



A Collaborative Framework for Sensing Abnormal Heart Rate Based on a Recommender System: Semantic Recommender System for Healthcare

Giovanni Guzmán¹ · Miguel Torres-Ruiz¹ · Vianney Tambonero¹ · Miltiadis D. Lytras^{2,3} · Blanca López-Ramírez⁴ · Rolando Quintero¹ · Marco Moreno-Ibarra¹ · Wadee Alhalabi⁵

Received: 22 March 2017 / Accepted: 23 April 2018 / Published online: 16 May 2018
© Taiwanese Society of Biomedical Engineering 2018

Abstract

According to different studies, people are not able to identify the physiological symptoms related to high risk cardiovascular condition that could require medical attention. In consequence, when they see a medical doctor, the heart damage could be quite advanced. Moreover, there are several studies focused on applying the Framingham or systematic coronary risk evaluation indexes; however, the combination with other physiological variables such as lifestyle, current activity, and maximal heart rate has not been deeply studied in the state-of-the-art. This paper proposes a collaborative framework for sensing physiological variables to determine possible high risk cardiovascular conditions, it will also provide a weighted ranking list of medical speciality centers. The framework will consist of two stages: in the first one, an ubiquitous heart rate monitoring by using an ID3 decision tree is applied to classify sensed-data for identifying the presence of a high risk cardiovascular condition. The second stage proposes a recommender system leading towards extracting and clustering of a set of hospitals, in which the medical specialities are defined in an application ontology. The clustering process matches the hospital attention factor, in order to estimate the number of possible medical doctors and the required cardiovascular medical speciality. In conclusion, the proposal applies different decision trees such as ID3, J48, NBTree, and BFTree in order to evaluate and compare the classification performance. The effectiveness of the ID3 decision tree was 85.71%.

Keywords Healthcare · Monitoring · Bluetooth sensor device · Recommender system · Medical ontology · Decision tree

✉ Miguel Torres-Ruiz
mtorres@cic.ipn.mx

Giovanni Guzmán
jguzman1@cic.ipn.mx

Vianney Tambonero
viannne@hotmail.com

Miltiadis D. Lytras
mlytras@acg.edu

Blanca López-Ramírez
bllopez@itroque.edu.mx

Rolando Quintero
quintero@cic.ipn.mx

Marco Moreno-Ibarra
marcomoreno@cic.ipn.mx

Wadee Alhalabi
walhalabi@effatuniversity.edu.sa

- ¹ Instituto Politécnico Nacional, CIC, UPALM Zacatenco, 07320 Mexico City, Mexico
- ² The American College of Greece, 6 Gravias Street GR-153 42 Aghia Paraskevi, Athens, Greece
- ³ King Abdulaziz University, 21589, Jeddah 80200, Saudi Arabia
- ⁴ Instituto Tecnológico de Roque - Tecnológico Nacional de México, Km.8 Carretera Celaya-Juventino Rosas, 38110 Celaya, Mexico
- ⁵ Virtual Reality Research Center, Effat University, 21478, Jeddah 34689, Saudi Arabia

1 Introduction

The population ageing has contributing factors such as decrease of mortality in the last 50 years, the increase of life expectancy in population, and the reduction in the birth rate [1]. According to official reports [2], adults are the most vulnerable group to have health damages (dead, disease and disability). The World Health Organization (WHO) described that when cardiovascular diseases are presented in people of median age, in most cases, the impact is significant in their economic status. In many under developed countries, cardiovascular diseases represent a great percent of social spending, because medical programs that respond to population requirements are more reduced in comparison with developed countries. Especially, issues related to early detection of some pathological conditions [3]. Moreover, the WHO estimated that distribution of the most diseases in 2011 year were caused: ischemic heart disease (7 million), stroke (6.2), lower respiratory infections (3.2), chronic obstructive lung disease (3), diarrhoea (1.9) and Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome, HIV/AIDS (1.6). According to [2], these diseases have remained as the main causes of death during the past decade [4]. On the other hand, tuberculosis is no longer among the 10 leading causes of death, but it is still among the top 15, killing one million people in 2011. Chronic diseases have originated most deaths worldwide, such as lung cancers (along with trachea and bronchus cancer) produced 1.5 million deaths in 2011, up from 1.2 million deaths in 2000. Similarly, diabetes produced 1.4 million deaths in 2011, up from 1.0 million deaths in 2000. The number of deaths associated with heart disease in the last decade had an increment from 5.9 million in 2000 to 7 million in 2011 [4].

In human computer interaction perspective, the human being requires to explicitly assign tasks to computers [5]. Ubiquitous computing offers a new perspective in which users are surrounded by numerous computing devices, which are implicitly used to perform daily activities. Thus, it is expected to be an acceptance of the use of communication technologies and information as a potential resource, with the purpose to improve healthcare systems worldwide [6]. The problems related to provide health services for future epidemiological challenges have not been solved. They are focused on giving access and equitable distribution of health services, as well as to reduce the growth in healthcare costs, which have mainly caused the increase in population life expectancy. Several authors consider that the use of ubiquitous computing in the healthcare is the evolution of telemedicine, because it goes a step beyond traditional desktop applications [6].

Moreover, the technological advances are reflected in some industries such as tourism services, which have opened

the stage for the emergence of recommender systems, that are focused on providing information in e-shops, where each recommendation satisfies both the needs and tastes of the user [7]. Some factors that are considered as basis for making a recommendation are the follows: user preferences among alternative products, product attributes, preferences and choices of other people, expert opinions and individual characteristics that can predict the user preferences.

Thus, the implications and potential of portable monitoring in cardiology are very important, because this information allows us to detect early signs of health deterioration, notifies healthcare providers of critical conditions, determine correlations between lifestyle and health condition, diffuse new dimensions of exercise workouts, providing detailed information about physiological signals in different exercise scenarios, and provide multiple health repositories of physiological data in real-time [8].

In this paper, we are focused on representing different features of the medical user profile by means of an application ontology, in order to find answers like the set of hospitals and arranging them by using a ranking approach called *attention factor*. We also try to determine the most adequate medical locations for attending some medical emergency that requires a specific (or related) medical speciality, inside an ubiquitous architecture. In fact, the emergency is automatically determined by using an ID-3 decision tree that continuously receives and analyzes the information by the obtained heart frequency by using a Bluetooth sensor device.

The rest of the paper is organized as follows: Sect. 2 focuses on the related work by focusing on continuous monitoring of cardiovascular signals. These are then analyzed and classified. Section 3 describes the proposed ubiquitous architecture for monitoring and ranking information. The experimental results are presented in Sect. 4. Conclusion and future work are outlined in Sect. 5.

2 Related Work

In this section, a description of the related works of mobile-based monitoring and the classification of health data, based on machine learning techniques is presented. Extensive researches have been proposed to develop mobile-based monitoring systems. They tackled different monitoring perspectives and addressed different issues related to handling the monitoring of cardiovascular signals mainly Electrocardiogram (ECG) by means of mobile devices, particular applications, or specific studies.

In [9], a heart rate monitoring task based on the R-peak detection method was proposed. The R wave represents the electrical stimulus in the heart as it passes through the main portion of the ventricular walls and is measured as the

highest peak in a single beat shown in an ECG. A method for non-contact measurement of heart rate using thermal video was presented in [10]. The measurements were performed on three different areas: the whole face, the upper half of the face and the supra orbital region. A tracker was used to trace regions in order to make the algorithm invulnerable to motion. After tracing, the videos were spatially filtered, in order to increase the signal to noise ratio; and then, the video frames were filtered applying a band pass filter for extracting the thermal variations caused by blood circulation. Finally, the heart rate was computed by using two methods, zero-crossing and Fast Fourier Transform. Another non-contact measurement approach is presented in [11]. This system is based on Doppler radar modules that detect movements due to heartbeat. Other and different techniques to estimate the heart rate have been proposed. The photoplethysmogram (PPG) is a method that depends on the optical properties of the skin [12]. 'FitBeat' is a system for heart rate tracking on a smart watch for monitoring physical exercises that is based on PPG [13]. This system integrates contact and motion sensing as well as an algorithm based on simple spectral analysis to eliminate some errors. The 'FitBeat' architecture consists of three components: (1) Contact sensing, which continuously monitors the amplitude and variance of PPG signal, (2) Noise reduction analyzes both PPG signal and accelerometer data to remove motion-induced noise by means of adaptive filtering, and (3) Pulse identification that reduces the residual motion-induced noise by co-analyzing the spectrum of PPG signal and accelerometer data, and performs spectral analysis. Recently, [14] an approach based on PPG signals for heart rate monitoring has been presented, during intensive physical exercises. The system is composed of the following tasks: (1) a pre-processing, where signals are normalized, (2) Noise signal, which is focused on attenuating the influence of motion artifacts by means of Wiener filtering, (3) A heart rate estimation, which performs a refinement to the initial computation by applying the Vocoder technique, and (4) User-adaptive post-processing (online and offline), to track the subject physiology; online post-processing is based on heuristic rules and thresholds, and offline post-processing is performed by Viterbi decoding. The performance of the heart rate estimation systems was tested with 23 PPG recordings; the average absolute error was 1.97 BPM.

In [15], a platform that is composed of communication protocols, signal acquisition and processing modules in a mobile environment to sense ECG signals is proposed. In [16, 17], a monitoring system that detects cardiac diseases such as tachycardia and acute myocardial is presented. Moreover, a real-time monitoring application that includes an intelligent data acquisition system with identification algorithms and threshold-based classification is proposed by [18]. In [19], a mobile-based monitoring system for different

cardiac diseases was proposed. The main outcomes were ECG filtering, feature extraction and arrhythmia detection, which were displayed on a mobile device. Other mobile systems were developed in Android in order to monitor in real-time ECG signals for detecting arrhythmia [20]. A real-time, scalable health monitoring system, based on cloud infrastructure and software services for data collection and analysis was proposed in [21]. In other context, frameworks to such as Armed-Cardio [22] and MobiHealth [23] have presented architectures and platforms by using Global System for Mobile Communication (GSM) technology for chronic heart diseases, offering access to remote management of such chronic conditions, and triggering emergency interventions. Other applications have been focused on reasoning systems, including wireless communications [24, 25]. In [26], some monitoring systems based on smartphones and wearable sensors have been described and divided into two groups. The first group uses smartphones equipped with biosensors, examples include: Alive technology, Vita-phone, Ventracor pocketview, and Welch Allyn Micropaq. The second group is oriented towards building platforms for remote health monitoring, such as Mobihealth, Tele-medicare, Osiris-SE, and PhMon. An ubiquitous healthcare monitoring system for Android tablets, which acquires and processes physiological data, including weighting, blood pressure and heart rate information from wearable sensors was proposed in [27]. In addition, an architecture for accurate heart rate detection system has been presented by using the inertial sensors of standard smartphones [28]. Moreover, a platform based on a web-based application for integrating a set of different services for cardiovascular clinical decision support is also proposed [29]. The services that the system provides are the follows: (i) a service for adding all information gathered by specialists, (ii) a tool to automatically determine a risk stratification, and (iii) visualization tools for representing ambulatory records and general reports about the patient's conditions.

