieved
4. resting electrocardiographic results (values 0, 1, 2)
5. fasting blood sugar >120 mg/dL
6. resting blood pressure
7. class attribute: presence or absence (of coronary artery disease)

7. Results and Discussion

Classification performance of our proposal was compared against fifty widely known classification models. Tables 1 and 2 show classification performance, sensitivity and specificity achieved by the twenty best-performing algorithms of the fifty widely known algorithms, available in WEKA 3: Data Mining Software in Java [44].

Table 1. Classification accuracy using Heart Disease Dataset. Algorithms are presented in alphabetical order.

No	Algorithm	Sensitivity	Specificity	Accuracy
1.	AdaBoostM1	85.30	78.30	82.22
2.	Bagging	87.30	79.20	83.70
3.	BayesNet	86.00	77.50	82.22
4.	Dagging	88.00	75.00	82.22
5.	DecisionTable	87.30	78.30	83.33
6.	DTNB	85.30	79.20	82.59
7.	FT	86.00	77.50	82.22
8.	LMT	86.00	77.50	82.22
9.	Logistic	87.30	79.20	83.70
10.	MultiClassClassifier	87.30	79.20	83.70
11.	NaiveBayes	87.30	78.30	83.33
12.	NaiveBayesSimple	86.70	78.30	82.96
13.	NveBayesUpdateable	87.30	78.30	83.33
14.	RandomCommittee	86.70	76.70	82.22
15.	RandomForest	89.30	76.70	83.70
16.	RandomSubSpace	86.70	76.70	82.22
17.	RBFNetwork	86.70	80.83	84.07
18.	RotationForest	86.70	77.50	82.59
19.	SimpleLogistic	86.00	77.50	82.22
20.	SMO	86.70	79.20	83.33
21.	IDAM (our proposal)	86.70	80.83	84.07

Table 2. Classification accuracy using e-Health Sensor Platform Dataset. Algorithms are presented in alphabetical order.

No	Algorithm	Sensitivity	Specificity	Accuracy
1.	AdaBoostM1	94.10	96.40	95.60
2.	Bagging	95.40	96.60	96.19
3.	BayesNet	97.90	96.80	97.21
4.	Dagging	94.60	98.00	96.77
5.	DecisionTable	93.70	96.80	95.75
6.	DTNB	98.30	97.10	97.51
7.	FT	97.50	96.60	96.92
8.	LMT	94.10	97.70	96.48
9.	Logistic	94.60	97.70	96.63
10.	MultiClassClassifier	94.60	97.70	96.63
11.	NaiveBayes	97.10	95.70	96.19
12.	NaiveBayesSimple	97.90	95.50	96.33
13.	NveBayesUpdateable	97.10	95.70	96.19
14.	RandomCommittee	95.40	97.10	96.48
15.	RandomForest	97.50	96.80	97.07
16.	RandomSubSpace	95.00	96.20	95.54
17.	RBFNetwork	95.80	95.90	95.90
18.	RotationForest	97.90	96.80	97.21
19.	SimpleLogistic	94.10	98.00	96.63
20.	SMO	95.80	97.50	96.92
21.	IDAM (our proposal)	98.33	97.51	97.80

According to the type of learning scheme, each of these can be grouped into one of the following types of classifiers: Functions based classifiers, Meta classifiers, Rules based classifiers, Bayesian classifiers and Decision Trees classifiers.

The twenty best-performing algorithms are as follows:

- Four functions based classifiers (Logistic [46], RBFNetwork [47], SimpleLogistic [48] and SMO [49]).
- Seven meta classifiers (AdaBoostM1 [50], Bagging [51], Dagging [52], MultiClassClassifier [53,54], RandomCommittee [53,54], RandomSubSpace [55], RotationForest [56]).
- Two rules based classifiers (DecisionTable [57] and DTNB [58]).
- Four algorithms based on the Bayesian approach (BayesNet [59], NaiveBayes [60], NaiveBayesSimple [61] and NaiveBayesUpdateable [60]).
- Three decision trees classifiers (FT [62], LMT [62], and RandomForest [63]).

Tables 3 and 4 show classification accuracy achieved by the five best-performing algorithms of the fifty widely known classification models.