In addition, the Chronic Care Model (CCM) is a conceptual framework based on evidence developed by the United States [30]. The World Health Organization (WHO) proposed an Innovative Care for Chronic Conditions (ICCC), which extends the CCM proposal to define a model to international benchmarks [2]. According to [31], the behavior of people with memory disorders was analyzed using Semantic Web technology that has been essential to monitoring chronic conditions, through the definition of a context model ontology. Recently, the care procedures like the Pervasive Self Care have considered ubiquitous computing services as a framework infrastructure [31]. The ubiquitous heart monitoring is a variant of different approaches for the ubiquitous computing, which is oriented towards healthcare domain. It is characterized by the union of portable medical sensors,

communication technologies and information for continuous or intermittent cardiac monitoring of a patient.

Thus, different architectures that were proposed for ubiquitous systems, in order to minimize the proper problems derived from the nature of ubiquitous systems. For instance, issues related to heterogeneity, mobility, communication, among ubiquitous platforms are described in [8]. Generally, traditional architectures cannot be individually implemented in a ubiquitous system, but they can combine their better characteristics to develop a specific platform. New proposals such as the Kameran project (Knowledge Management in Ambient Intelligence) [31], are focused on developing a platform for context-sensitive service for home health care. This system works with a collaborative action between the family, the patient and the medical center or hospital. Family members provide care services for the patient, by means of monitoring devices. A personal computer is connected via Asymmetric Digital Subscriber Line (ADSL) to a server located in the hospital. When the alarm is activated, the computer sends a message to the hospital with some relevant data such as patient identifier, the intensity level of the alarm and the shooting conditions. In Case of Emergency (ICE) is a mobile application that stores information about possible patients, it is vital for rescuers and medical specialists when an emergency is presented [32]. The stored data contain emergency contact, insurance information, name and number of personal medical doctor, allergies, chronic medical conditions, current medications and additional informational and instructions. The patient or another external user can make the call from the application, even when the cell phone is locked. Thus, the specialists receive information before the patient arrives to the hospital.

The development of frameworks in this thematic has played an important role in the components for the architecture of ubiquitous systems. Bottazzi [33] presented the need of ubiquitous systems in order to collect, process and distribute contextualize information to determine appropriate assistance strategies. An ubiquitous system tends to consider different pieces that compose the context information. For instance, geographic location, gestures and heart rate of elderly persons. According to [33], it is necessary to adapt infrastructure in the daily routine of a patient activities; for example, wireless biomedical devices identity cards based on Radio Frequency Identification (RFID) technology, wireless, bluetooth, and magnetic cards. Allocation and Group Aware Pervasive Environment (AGAPE) supports context adaptive management of scenarios in the care of senior citizens [33]. Additionally, it offers a set of tools and mechanisms for the collection and incorporation of context information from heterogeneous and distributed sources, i.e., sensors and medical records repository. Other application such as UCARE was developed by the Research Centre of Sensor Networks and Application of Beijing, China [34].

It is a platform for the ubiquitous healthcare system. The architecture is based on two components: a server device and UCare. It consists of sensors and a laptop, in which the sensors are treated as nodes that perform data collection, pre-processing and data transmission to the computer, using a wireless protocol. Those sensors are divided into two types, the first category covers context sensors (accelerometers, microphones and thermometers), and the second type uses physiological sensors (ECG, blood pressure, pulse oximetry). Unfortunately, the system only works with ECG, accelerometers and temperature sensors. Portable processing is done by using a smartphone, data are received via Bluetooth and they are transmitted to the server using Wi-Fi.

The assistant sensitive real-time context combines the smart use of vital signs remote monitoring with context (CARA) [35]. It is an architecture that improves the healthcare. Vital signs are monitored using a Body Area Network (BAN) based on sensors that can monitor: ECG, blood pressure, blood oxygen and the user's location. It only requires a device equipped with web browser and camera flash. The most important aspects of this application are the internal display between the caregiver and the patient, interactive real-time medical assistance, and review of inquiry records by a medical professional. This functionality provides the basis for performing automatic analysis.

The *geolocation platform* uses the Health Level Seven (HL7) communication protocol, mobile phone technology, and Global Positioning System (GPS). This platform allows the computation of the distance and location of various medical centers based on the geographic location of the mobile user [36]. It defines three types of queries: *general*, *category* and *service*. In the first one, only the coverage radius is specified. In *category*, both the radius and the type of required medical center are established. Finally, in the service, the medical specialist is provided (i.e, cardiology). The platform performs a validation process of the query request, and displays the health centers located inside the specified radius, according to the selected criteria. In addition, the hospital location is visualized in a map.

With respect to health data classification, different supervised and unsupervised machine learning approaches have been proposed in the state-of-the-art. In [37], ECG feature extraction was performed by applying a cross-wavelet transform-based approach. In [37], a classification of ventricular arrhythmia based on support vector machine and morphological structures were applied. Moreover, neural networks were used for classifying ECG arrhythmia and myocardial infarction [38, 39]. In [40], a method to classify ECG signals according to a combination of projected and dynamic features is described. It consists of deriving from a random projection matrix, in which each column is normalized and each row is transformed by discrete cosine transform. In addition, a support vector machine classifier was applied

to cluster heartbeats into different classes. Due to the presence of noise in the ECG signals some diseases such as the arrhythmia are difficult to recognize and classify. Thus, in [41], a combination between linear and non-linear features is proposed. This combination is based on five types of beat classes of arrhythmia and the application of discrete wavelet transform coefficients is considered. The features were tested for their ability to select different classes of data by using different classifiers such as support vector machine and neural network techniques. [42] presented an approach that consists of using a block-based neural network and is proposed to classify ECG heartbeats of a patient in five types. The optimization of the network structure and the weights was performed by applying particle swarm optimization algorithm. Other methods based on fuzzy logic classifiers have been identify patterns of ECG arrhythmic classification [43].

Other initiatives based on Markov models were designed to classify patients according to the dynamics of features extracted from their ECG [44]. In [45], a prediction of heart failure based on the statistical analysis of seven clinical variable factors taken from electronic databases was proposed. It consists of computing each risk variable by applying a random-effects model weighted by inverse variance. In [46] the authors describe automatic methods for classification and identification of heartbeat abnormalities based on ECG analysis. These methods use ECG signal preprocessing, segmentation techniques, feature description methods and learning algorithms.

Most of the above studies and researches proposed solutions to support monitoring of patients using mobile devices and other artifacts. However, they are mainly focused on exploring ECG signals in order to gather information for medical specialists. In some cases, these initiatives have applied machine learning techniques, such as support vector machines, artificial neural networks, and probabilistic approaches for clustering and classifying in different classes heart diseases. Those works are oriented towards medical specialists in order to analyze and make decisions about possible treatments and diagnostics.

Moreover, it is important to mention that there are not particular or specific works towards applying recommender system. The combination of web semantic technologies, and geoprocessing approaches to develop a framework focused on healthcare applications, and particularly with the mobile monitoring based on wearable sensors.

Thus, the proposed framework is focused on recommending and conducting the application of decision trees for determining a high risk cardiovascular condition, taking into consideration different criteria such as Framingham and SCORE cardiovascular risk indexes that have been deeply explored in the last years, as well as lifestyle, level of activity, among others.

3 Methodology

The methodology related to the proposed research is presented as a conceptual framework, which is depicted in Fig. 1). It consists of two main stages: the first one, is related to the heart rate monitoring in order to detect an anomaly, while the second stage, is based on a recommender system of medical services taking into account the geographic location and the design of an application ontology.

3.1 Detection of Abnormal Heart Rate

The conceptual design of the application and the interaction between all components is depicted in Fig. 1. From a general perspective, sensors interact with the smartphone by means of sending recollected data of the physiological monitoring and the physical activity of the user. The mobile device is involved with the reception, analysis and temporal tasks of all received data. Additionally, it emits the alert if the decision module determines a high risk cardiovascular condition. It also interacts with the server sending updated data that were received from sensors and values determined in the analysis task. The features that compose this stage are the follows:

1. *Sensors* They refer to the accelerometer and the portable heart rate monitor (for this research Zephyr HxM device was used). The accelerometer determines the current physical state of the patient.
2. *Web Server* It is in charged of storing, managing, and presenting processes of the collected data by the mobile device.

3.1.1 The Monitoring Management

The *monitoring management* includes several tasks that are related to heart rate monitoring and alert generation. The communication with the database and the preferences is essential to the decision tree. By using the preferences, it is possible to obtain the user's clinical data and the database contains all files generated in the heart rate sampling process.

3.1.1.1 Analysis of physiological data In order to obtain the vital signs, the heart rate monitor is used. Moreover, in the user preferences the value that indicates the sampling analysis is established. The heart frequency is classified into two categories: *normal* or *abnormal*. During periods of activity, the heart rate is established as normal or abnormal by means of computing the maximal heart rate (see Eq. 1), proposed by [47]. Later, a comparison of this value and the sampled frequency over a period of activity is performed, in order to

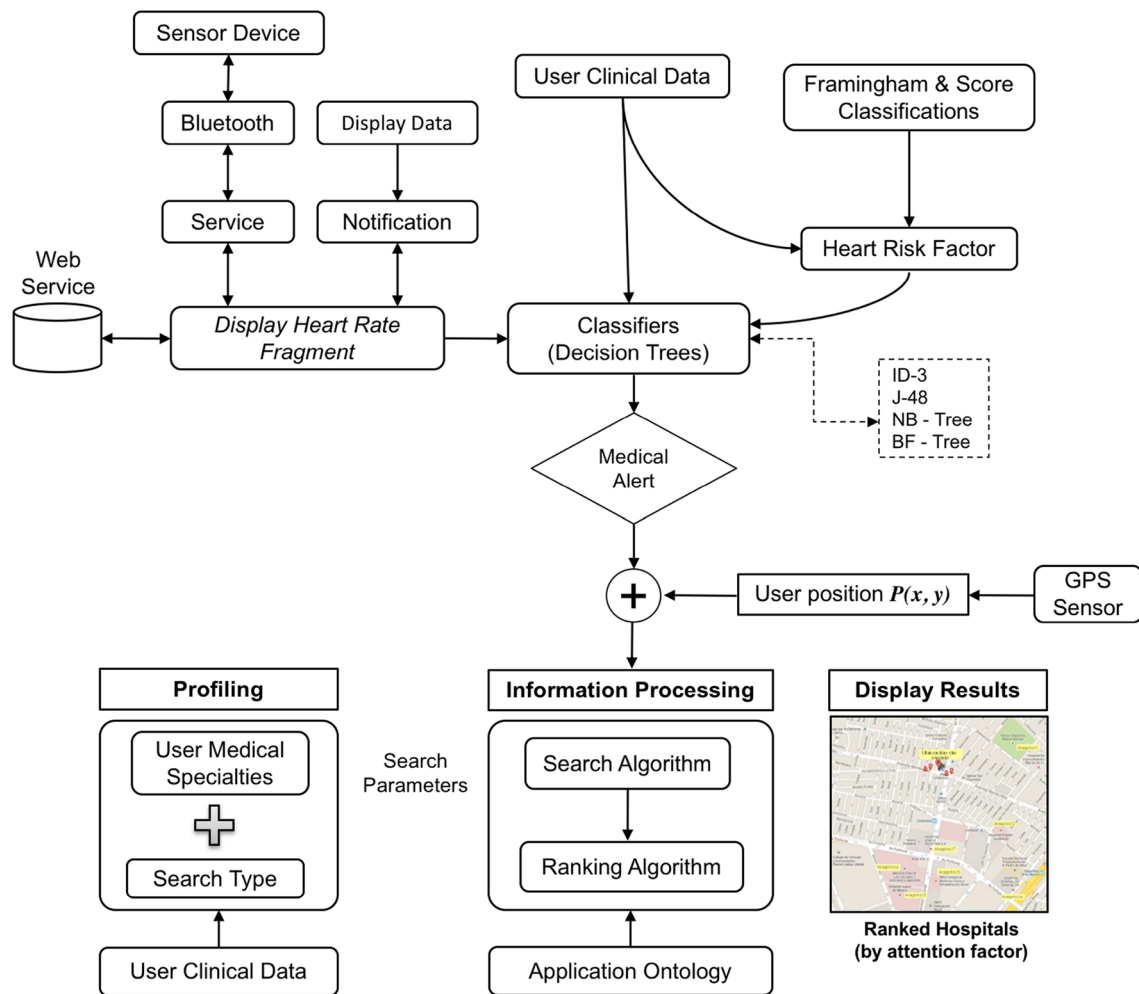


Fig. 1 The general framework that represents the architecture of the system

determine if the current heart rate is normal (less or equal than maximal heart rate) or abnormal (greater than maximal heart rate). Additionally, it is necessary to determine the average speed by using a sensor on the mobile device, i.e. accelerometer, and classify this value according to the state: *inactive*, *walking* or *running*. In order to determine the lifestyle (*normal*, *sedentary* or *active*), it is necessary to classify the heart rate frequency in passive state, the values proposed by Wilmore and Costill [48] are used for this classification. Table 1 describes these values.

$$HR_{max} = 208.75 - (0.73 * \text{years_old}) \tag{1}$$

3.1.1.2 Context Data Analysis Information of user context is generated with some sensors such as thermometer, accelerometer or GPS. The accelerometer sensor obtains the values that are used to compute the patient speed. Both features, the speed and the heart rate, are received in the

Table 1 Ranges used in monitoring management

Element to classify	Range	Unit	Result
Heart frequency	$60 < hf < 80$	bmp	Normal
	$hf > 100$	bmp	Sedentary
	$20 < hf < 40$	bmp	Active
Physical activity	$speed = 0.0$	m/s	Inactive
	$4 \leq speed \leq 7$	m/s	Walking
	$7 < speed < 11$	m/s	Running
Cardiovascular risk	$cr > 20$	Percent	High
	$10 \leq cr \leq 20$	Percent	Moderate
	$0 < cr < 10$	Percent	Low

same message transmitted by the mobile device. It is necessary the comparison of received information with the ranges described in Table 1, in order to deduce the physical activity.

3.1.1.3 Cardiovascular Risk Computation The Framingham equations [49] to determine the percentage of cardiovascular risk were applied. When the risk rate is obtained, Table 1 is used to establish the type of cardiovascular risk. According to [50] the Framingham equations are more adequate to estimate the cardiovascular risk that SCORE index, in Mexican patients.

3.1.1.4 Decision-Making Module This module works with the ID3 decision tree method, with the following input variables: data analysis of heart frequency, sensed samples obtained from the accelerometer, the lifestyle, and the cardiovascular risk values. Thus, Algorithm 1 presents the process for generating the decision tree.

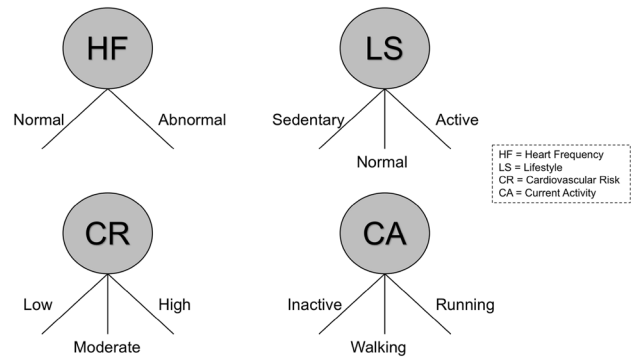


Fig. 2 Nodes and attributes for the decision tree

```

1  S is use to denote the set of samples applied, and  $\Lambda(S)$  to denote the set of attributes
2  if  $(s = \min(S)) > 0, \forall s \in S$  then
3  | return a node with the s label.
4  if  $(s = \max(S)) < 0, \forall s \in S$  then
5  | return a node with the s label.
6  if  $|\Lambda| = \emptyset$  then
7  | return a node labeled with the most frequent value of attribute – object from S.
8  else
9  | Let  $A \in \Lambda$  be the attribute that best classifies S
10 | Create a tree T, that contains a node labeled with A.
11 | for each valid value  $v \in A$  do
12 | | Add an arc with label v to T.
13 | | Let  $S(v)$  the set of samples such that  $A = v$ 
14 | | if  $S(v) = \emptyset$  then
15 | | | Connect the arc with a node labeled with the most frequent value of
16 | | | attribute – object from S.
17 | | else
18 | | | Connect the arc with the subtree  $ID3(S(v), \text{attribute – object}, \Lambda(A))$ .
18 return T

```

Algorithm 1: Generation of the decision tree.

The nodes and attributes of the decision tree are depicted in Fig. 2, taking into consideration the values and attributes of the heart frequency, cardiovascular risk, current activity and lifestyle. On the other hand, Table 2 describes the dataset applied in the training process of the decision tree, which is based on the criteria defined by Wilmore and Costill [48]. However, different scenarios according to medical requirements can be defined. For instance, it is possible to assign strict alarm emission in medical cases where patients or relative members have high risk cardiovascular records.

The main objective of the proposed approach is the possibility to implement the high risk cardiovascular condition algorithm by using mobile devices. By this reason, the use of decision trees is adequate for the proposed method, because they only require the use of *if-else if* code statements, which

is very important for improving the performance and computation task in the mobile device. The ID3, J48, NBTree, and BFTree were applied to the training dataset presented in Table 2, using normal alarm emission conditions. On the other hand, Tables 3 and 4 describes the obtained results implementing different decision trees in order to demonstrate the accuracy of these classification algorithms. The implementation was performed in the Weka software with the version 3.6.12.

By using the ID3 algorithm, the decision tree was correctly obtained, and it covered all cases. In Fig. 3 a graphical representation of the decision tree is depicted. Additionally, the generated result using strict alarm emission values is represented in Fig. 4. Moreover, in Figs. 5, 6 and 7 a comparison between the *Precision*, *Recall* and *F* measures is presented.

Table 2 Training dataset for the decision tree

Cardiovascular frequency	Cardiovascular risk	Current activity	Lifestyle	Normal alarm emission	Strict alarm emission
Normal	Low	Inactive	Sedentary	No	No
Normal	Low	Inactive	Normal	No	No
Normal	Low	Inactive	Active	No	No
Normal	Low	Walking	Sedentary	No	No
Normal	Low	Walking	Normal	No	No
Normal	Low	Walking	Active	No	No
Normal	Low	Running	Sedentary	No	Yes
Normal	Low	Running	Normal	No	No
Normal	Low	Running	Active	No	No
Normal	Moderate	Inactive	Sedentary	No	No
Normal	Moderate	Inactive	Normal	No	No
Normal	Moderate	Inactive	Active	No	No
Normal	Moderate	Walking	Sedentary	No	No
Normal	Moderate	Walking	Normal	No	No
Normal	Moderate	Walking	Active	No	No
Normal	Moderate	Running	Sedentary	Yes	Yes
Normal	Moderate	Running	Normal	No	Yes
Normal	Moderate	Running	Active	No	Yes
Normal	High	Inactive	Sedentary	No	No
Normal	High	Inactive	Normal	No	No
Normal	High	Inactive	Active	No	No
Normal	High	Walking	Sedentary	No	No
Normal	High	Walking	Normal	No	No
Normal	High	Walking	Active	No	No
Normal	High	Running	Sedentary	Yes	Yes
Normal	High	Running	Normal	No	Yes
Normal	High	Running	Active	No	Yes
Abnormal	Low	Inactive	Sedentary	No	Yes
Abnormal	Low	Inactive	Normal	No	Yes
Abnormal	Low	Inactive	Active	No	Yes
Abnormal	Low	Walking	Sedentary	No	Yes
Abnormal	Low	Walking	Normal	No	Yes
Abnormal	Low	Walking	Active	No	Yes
Abnormal	Low	Running	Sedentary	No	Yes
Abnormal	Low	Running	Normal	No	Yes
Abnormal	Low	Running	Active	No	Yes
Abnormal	Moderate	Inactive	Sedentary	Yes	Yes
Abnormal	Moderate	Inactive	Normal	No	Yes
Abnormal	Moderate	Inactive	Active	No	Yes
Abnormal	Moderate	Walking	Sedentary	Yes	Yes
Abnormal	Moderate	Walking	Normal	No	Yes
Abnormal	Moderate	Walking	Active	No	Yes
Abnormal	Moderate	Running	Sedentary	Yes	Yes
Abnormal	Moderate	Running	Normal	No	Yes
Abnormal	Moderate	Running	Active	No	Yes
Abnormal	High	Inactive	Sedentary	Yes	Yes
Abnormal	High	Inactive	Normal	Yes	Yes
Abnormal	High	Inactive	Active	Yes	Yes
Abnormal	High	Walking	Sedentary	Yes	Yes