Table 3. Classification accuracy of the five best-performing algorithms using Heart Disease Dataset.

No	Algorithm	Sensitivity	Specificity	Accuracy
1.	Bagging	87.30	79.20	83.70
2.	Logistic	87.30	79.20	83.70
3.	RandomForest	89.30	76.70	83.70
4.	RBFNetwork	86.70	80.83	84.07
5.	IDAM (our proposal)	86.70	80.83	84.07

Table 4. Classification accuracy of the five best-performing algorithms using e-Health Sensor Platform Dataset.

No	Algorithm	Sensitivity	Specificity	Accuracy
1.	BayesNet	97.90	96.80	97.21
2.	DTNB	98.30	97.10	97.51
3.	RotationForest	97.90	96.80	97.21
4.	SimpleLogistic	94.10	98.00	96.63
5.	IDAM (our proposal)	98.33	97.51	97.80

As shown in Table 1, the algorithm that best identifies sick patients, using coronary artery disease dataset, is RandomForest with a Sensitivity value of 89.30. The model that best identifies healthy patients is Improved Delta Associative Memory which achieved a Specificity value of 80.83. As shown in Table 3, RBFNetwork algorithm and Improved Delta Associative Memory achieved the highest classification accuracy.

Performance achieved by Improved Delta Associative Memory is very competitive, as can be seen in Table 2. There are two algorithms that best identify healthy patients, using e-Health Sensor Platform Dataset, SimpleLogistic and Dagging with a Specificity value of 98.00. The model that best identifies sick patients is Improved Delta Associative Memory which achieved a Sensitivity value of 98.33. As shown in Table 4, Improved Delta Associative Memory algorithm achieved the highest classification accuracy.

As shown in Tables 1 and 2, there is no particular method that surpasses all other algorithms. Wolpert and Macready [64] proved that what an algorithm gains in performance on one class of problems is necessarily offset by its performance on the remaining problems.

As shown in Tables 3 and 4, Improved Delta Associative Memory has a competitive performance compared against the performance achieved by the fifty widely known algorithms, available in WEKA 3: Data Mining Software in Java [44]. It is worth noting that Improved Delta Associative Memory

achieved the best performance using the e-Health Sensor Platform Dataset. Similarly, it should be noted that Improved Delta Associative Memory achieved the best classification accuracy averaged over the two datasets.

It is necessary to highlight that, using e-Health Sensor Platform Dataset, Improved Delta Associative Memory model achieved the best classification performance as well as the highest capacity to identify sick patients.

As can be seen in Table 4, in the two most important values for medical diagnosis and decision making, Improved Delta Associative Memory model delivered the best performance.

8. Conclusions

The proposed model, called Improved Delta Associative Memory, showed competitive performance compared against the performance achieved by the fifty widely known algorithms, available in WEKA 3: Data Mining Software in Java [44].

It should be noted that Improved Delta Associative Memory achieved the best classification accuracy averaged over all datasets. Likewise, in the two most important values for medical diagnosis and decision making, Improved Delta Associative Memory model delivered the best performance.

Classification performance of Improved Delta Associative Memory demonstrates associative memories potential to develop applications based on artificial intelligence techniques to support medical specialists in healthcare monitoring and decision making.

The results presented in this paper demonstrate associative memories potential to predict the presence or absence of coronary artery disease for pattern classification systems.