Table 2 (continued)

Cardiovascular frequency	Cardiovascular risk	Current activity	Lifestyle	Normal alarm emission	Strict alarm emission
Abnormal	High	Walking	Normal	Yes	Yes
Abnormal	High	Walking	Active	Yes	Yes
Abnormal	High	Running	Sedentary	Yes	Yes
Abnormal	High	Running	Normal	Yes	Yes
Abnormal	High	Running	Active	Yes	Yes

Table 3 Summary of evaluation on training dataset

Description	ID3	J48	NBTree	BFTree
Correctly classified instances	54	100%	52	96.296%
Incorrectly classified instances	0	0%	2	3.703%
Kappa statistic	1	–	0.8989	–
Mean absolute error	0	–	0.0576	–
Root mean squared error	0	–	0.1697	–
Relative absolute error	0%	–	14.840%	–
Root relative squared error	0%	–	38.722%	–
Total number of instances	54	–	54	–

Table 4 Detailed accuracy by class

Classifier	TP rate	FP rate	Precision	Recall	F-measure	ROC area	Class
ID3							
–	1	0	1	1	1	1	No
–	1	0	1	1	1	1	Yes
Weighted average	1	0	1	1	1	1	–
J48							
–	1	0.143	0.952	1	0.976	0.988	No
–	0.857	0	1	0.857	0.923	0.988	Yes
Weighted average	0.963	0.106	0.965	0.963	0.962	0.988	–
NBTree							
–	1	0.143	0.952	1	0.976	1	No
–	0.857	0	1	0.857	0.923	1	Yes
Weighted average	0.963	0.106	0.965	0.963	0.962	1	–
BFTree							
–	0.95	0.071	0.974	0.95	0.962	0.988	No
–	0.929	0.05	0.867	0.929	0.897	0.988	Yes
Weighted average	0.944	0.066	0.946	0.944	0.945	0.988	–

3.1.1.5 Complexity of Decision-Making Algorithm In the worst case, the complexity computation of a decision tree occurs when a full tree is built, where each generated path inspects each input feature. In some *i*-th level, the algorithm inspects the *m* – 1 remaining features for all level instances for determining the gain of information. However, by apply-

ing the a-priori learning process, a linear complexity is obtained, i.e. $O(nm)$, because the algorithm generates a partial tree instead of a full decision tree (see Eq. 2).

$$\sum_{i=1}^m i * n = O(nm^2) \tag{2}$$

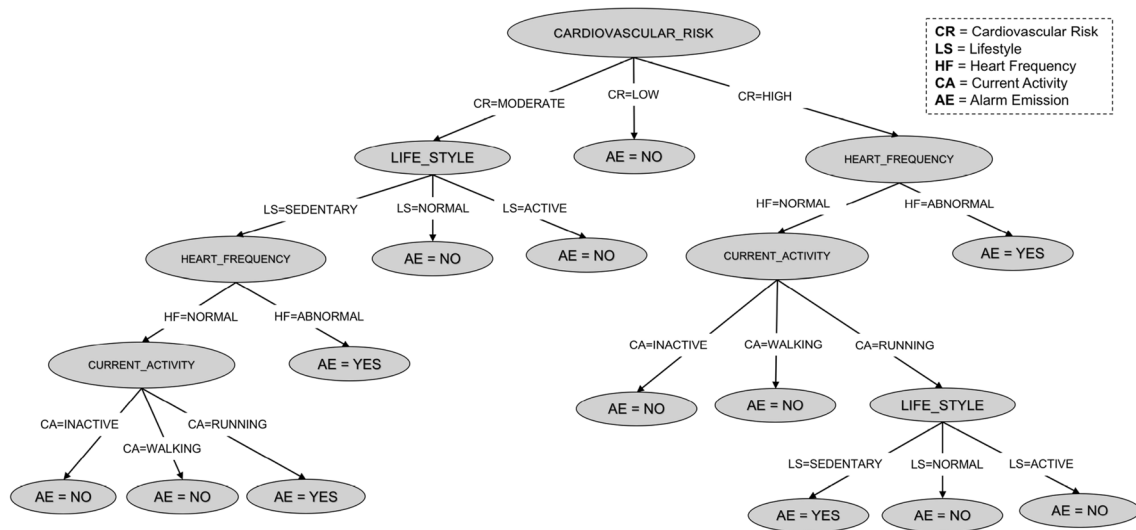


Fig. 3 ID3 decision tree using normal alarm emission

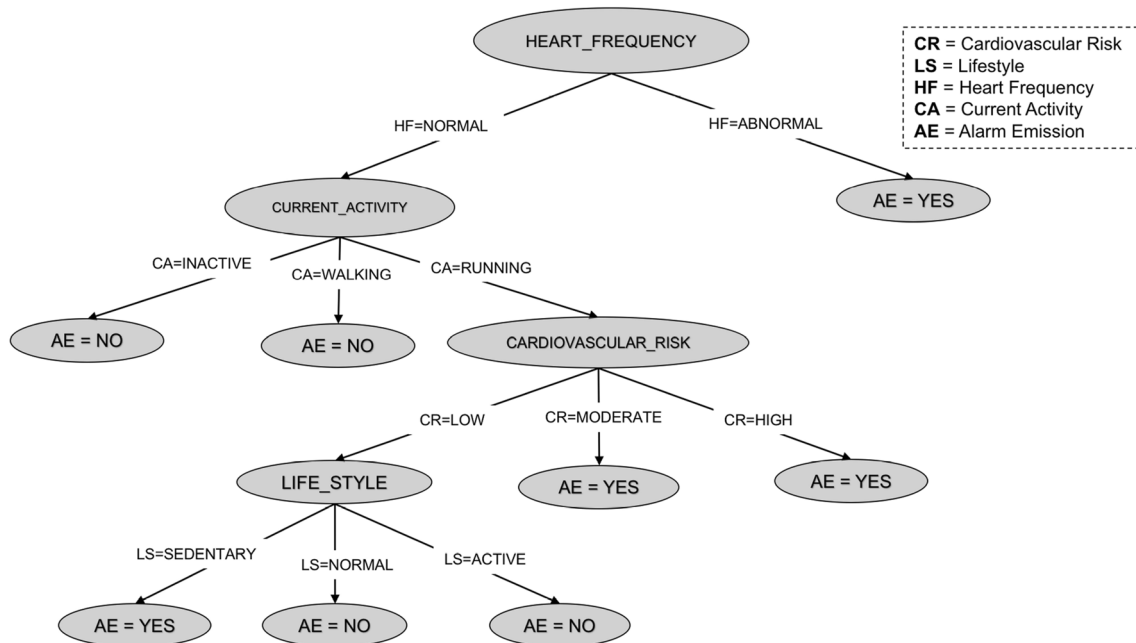


Fig. 4 ID3 decision tree using strict alarm emission

3.2 The Recommender System Stage

When a high-risk condition is detected, the patient requires assisting to a medical center for a partial or full medical revision. In this case, it is possible one of the next two scenarios: in the first one, the patient can easily identify the adequate medical center, because there is a priori knowledge that is represented by the ontology. In the second one, there is not a priori information about the most adequate hospital for attending the emergency of the patient. For example,

the nearest hospital does not have a medical doctor with a cardiology speciality; however, for this scenario, it is necessary that a cardiologist aids this emergency. According to the example, a recommender system stage is proposed. The system is in charged of performing the following tasks: recommendations, a personalization of answers, semantic information processing and a map rendering. Fig. 1 describes the recommender system, in order to guarantee the consistency in querying and retrieving processes, all information

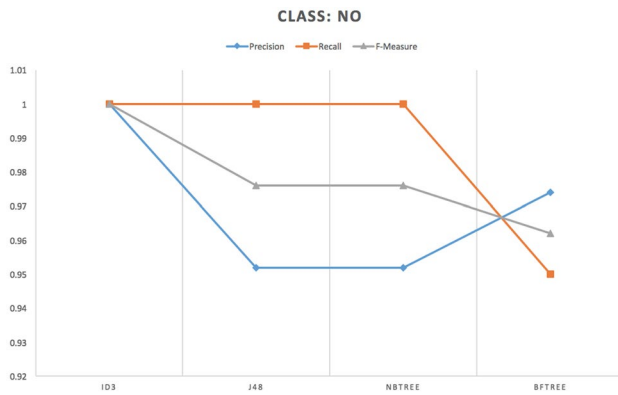


Fig. 5 The Precision, Recall and F measures when the alarm is not generated

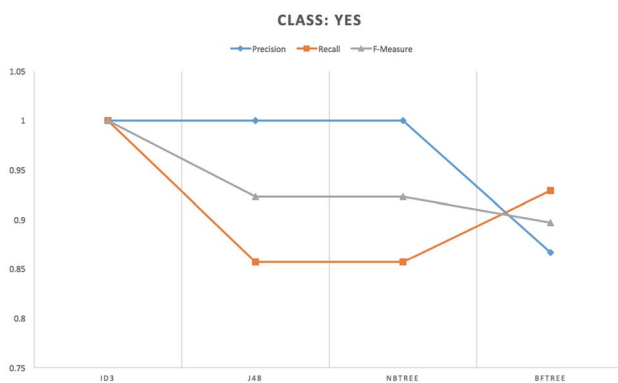


Fig. 6 The Precision, recall and F measures when the alarm is generated

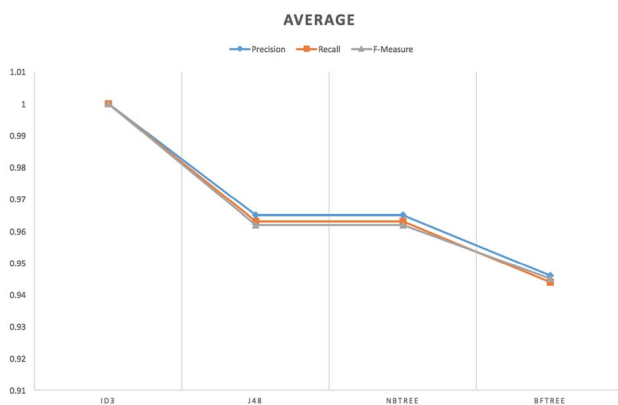


Fig. 7 Averages of the precision, recall and F measures

is processed by means of a medical ontology, which is presented in Fig. 8.