References

1. Schieber, G.; Maeda, A. Health care financing and delivery in developing countries. *Health Aff.* **1999**, *18*, 193–205. [CrossRef]
2. Böhm, K.; Schmid, A.; Götze, R.; Landwehr, C.; Rothgang, H. Five types of OECD healthcare systems: Empirical results of a deductive classification. *Health Policy* **2013**, *113*, 258–269. [CrossRef] [PubMed]
3. Font, J.C. The State and Healthcare: Comparing OECD Countries. *West Eur. Politics* **2012**, *35*, 439–440. [CrossRef]
4. Senouci, M.R.; Mellouk, A. 1-Wireless Sensor Networks. In *Deploying Wireless Sensor Networks*; Senouci, M.R., Mellouk, A., Eds.; Elsevier: New York, NY, USA, 2016; pp. 1–19.
5. Botta, A.; de Donato, W.; Persico, V.; Pescapé, A. Integration of Cloud computing and Internet of Things: A survey. *Future Gener. Comput. Syst.* **2016**, *56*, 684–700. [CrossRef]
6. Llaves, A.; Corcho, Ó.; Taylor, P.; Taylor, K. Enabling RDF Stream Processing for Sensor Data Management in the Environmental Domain. *Int. J. Semant. Web Inf. Syst.* **2016**, *12*, 1–21. [CrossRef]
7. Guinea, A.S.; Nain, G.; Traon, Y.L. A systematic review on the engineering of software for ubiquitous systems. *J. Syst. Softw.* **2016**, *118*, 251–276. [CrossRef]
8. Qi, J.; Yang, P.; Min, G.; Amft, O.; Dong, F.; Xu, L. Advanced internet of things for personalised healthcare systems: A survey. *Pervasive Mob. Comput.* **2017**, *41*, 132–149. [CrossRef]

9.　　Hamdi, O.; Chalouf, M.A.; Ouattara, D.; Krief, F. eHealth: Survey on research projects, comparative study of telemonitoring architectures and main issues. *J. Netw. Comput. Appl.* **2014**, *46*, 100–112. [CrossRef]

10.　Della Mea, V. What is e-Health (2): The death of telemedicine? *J. Med. Internet Res.* **2001**, *3*, e22. [CrossRef] [PubMed]

11.　Rashvand, H.F.; Traver Salcedo, V.; Montón Sánchez, E.; Iliescu, D. Ubiquitous wireless telemedicine. *IET Commun.* **2008**, *2*, 237–254. [CrossRef]

12.　Segato, F.; Masella, C. Telemedicine services: How to make them last over time. *Health Policy Technol.* **2017**, *6*, 268–278. [CrossRef]

13.　Suryadevara, N.; Mukhopadhyay, S.; Wang, R.; Rayudu, R. Forecasting the behavior of an elderly using wireless sensors data in a smart home. *Eng. Appl. Artif. Intell.* **2013**, *26*, 2641–2652. [CrossRef]

14.　Garcia-Perez, C.; Diaz-Zayas, A.; Rios, A.; Merino, P.; Katsalis, K.; Chang, C.Y.; Shariat, S.; Nikaein, N.; Rodriguez, P.; Morris, D. Improving the efficiency and reliability of wearable based mobile eHealth applications. *Pervasive Mob. Comput.* **2017**, *40*, 674–691. [CrossRef]

15.　Suominen, H. Text mining and information analysis of health documents. *Artif. Intell. Med.* **2014**, *61*, 127–130. [CrossRef] [PubMed]

16.　Talaminos-Barroso, A.; Estudillo-Valderrama, M.A.; Roa, L.M.; Reina-Tosina, J.; Ortega-Ruiz, F. A Machine-to-Machine protocol benchmark for eHealth applications-Use case: Respiratory rehabilitation. *Comput. Methods Programs Biomed.* **2016**, *129*, 1–11. [CrossRef] [PubMed]

17.　Johnson, W.O.; Ward, E.B.; Gillen, D.L. Chapter 15-Bayesian Methods in Public Health. In *Disease Modelling and Public Health, Part A*; Rao, A.S.S., Pyne, S., Rao, C., Eds.; Handbook of Statistics; Elsevier: New York, NY, USA , 2017; Volume 36, pp. 407–442.

18.　Paradarami, T.K.; Bastian, N.D.; Wightman, J.L. A hybrid recommender system using artificial neural networks. *Expert Syst. Appl.* **2017**, *83*, 300–313. [CrossRef]

19.　Erkaymaz, O.; Ozer, M.; Perc, M. Performance of small-world feedforward neural networks for the diagnosis of diabetes. *Appl. Math. Comput.* **2017**, *311*, 22–28. [CrossRef]

20.　López-Vallverdú, J.A.; Riaño, D.; Bohada, J.A. Improving medical decision trees by combining relevant health-care criteria. *Expert Syst. Appl.* **2012**, *39*, 11782–11791. [CrossRef]

21.　Aldape-Pérez, M.; Yáñez-Márquez, C.; Camacho-Nieto, O.; Argüelles-Cruz, A.J. A New Tool for Engineering Education: Hepatitis Diagnosis using Associative Memories. *Int. J. Eng. Educ.* **2012**, *28*, 1399–1405.