The goal of this ontology is to conceptualize and describe the most important medical specialities. This ontology was implemented by taken into consideration the Association of

American Medical Colleges Standard [51]. It is important to mention that there are other medical ontologies such as Ontology for Biomedical Investigations (OBI) [52], Gene Ontology [53], Generalized Architecture for Languages, Encyclopaedias and Nomenclatures in medicine (GALEN) [54], MedO [55], Systematized Nomenclature of Medicine (SNOMED) [56], among others. However, they are built in different languages like Open Biomedical Ontologies (OBO), Web Ontology Language (OWL) and Resource Description Framework (RDF), their structures are very complex, the content is around of 800,000 concepts, their query processing is slow, and medical specialities are not explicitly described. So, the proposed research only uses an application ontology as a hierarchy, with specific partitions related to medical specialities, only applying a *is-a* relationship. Thus, general and particular medical concepts, diseases, procedures, and other medical particularities are not necessary in this work. Moreover, the application of public medical ontologies could infer directly in the performance and semantic information integration approaches.

In other words, the proposed conceptualization was developed by efficiency purposes, single hierarchical organization, specialization level, and granularity, which fit with the proposed application that is to recommend medical specialities. The proposed ontology was modelled in OWL, using Protégé version 4.3, especially considering the semantic relation *is-a*. The root node contains three properties: the first one determines if the speciality is a diagnostic or therapeutic. The second one indicates the age range of patients (paediatric, adults, geriatric, all), and the third property describes if the speciality is organ-based, technique-based or both.

3.2.1 Personalization Task

The personalization task defines a user profile and the specification of the search type. The profile contains the following information:

- *Type of Medical Service* The user indicates if he requires public or private medical services.
- *Hospitals* All hospitals that satisfy the type of medical service are displayed.
- *Medical Specialities* They require to select the particular specialities for each hospital previously enlisted.

3.2.2 Information Processing Task

Information processing task receives the search type as an input parameter. By using this value, the system performs queries in the ontology by means of the SPARQL, which is an ontology query language. The purpose is to retrieve information of all hospitals that accomplish the user's criteria.

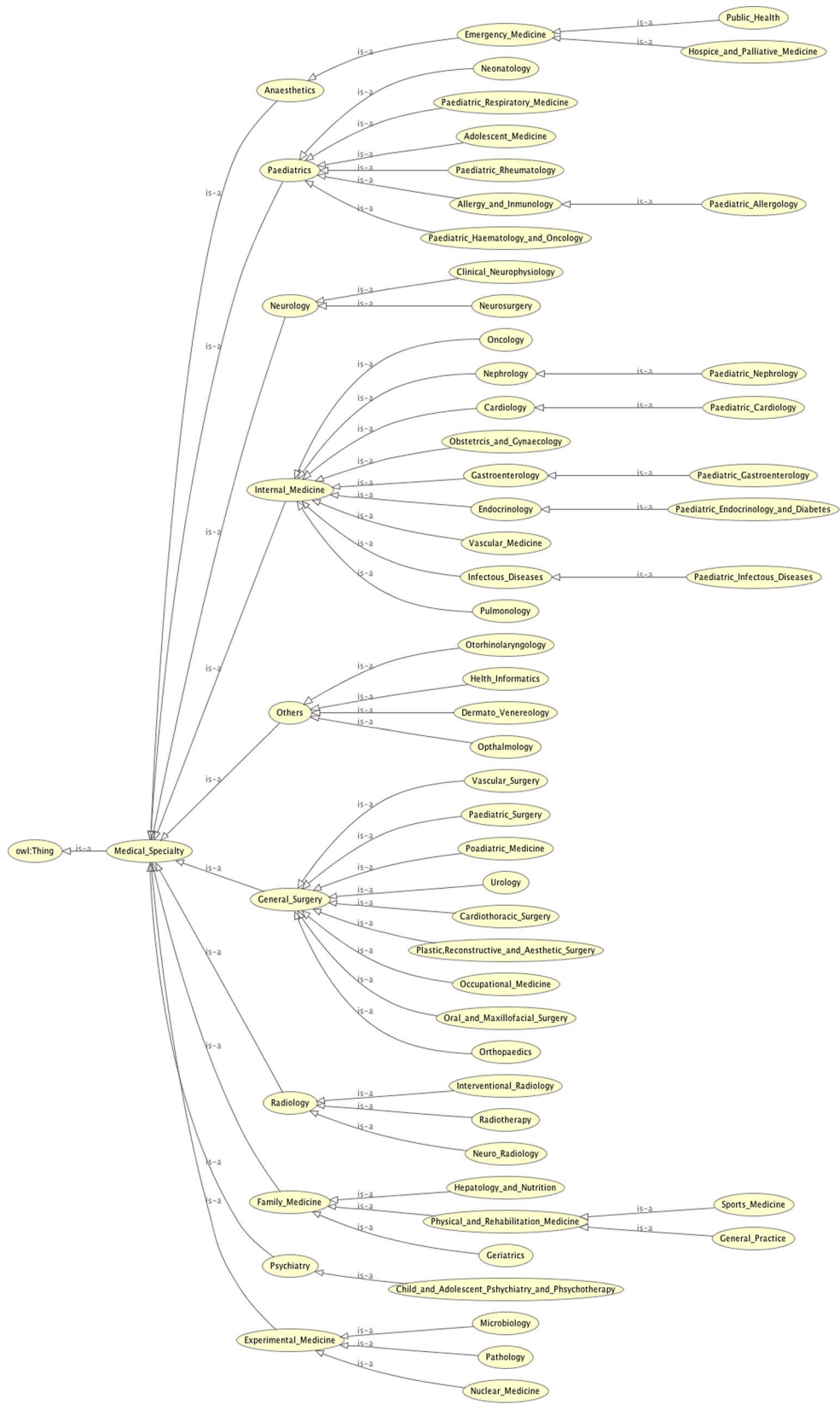


Fig. 8 Medical ontology to describe medical specialties used in the recommender system

The search type adopts one of three possible values: *general*, *emergency* or *oriented-profile*. In the general search, the list of retrieved hospitals are ordered according to the geographic distances measured by user’s location. In consequence, the first hospital is the nearest patient. In case of emergency, the “ER” string is assigned to the search type variable, the previous steps are performed, but the query result only contains hospitals that provide emergency service. Finally, in an oriented-profile query, all information stored in the user profile is employed to request the nearest hospitals with the medical specialties described in the profile. The final result of this process is a list of hospitals denoted by the set $H = \{h_1, h_2, \dots, h_n\}$.

3.2.3 The Ranking Algorithm for Computing the Attention Factor

In this task an algorithm is proposed to rank the attention factor. It consists of numerically quantifying the attention

user position $p(x, y)$, Ξ is the set of all medical specialties described by the ontology O , S_i is a subset of Ξ and denotes the number of medical specialties offered by some hospital $h_i \in H_R$, the value of $nmd(h_i)$ for some h_i is obtained by applying Eq. 3.

$$nmd(h_i) = \sum_{\forall s \in S_i} \text{doctors}(s, h_i) * \text{sim}(s, s_u), h_i \in H_R \tag{3}$$

where $\text{doctors}(s, h_i)$ denotes the number of available medical doctors with a speciality s in the hospital h_i , and $\text{sim}(s, s_u)$ determines the semantic similarity between specialties s and s_u . In order to compute the similarity, the Resnik semantic similarity measure was adopted [57]. However, it is possible to use other measures such as Rada et al. among others that are based on information content [58–60]. Algorithm 2 presents the method to rank the attention factor.

Input:
 $H_r = \{h_1, h_2, \dots, h_n\}$, set of disordered hospitals.
 s , medical specialty required.
 $ratio$, search ratio from user current location $p(x,y)$
 $patients(h_i) \forall h_i \in H_R$, number of patients in hospital h_i .

Output:
 $RH_r = \{rh_1, rh_2, \dots, rh_n\} \mid rf(rh_i) > rf(rh_j), i < j$, set of ranked hospitals.

```

1 for  $i = 1 \rightarrow n$  do
2    $p_1(x_1, y_1) \leftarrow patient\_geographic\_coordinates$ 
3    $p_2(x_2, y_2) \leftarrow hospital\_geographic\_coordinates, h_i$ 
4    $norm\_distance(h_i) = 1.0 - \frac{geographic\_distance(p_1, p_2)}{ratio}$ 
5    $nmd(h_i) = 0$ 
6   for  $j = 1 \rightarrow m \forall s_j \in S_i$  do
7      $nmd(h_i) = nmd(h_i) + \text{doctors}(s_j, h_i) * \text{sim}(s_j, s)$ 
8      $attention\_factor(h_i) = \frac{nmd(h_i)}{patients(h_i)}$ 
9      $rf(h_i) = norm\_distance(h_i) * attention\_factor(h_i)$ 
10  $RH_r = \{rh_1, rh_2, \dots, rh_n\}, \mid rf(rh_i) > rf(rh_j), i < j$ 

```

Algorithm 2: Ranking algorithm based on the attention factor

factor for a set of hospitals. In an emergency, suppose the case of a heart attack, the ideal condition is that a cardiologist care of patient. However, in the situation when this medical is not available, it is important to compute the similarity and find all possible doctors with some and similar knowledge, in order to treat the patient. In consequence, Eq. 3 is proposed to compute the number of doctors (nmd) that can aid some emergency related to some medical specialties s_u . The nmd determines the medical specialties that are semantically related to the required medical speciality, which is defined in the user profile.