22.　Cerón-Figueroa, S.; López-Yáñez, I.; Alhalabi, W.; Camacho-Nieto, O.; Villuendas-Rey, Y.; Aldape-Pérez, M.; Yáñez-Márquez, C. Instance-based ontology matching for e-learning material using an associative pattern classifier. *Comput. Hum. Behav.* **2017**, *69*, 218–225. [CrossRef]

23.　Aldape-Pérez, M.; Yáñez-Márquez, C.; Camacho-Nieto, O.; Argüelles-Cruz, A.J. An associative memory approach to medical decision support systems. *Comput. Methods Programs Biomed.* **2012**, *106*, 287–307. [CrossRef] [PubMed]

24.　Aldape-Pérez, M.; Yáñez-Márquez, C.; Camacho-Nieto, O.; López-Yáñez, I.; Argüelles-Cruz, A.J. Collaborative learning based on associative models: Application to pattern classification in medical datasets. *Comput. Hum. Behav.* **2015**, *51*, 771–779. [CrossRef]

25.　Ramírez-Rubio, R.; Aldape-Pérez, M.; Yáñez-Márquez, C.; López-Yáñez, I.; Camacho-Nieto, O. Pattern classification using smallest normalized difference associative memory. *Pattern Recognit. Lett.* **2017**, *93*, 104–112. [CrossRef]

26.　Kahramanli, H.; Allahverdi, N. Design of a hybrid system for the diabetes and heart diseases. *Expert Syst. Appl.* **2008**, *35*, 82–89. [CrossRef]

27.　Dheeru, D.; Karra Taniskidou, E. UCI Machine Learning Repository. 2017. Available online: http://archive.ics.uci.edu/ml (accessed on 7 August 2018).

28.　Polat, K.; Güneş, S. A new feature selection method on classification of medical datasets: Kernel F-score feature selection. *Expert Syst. Appl.* **2009**, *36*, 10367–10373. [CrossRef]

29.　McSherry, D. Conversational case-based reasoning in medical decision making. *Artif. Intell. Med.* **2011**, *52*, 59–66. [CrossRef] [PubMed]

30.　Anooj, P. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *J. King Saud Univ. Comput. Inf. Sci.* **2012**, *24*, 27–40. [CrossRef]

31.　Nahar, J.; Imam, T.; Tickle, K.S.; Chen, Y.P.P. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Syst. Appl.* **2013**, *40*, 1086–1093. [CrossRef]

32. Biswas, S.K.; Sinha, N.; Purakayastha, B.; Marbaniang, L. Hybrid expert system using case based reasoning and neural network for classification. *Biol. Inspir. Cogn. Archit.* **2014**, *9*, 57–70. [CrossRef]

33. Nguyen, T.; Khosravi, A.; Creighton, D.; Nahavandi, S. Classification of healthcare data using genetic fuzzy logic system and wavelets. *Expert Syst. Appl.* **2015**, *42*, 2184–2197. [CrossRef]

34. Leema, N.; Nehemiah, H.K.; Kannan, A. Neural network classifier optimization using Differential Evolution with Global Information and Back Propagation algorithm for clinical datasets. *Appl. Soft Comput.* **2016**, *49*, 834–844. [CrossRef]

35. Nahato, K.B.; Nehemiah, K.H.; Kannan, A. Hybrid approach using fuzzy sets and extreme learning machine for classifying clinical datasets. *Inform. Med. Unlocked* **2016**, *2*, 1–11. [CrossRef]

36. Acevedo-Mosqueda, M.E.; Yáñez-Márquez, C.; López-Yáñez, I. Alpha-Beta bidirectional associative memories: Theory and applications. *Neural Process. Lett.* **2007**, *26*, 1–40. [CrossRef]

37. Shah, S.; Batool, S.; Khan, I.; Ashraf, M.; Abbas, S.; Hussain, S. Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis. *Phys. A Stat. Mech. Its Appl.* **2017**, *482*, 796–807. [CrossRef]