Let $f : H_R \rightarrow S$ be a function, where H_R is the number of all hospitals $H_R = \{h_1, h_2, \dots, h_n\}$ inside a ratio R obtained by applying the buffer spatial operator from current

On the other hand, for computing the geographic distance, the Haversine function was applied [61], which is denoted by Eqs. 4–6.

$$geographic_distance(p_1(x_1, y_1), p_2(x_2, y_2)) = R \times c \tag{4}$$

$$a = \sin^2\left(\frac{\text{deg2rad}(\Delta_{lat})}{2}\right) + \cos(\text{deg2rad}(x_1)) \times \cos(\text{deg2rad}(x_2)) \times \sin^2\left(\frac{\text{deg2rad}(\Delta_{lon})}{2}\right) \tag{5}$$

$$c = \text{rad2deg}\left(2 \times \arcsin\left(\min(1, \sqrt{a})\right)\right) \tag{6}$$

where

$\Delta_{lon} = y_2 - y_1$ is the difference of longitude coordinates.

$\Delta_{lat} = x_2 - x_1$ is the difference of latitude coordinates.

$\text{deg2rad}(\text{value})$ is a function to convert value from degrees to radians.

$\text{rad2deg}(\text{value})$ is a function to convert value from radians to degrees.

R is the Earth radius in km, equal to 6378.137 m.

3.2.4 Rendering Task

The rendering task consists of interpreting the obtained results from the ranking algorithm, in order to display them on a map using Google Maps API version 2. This map server was used because processes like download, visualization, and response times for requests in mobile devices are optimal, in comparison with other map servers (i.e. Yahoo Maps, Microsoft Bing Maps, etc.). Additionally, it is possible to use a temporal layer to add different visual analysis options (markers, polygons, route lines, and change of map type). This task processes the input data that contains a geographic distance ordered (general and oriented-profile queries) or ranked (emergency search) list of hospitals, in order to add markers for all hospitals enlisted in the input list. Each hospital is represented by a geographic marker and a label that describes the name and related information of the medical

unit, the address and distance (measured in km) are also presented.

4 Tests and Results

In this section, the corresponding tests and obtained results are presented. The first discussion is related to the abnormal heart rate detection. The next section explains the results of the recommender system, which is focused on the emergency search query.

The experiments were performed on 21 patients, between 45 and 70 years old, in Cuernavaca, Mexico. All patients were in monthly control with his doctor. Unfortunately, a greater number of patients were not considered due to the hospital restrictions. The selected patients to participate in the experiments used a mobile device and the Zephyr H × M. Initially, the patients registered in the application their clinical data, such as their weight, height, blood pressure, among others. Subsequently, users select the option “Cardiovascular risk”, which is computed by considering their clinical data. The user can monitor their heart rate through the H × M Zephyr device and display it on the mobile device, then an alert can be issued, which is based on the analysis of ID3 decision tree. Patients used the mobile device and the Zephyr H × M in their everyday activities in order to continuously monitor their heart rate.

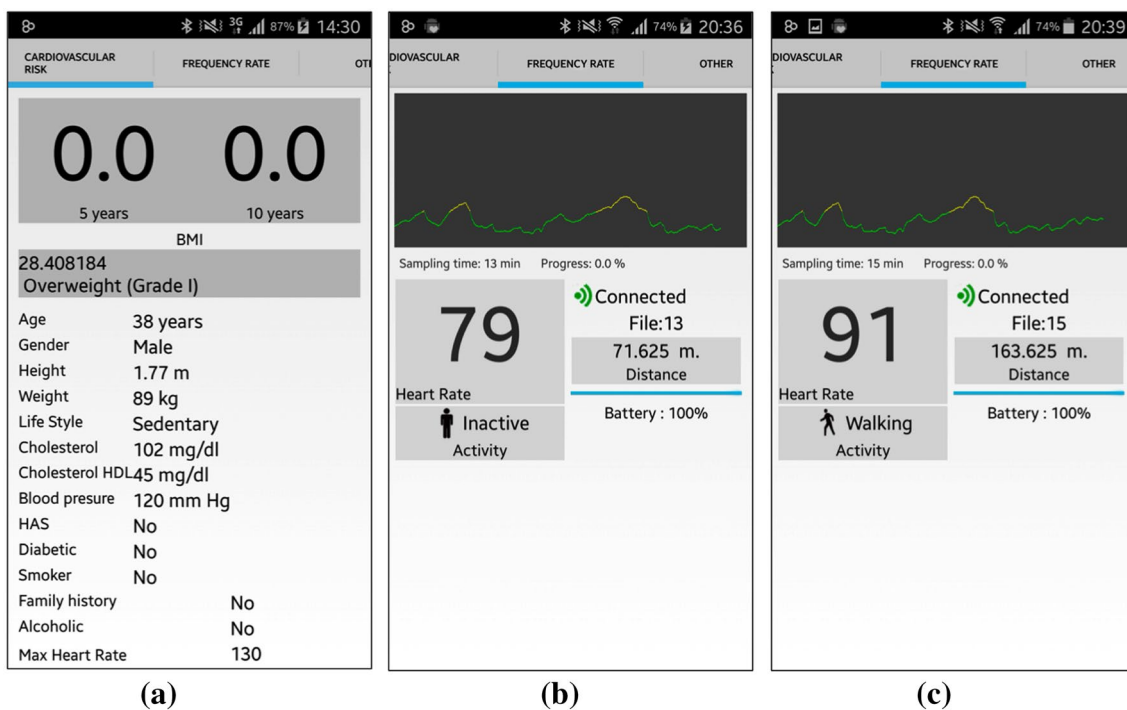


Fig. 9 Monitoring application **a** Cardiovascular risk **b** Heart frequency on inactive physical state **c** Heart frequency on walking activity

The main limitations of this research are centered in increasing the sample, because it is necessary to acquire more mobile devices to monitor medical data, the lack for establishing a crowdsensing method due to cultural and economical issues related directly to the region. The project started as a prototype and at this moment it is difficult to find other medical databases (at least in Mexico) for increasing the sample and performing additional tests, as well as comparing the outcomes of this approach with other methods that have been proposed in the the state of the art. However, the main contribution of this work is the methodological framework for monitoring data by means of a mobile device, the semantic processing to recommend medical specialities, and the ranking algorithm for computing the attention factor.

Moreover, the goals of the research are not focused on having a large dataset, because the proposed methodology is based on indexes such as Framingham and SCORE, which has collected large samples of patients in order to estimate the cardiovascular risk that could be applied in the worldwide. Therefore, the application of these indexes guarantee the accuracy in the computation of a high risk cardiovascular condition that is suggested in this research. In addition the

ID3 decision tree was completed to show all possible cases. It is important to mention that there is public dataset such as used by [49], but they do not combine variables that describe the current activity and maximal heart rate.

4.1 Monitoring Task

4.1.1 Cardiovascular Risk Computation

It is necessary to select the “cardiovascular risk” option in the mobile application, an interface similar to the screenshot presented in Fig. 9a is loaded to the user. This interface contains information related to the computation of cardiovascular risk, according to the medical profile provided by the application.

4.1.2 Heart Rate Monitoring

The option “heart frequency” is associated with the display of the heart rate values that were obtained from the monitoring device. The form has a graphical representation of the heart rate behavior, type of physical activity, remaining

Table 5 Equivalence class for the cardiovascular risk computation

Input condition	Data type	Valid equivalence class	No valid equivalence class
Age	Numerical	(1) Age > 0, (2) Age < 120	(3) No numerical value, (4) Age < 0, (5) Age > 120
Genre	Set	(6) Female, (7) Male	(8) Empty string and all different of valid strings
Diabetic	Logical	(9) Yes, (10) No	(11) Different value
Smoker	Logical	(12) Yes, (13) No	(14) Different value
Diabetic	Range	(15) Cholesterol > 0 (16) Cholesterol < 300	(17) No numerical value (18) Cholesterol < 0 (19) Cholesterol > 300
HDL ^a	Range	(20) HDL > 0 (21) HDL < 300	(22) No numerical value (23) HDL < 0 (24) HDL > 300
SBP ^b	Range	(25) SBP > 0 (26) SBP < 300	(27) No numerical value (28) SBP < 0 (29) SBP > 300
DBP ^c	Range	(30) DBP > 0 (31) DBP < 300	(32) No numerical value (33) DBP < 0 (34) DBP > 300
ECG ^d	Boolean	(35) Yes (36) No	(37) Different value

^aHigh Density Lipoprotein

^bSystolic Blood Pressure

^cDiastolic Blood Pressure

^dElectrocardiogram

Table 6 Test cases for the the cardiovascular risk computation

Case	Equivalence class	Values	Result
1	1, 6, 9, 12, 15, 20, 25, 30, 35	Age: 58, Genre: Female, Diabetic: Yes, Smoker: Yes, Cholesterol: 158, HDL: 95, SBP: 100, DBP: 60, ECG: Yes,	Computation of the cardiovascular risk percentage is determined
2	3, 8, 11, 14, 17, 22, 27, 32, 37	Age: x, Genre: Nurse, Diabetic: 14, Smoker: null, Cholesterol: null, HDL: null, SBP: null, DBP: z, ECG: null,	Computation of the cardiovascular risk percentage is not determined
3	2, 7, 10, 13, 16, 21, 26, 31, 36	Age: 58, Genre: Male, Diabetic: No, Smoker: No, Cholesterol: 158, HDL: 95, SBP: 100, DBP: 60, ECG: No,	Computation of the cardiovascular risk percentage is determined
4	4, 6, 9, 12, 18, 23, 28, 33, 36	Age: - 8, Genre: Female, Diabetic: Yes, Smoker: Yes, Cholesterol: - 20, HDL: - 4, SBP: - 100, DBP: - 20, ECG: No,	Computation of the cardiovascular risk percentage is not determined
5	5, 8, 11, 12, 19, 24, 29, 34, 35	Age: 58, Genre: Pencil, Diabetic: x, Smoker: Yes, Cholesterol: 458, HDL: 395, SBP: 400, DBP: 600, ECG: Yes,	Computation of the cardiovascular risk percentage is determined

Table 7 Equivalence classes for the quantification of the decision-making

Input condition	Type	Equivalence class valid	Equivalence class not valid
Heart frequency (HF)	Range	(1) HF > 30, (2) HF < 240	(3) No numerical value, (4) HF < 30, (5) HF > 240
Lifestyle	Set	(6) "Sedentary", (7) "Normal", (8) "Active"	(9) Empty string or different to valid options
Physical activity (PA)	Range	(10) PA > 0, (11) PA < 19	(12) No numerical value, (13) PA < 0, (14) PA > 19
Cardiovascular risk (CR)	Range	(15) CR > 0, (16) CR < 100	(17) No numerical value, (18) CR < 0, (19) CR > 100

Table 8 Equivalence classes for the decision-making tree

Case	Equivalence class	Values	Result
1	1, 6, 10, 15	HF: 80, LS: Sedentary, PA: 4.0, CR: 0.3	Decision-making is applied
2	2, 7, 11, 16	HF: 120, LS: Normal, PA: 11.0, CR: 0.3	Decision-making is applied
3	3, 9, 12, 17	HF: null, LS: total, PA: z, CR: y	Error, Decision-making is aborted
4	4, 8, 13, 16	HF: 20, LS: Active, PA: - 3, CR: 30	Error, Decision-making is aborted
5	5, 7, 11, 18	HF: 280, LS: Normal, PA: 8.0, CR: - 3	Error, Decision-making is aborted

Table 9 Similarity distances obtained by using the Resnik measure

Speciality	Oncology	Nephrology	Cardiology	Urology	Orthoapedics	Neonatology	Internal medicine
Oncology	1.00	0.75	0.75	0.50	0.50	0.50	0.90
Nephrology	0.75	1.00	0.75	0.50	0.50	0.50	0.90
Cardiology	0.75	0.50	1.00	0.50	0.50	0.50	0.90
Urology	0.50	0.50	0.50	1.00	0.75	0.50	0.30
Orthoapedics	0.50	0.50	0.50	0.75	1.00	0.50	0.30
Neonatology	0.50	0.50	0.50	0.50	0.50	1.00	0.30
Internal medicine	0.90	0.90	0.90	0.30	0.30	0.30	1.00

battery charge, among others. The first time the application is run, a service is registered in the mobile device, in order to maintain the monitoring process, even though the application

is not running in foreground. Figure 9b depicts graphic information that describes the inactivity or physical activity state, and Fig. 9c shows a walking activity monitoring.

Table 10 Number of doctors used in the second scenario

Hospital	Oncology	Nephrology	Cardiology	Urology	Orthopaedics	Neonatology	Internal medicine
Hospital 1	1.00	0.00	2.00	1.00	0.00	0.00	2.00
Hospital 2	0.00	0.00	1.00	1.00	1.00	1.00	1.00
Hospital 3	1.00	1.00	2.00	1.00	0.00	1.00	2.00
Hospital 4	0.00	2.00	1.00	1.00	0.00	1.00	1.00

Table 11 Results by applying ranking algorithm for computing the attention factor

Hospital	Latitude	Longitude	Patients	Distance (Km)	nmd	Attention factor	Ranking
Hospital 1	12.35	22.36	10.00	0.45	5.05	0.51	0.23
Hospital 2	13.46	23.46	12.00	0.26	3.40	0.28	0.07
Hospital 3	13.72	20.19	12.00	0.53	6.30	0.53	0.28
Hospital 4	14.36	21.53	11.00	0.37	4.40	0.40	0.40

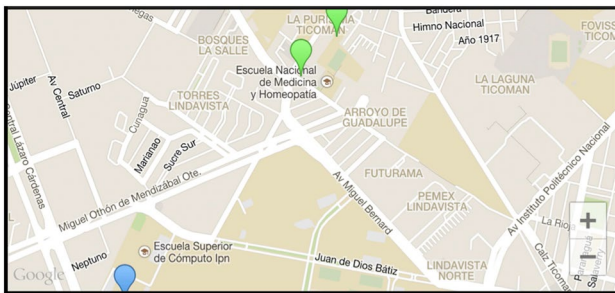


Fig. 10 Result of the general search

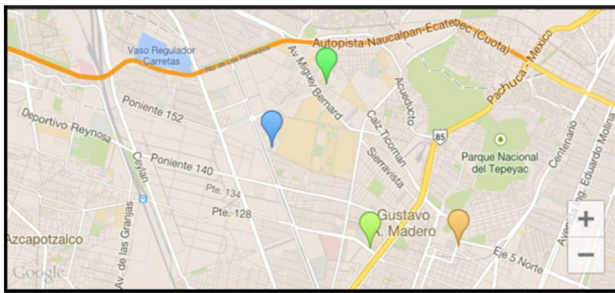


Fig. 11 Result of the emergency query

1.61km	Hospital_General_Ticomán Plan de San Luis 7 La Purísima Ticoman, Gustavo A. Madero 07320 Ciudad de México, D.F.
2.6km	Hospital_Angeles_Lindavista Río Bamba 639 Magdalena de Las Salinas, Gustavo A. Madero 07760 Ciudad de México, D.F.
3.93km	Hospital_Pediatrico_la_Villa Hidalgo 200 Villa Gustavo a Madero, Gustavo A. Madero Ciudad de México, D.F.
4.12km	Hospital_Materno_Infantil_Cuauhtpec Emiliano Zapata Zona Escolar Oriente, Gustavo A. Madero 07320 Ciudad de México, D.F.
5.27km	Hospital_General_la_Villa Eje 5 Norte 259 Granjas Modernas, Gustavo A. Madero 07460 Ciudad de México, D.F.
6.13km	Hospital_Central_Norte Campo Matillas 32 San Antonio, Azcapotzalco 02760 Ciudad de México, D.F.

Fig. 12 Final result in emergency query using ranking algorithm

4.1.3 Equivalent Partition

It is based on making a partition of the input field into a set of data classes, commonly denoted as equivalence classes. The equivalence classes are defined as valid or invalid, the main difference in this classification is that an expected value is generated by a valid class, whereas a non-expected value is defined by a non-valid class.

4.1.3.1 Equivalence Class for the Cardiovascular Risk Computation It consists of the estimation of the cardiovascular risk, and requires all input conditions in order to obtain the Framingham index. Table 5 describes the equivalence classes obtained in the cardiovascular risk analysis.

4.1.3.2 Test Cases for the Cardiovascular Risk Computation Table 6 presents five test cases that were used to determine the cardiovascular risk of the abnormal heart rate stage.

4.1.3.3 Equivalence Classes for the Decision Tree The input conditions for the equivalence classes are integrated by the node members of the decision tree. In Table 7 the

Table 12 Specifications of the mobile devices

Model	Android version	CPU	Internal memory	Screen size	Time consumed (s)
Samsung S6	Android 7.0	4 × 2.1 GHz Cortex-A57 (Octacore)	32 GB (3 GB RAM)	5.1" (1440 × 2560 pixels)	8.03
Sony Xperia Z5	Android 7.1.1	4 × 1.5 GHz Cortex-A53 (Octacore)	32 GB (3 GB RAM)	5.2" (1080 × 1920 pixels)	8.72
Motorola G4	Android 7.0	4 × 1.5 GHz Cortex-A53 (Octacore)	32 GB (2GB RAM)	5.5" (1080 × 1920 pixels)	9.12
LG G5	Android 7.0	2 × 2.15 GHz Kryo (Quadcore)	32 GB (2 GB RAM)	5.3" (1440 × 2560 pixels)	11.86
Samsung S5 mini	Android 5.1.1	2 × 1.4 GHz Cortex-A7 (Quadcore)	16 GB (1.5 GB RAM)	4.5" (720 × 1280 pixels)	13.37

corresponding equivalence classes for the decision tree are described.

4.1.4 Test Cases for the Estimation of the Decision Tree

Table 8 presents the test cases used in the estimation for the decision tree. As it can be seen, only three cases out of twenty-one, the system did not estimate the risk condition of the patient. In this case, the effectiveness of decision tree was 85.71%.

4.2 The Recommender System Stage

In this section, the obtained results in the recommender system stage are shown. For discussion purposes, two scenarios are explained. Initially, in some geographic location, the user requests a *general* search. So, in Fig. 10 the result is depicted, the blue marker denotes the user location, whereas the other markers represent the closest hospitals to the user's location. The second test considers an *emergency condition* as query type; in consequence, the hospitals are arranged by using the proposed ranking algorithm based on the attention factor approach (see Fig. 11). In Tables 9, 10 and 11 the results of applying the ranking algorithm for computing the attention factor are described. The final result of the recommender system stage is shown in Fig. 12.

The time required to detect a high risk cardiovascular condition of a patient and determine the appropriate hospital to attend, depends on the mobile device. Table 12 shows the technical details of the 5 mobile models that were used to perform the tests. According to Table 12 the Samsung S6 was the model that obtained the better performance, with 8.03 s. The rest of results related to the other mobile models is presented in Table 12. So, when a high risk condition is detected, the recommender system proposes a set of ranked hospitals to take care of the emergency.

5 Conclusions and Future Work

In this paper, a collaborative architecture and methodological framework for sensing abnormal heart rate based on a recommender system is proposed. The heart rate monitoring allows a continuous supervision of the heart frequency. This information combined with a priori data, and described in a medical user-profile, is processed by using an ID-3 decision tree, in order to determine a high risk cardiovascular condition.

The main contribution presented in this paper is related to the recommender system stage in which an attention factor approach was proposed. It estimates which hospital or medical unit with better conditions (lower number of patients at the emergency moments and higher number of medical specialists), by means of the semantic similarity computation between medical specialties. The ranking of hospitals that meets those criteria are visualized in a web-mapping application. According to the outcomes, almost immediately, the patient can obtain an alert and get the recommendation of a hospital to provide the proper care to increase the chances of saving the patients life.

On the other hand, by using the Bluetooth Health Device Profile (HDP), any wireless sensor that satisfies this specification can be used with our mobile application to obtain other physiological variables of the patient such as glucose level, pulse oxymetry, among others.

Future works are oriented towards implementing a system that takes into consideration the expertise of medical institutions and collects data by volunteered geographic information, to analyze and obtain historical information about the use and advantages of our approach. In addition, the sample will be enriched by searching public datasets or establishing a crowd-sensing approach to add more data related to the heart rate frequency. Limitations associated with the energy consumption will be attended to guarantee the efficient use of computing resources. Applications to replicate experiments with other machine learning techniques should be considered for future research.