38. Steinbuch, K. Die Lernmatrix. *Kybernetik* **1961**, *1*, 36–45. [CrossRef]

39. Steinbuch, K.; Frank, H. Nichtdigitale lernmatrizen als perzeptoren. *Kybernetik* **1961**, *1*, 117–124. [CrossRef] [PubMed]

40. Steinbuch, K.; Piske, U.A.W. Learning Matrices and Their Applications. *IEEE Trans. Electron. Comput.* **1963**, *EC-12*, 846–862. [CrossRef]

41. Steinbuch, K. Adaptive networks using learning matrices. *Kybernetik* **1964**, *2*, 148–152.

42. Steinbuch, K.; Widrow, B. A Critical Comparison of Two Kinds of Adaptive Classification Networks. *IEEE Trans. Electron. Comput.* **1965**, *EC-14*, 737–740. [CrossRef]

43. Ting, K.M., Sensitivity and Specificity. In *Encyclopedia of Machine Learning*; Sammut, C., Webb, G.I., Eds.; Springer: Boston, MA, USA, 2010; pp. 901–902.

44. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *SIGKDD Explor.* **2009**, *11*, 10–18. [CrossRef]

45. Kohavi, R.; John, G.H. Wrappers for Feature Subset Selection. *Artif. Intell.* **1997**, *97*, 273–324. [CrossRef]

46. Le Cessie, S.; van Houwelingen, J. Ridge Estimators in Logistic Regression. *Appl. Stat.* **1992**, *41*, 191–201. [CrossRef]

47. Buhmann, M.D. *Radial Basis Functions: Theory and Implementations (Cambridge Monographs on Applied and Computational Mathematics)*; Cambridge University Press: Cambridge, UK, 2003.

48. Sumner, M.; Frank, E.; Hall, M. Speeding up logistic model tree induction. In *9th European Conference on Principles and Practice of Knowledge Discovery in Databases*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 675–683.

49. Platt, J. *Fast Training of Support Vector Machines Using Sequential Minimal Optimization*; Advances in Kernel Methods-Support Vector Learning; MIT Press: Cambridge, MA, USA, 1998.

50. Freund, Y.; Schapire, R. Experiments with a new boosting algorithm. In Proceedings of the Thirteenth International Conference on Machine Learning, Bari, Italy, 3–6 July 1996; pp. 148–156.

51. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

52. Ting, K.M.; Witten, I.H. Stacking Bagged and Dagged Models. In *Fourteenth International Conference on Machine Learning*; Fisher, D.H., Ed.; Morgan Kaufmann Publishers: San Francisco, CA, USA, 1997; pp. 367–375.

53. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. *Weka 3: Data Mining Software in Java*; University of Waikato: Hamilton, Waikato, 2010.

54. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems)*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 2005.

55. Ho, T.K. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 832–844.

56. Rodriguez, J.J.; Kuncheva, L.I.; Alonso, C.J. Rotation Forest: A new classifier ensemble method. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 1619–1630. [CrossRef] [PubMed]

57. Kohavi, R. The Power of Decision Tables. In Proceedings of the 8th European Conference on Machine Learning, Heraclion, Crete, Greece, 25–27 April 1995; pp. 174–189.

58. Hall, M.; Frank, E. Combining Naive Bayes and Decision Tables. In Proceedings of the 21st Florida Artificial Intelligence Society Conference (FLAIRS), Coconut Grove, FL, USA, 15–17 May 2008; pp. 318–319.

59. Christofides, N. *Graph Theory: An Algorithmic Approach (Computer Science and Applied Mathematics)*; Academic Press, Inc.: Orlando, FL, USA, 1975.

60. John, G.; Langley, P. Estimating continuous distributions in Bayesian classifiers. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Montreal, QU, Canada, 18–20 August 1995; pp. 338–345.

61. Duda, R.; Hart, P. *Pattern Classification and Scene Analysis*; Wiley: New York, NY, USA, 1973.

62. Landwehr, N.; Hall, M.; Frank, E. Logistic Model Trees. *Mach. Learn.* **2005**, *59*, 161–205. [CrossRef]

63. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

64. Wolpert, D.H.; Macready, W.G. No Free Lunch Theorems for Optimization. *IEEE Trans. Evolut. Comput.* **1997**, *1*, 67–82. [CrossRef]