Acknowledgements This work was partially sponsored by the Instituto Politécnico Nacional (IPN), Consejo Nacional de Ciencia y Tecnología (CONACYT) under grant PN-2016/2110, and the Secretaría de Investigación y Posgrado (SIP) under Grants Nos. 20180308, 20181568, 20182159 and 20180409. Additionally, we are thankful to the reviewers for their invaluable and constructive feedback that helped improve the quality of the paper.

References

- United Nations. (2010). *Department of Economic. World population ageing 2009*. New York: United Nations Publications.
- World Health Organization. (2009). *Global status report on road safety: time for action*. Geneva: World Health Organization.
- Beaglehole, R., Epping-Jordan, J., Patel, V., Chopra, M., Ebrahim, S., Kidd, M., et al. (2008). Improving the prevention and management of chronic disease in low-income and middle-income countries: A priority for primary health care. *The Lancet*, 372(9642), 940–949.
- World Health Organization et al. (2011) *The top 10 causes of death*. Geneva: World Health Organization.
- Kendall, K. E., Kendall, J. E. (2007). *Systems analysis and design*, (Vol. 7). New Jersey: Prentice-Hall.
- Mihailidis, A., & Bardram, J. E. (2010). *Pervasive computing in healthcare*. Boca Raton: CRC Press.
- Noguera, J. M., Barranco, M. J., Segura, R. J., & Martínez, L. (2012). A mobile 3D-GIS hybrid recommender system for tourism. *Information Sciences*, 215, 37–52.
- Kumar, S., Kambhatla, K., Fei, H., Lifson, M., & Xiao, Y. (2008). Ubiquitous computing for remote cardiac patient monitoring: A survey. *International Journal of Telemedicine and Applications*, 2008, 3.
- Griffiths, E., Saponas, T. S., & Brush, A. J. (2014). Health chair: Implicitly sensing heart and respiratory rate. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 661–671). ACM.
- Hamedani, K., Bahmani, Z., & Mohammadian, A. (2016). Spatio-temporal filtering of thermal video sequences for heart rate estimation. *Expert Systems with Applications*, 54, 88–94.
- Ravichandran, R., Rahman, T., Adams, A., Choudhury, T., Kientz, J., & Patel, S. (2015). Real time heart rate and breathing detection using commercial motion sensors. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers* (pp. 325–328). ACM.
- Tamura, T., Maeda, Y., Sekine, M., & Yoshida, M. (2014). Wearable photoplethysmographic sensors-past and present. *Electronics*, 3(2), 282–302.
- Tu, L., Huang, J., Bi, C., & Xing, G. (2017). Fitbeat: A lightweight system for accurate heart rate measurement during exercise. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 1–8). IEEE.
- Temko, A. (2017). Accurate heart rate monitoring during physical exercises using PPG. *IEEE Transactions on Biomedical Engineering*, 64(9), 2016–2024.
- Chai, J. (2013). The design of mobile ECG monitoring system. In *2013 IEEE 4th International Conference on Electronics Information and Emergency Communication (ICEIEC)* (pp. 148–151). IEEE.
- Alzate, E. B., & Martinez, F. M. (2010). ECG monitoring system based on ARM9 and mobile phone technologies. In *2010 IEEE on ANDESCON* (pp. 1–6). IEEE.
- Kai, L., Zhang, X., Wang, Y., Suibiao, H., Ning, G., & Wangyong, P., et al. (2011). A system of portable ecg monitoring based on bluetooth mobile phone. In *2011 International Symposium on IT in Medicine and Education (ITME)* (Vol. 2, pp. 309–312). IEEE.
- Qidwai, U., Chaudhry, J. A., Shakir, M., & Rittenhouse, R. G. (2012). Ubiquitous monitoring system for critical cardiac abnormalities. In *Computer Applications for Bio-technology, Multimedia, and Ubiquitous City* (pp. 124–134). Berlin: Springer.
- Patel, A. M., Gakare, P. K., & Cheeran, A. N. (2012). Real time ECG feature extraction and arrhythmia detection on a mobile platform. *International Journal of Computer Applications*, 44(23), 40–45.
- Gradl, S., Kugler, P., Lohmuller, C., & Eskofier, B. (2012). Real-time ECG monitoring and arrhythmia detection using android-based mobile devices. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2452–2455). IEEE.
- Pandey, S., Voorsluys, W., Niu, S., Khandoker, A., & Buyya, R. (2012). An autonomic cloud environment for hosting ECG data analysis services. *Future Generation Computer Systems*, 28(1), 147–154.
- Salvador, C. H., Carrasco, M. P., De Mingo, M. G., Carrero, A. M., & Montes, J. M., et al. (2005). Airmed-cardio: A GSM and internet services-based system for out-of-hospital follow-up of cardiac patients. *IEEE Transactions on Information Technology in Biomedicine*, 9(1), 73–85.
- Jones, V., Van Halteren, A., Dokovsky, N., Koprnikov, G., Peuscher, J., & Bults, R., et al. (2006). Mobihealth: Mobile services for health professionals. In *M-Health* (pp. 237–246). Boston: Springer.
- Shih, D.-H., Chiang, H.-S., Lin, B., & Lin, S.-B. (2010). An embedded mobile ECG reasoning system for elderly patients. *IEEE Transactions on Information Technology in Biomedicine*, 14(3), 854–865.
- Ren, Y., Werner, R., Pazzi, N., & Boukerche, A. (2010). Monitoring patients via a secure and mobile healthcare system. *IEEE Wireless Communications*, 17(1), 59–65.
- Gay, V., & Leijdekkers, P. (2007). A health monitoring system using smart phones and wearable sensors. *International Journal of ARM*, 8(2), 29–35.
- Bao, X., Chen, X., Fang, Z., & Xia, S. (2016). Design and development of a ubiquitous healthcare monitoring system based on android platform. *DEStech Transactions on Engineering and Technology Research*, (ICMITE2016).
- Mohamed, R., & Youssef, M. (2017). Heartsense: Ubiquitous accurate multi-modal fusion-based heart rate estimation using smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3), 97.
- Novo, J., Hermida, A., Ortega, M., Barreira, N., & Penedo, M., et al. (2017). A web-based system for cardiovascular analysis, diagnosis and treatment. *Computer Methods and Programs in Biomedicine*, 139, 61–81.
- Epping-Jordan, J. E., Pruitt, S. D., Bengoa, R., & Wagner, E. H. (2004). Improving the quality of health care for chronic conditions. *Quality and Safety in Health Care*, 13(4), 299–305.
- Paganelli, F., & Giuli, D. (2007). An ontology-based context model for home health monitoring and alerting in chronic patient care networks. *AINA Workshops*, 2, 838–845.
- Appventive LLC. (2014). In case of emergency (ICE).
- Bottazzi, D., Corradi, A., & Montanari, R. (2006). Context-aware middleware solutions for anytime and anywhere emergency assistance to elderly people. *IEEE Communications Magazine*, 44(4), 82–90.
- Wang, S., Ji, L., Li, A., & Jiankang, W. (2011). Body sensor networks for ubiquitous healthcare. *Journal of Control Theory and Applications*, 9(1), 3–9.

35. Yuan, B., & Herbert, J. (2011). Web-based real-time remote monitoring for pervasive healthcare. In *2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)* (pp. 625–629). IEEE.
36. Soto, J. (2011). Plataforma de geolocalización de centros de salud con tecnología móvil implementando el protocolo de comunicación hl7. *TELEMATIQUE*, 9(3), 79–101.
37. Banerjee, S., & Mitra, M. (2014). A cross wavelet transform based approach for ECG feature extraction and classification without denoising. In *2014 International Conference on Control, Instrumentation, Energy and Communication (CIEC)* pp. 162–165. IEEE.
38. Zeybekoglu, S., & Özkan, M. (2010). Classification of ECG arrhythmia beats with artificial neural networks. In *2010 15th National Biomedical Engineering Meeting (BIYOMUT)* (pp. 1–4). IEEE.
39. Arif, M., Malagore, I. A., & Afsar, F. A. (2010). Automatic detection and localization of myocardial infarction using back propagation neural networks. In *2010 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)* (pp. 1–4). IEEE.
40. Chen, S., Hua, W., Li, Z., Li, J., & Gao, X. (2017). Heartbeat classification using projected and dynamic features of ECG signal. *Biomedical Signal Processing and Control*, 31, 165–173.
41. Fatin, A., Salim, N., Harris, A. R., Swee, T. T., & Ahmed, T. (2016). Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Computer Methods and Programs in Biomedicine*, 127, 52–63.
42. Shadmand, S., & Mashoufi, B. (2016). A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization. *Biomedical Signal Processing and Control*, 25, 12–23.
43. Tan, W. W., Foo, C. L., & Chua, T. W. (2007). Type-2 fuzzy system for ECG arrhythmic classification. In *FUZZ-IEEE 2007. IEEE International on Fuzzy Systems Conference* (pp. 1–6). IEEE.
44. Dumont, J., Hernández, A. I., Fleureau, J., & Carrault, G. (2008). Modelling temporal evolution of cardiac electrophysiological features using hidden semi-markov models. In *30th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, 2008. EMBS 2008* (pp. 165–168). IEEE.
45. Yang, H., Negishi, K., Otahal, P., & Marwick, T. H. (2015). Clinical prediction of incident heart failure risk: A systematic review and meta-analysis. *Open Heart*, 2(1), e000222.
46. Luz, E. J. D. S., Schwartz, W. R., Cámara-Chávez, G., & Menotti, D. (2016). ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine*, 127, 144–164.
47. Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, 37(1), 153–156.
48. Wilmore, J. H., & Costill, D. L. (2004). *Fisiología del esfuerzo y del deporte*. Editorial Paidotribo.
49. Álvarez Cosmea, A. (2001). Las tablas de riesgo cardiovascular. una revisión crítica. *Medifam*, 11, 122–139.
50. Alcocer, L. A., Lozada, O., Fanghanel, G., Sánchez-Reyes, L., & Campos-Franco, E. (2011). Estratificación del riesgo cardiovascular global. comparación de los métodos framingham y score en población mexicana del estudio prit. *Cir Cir*, 79(1), 168–174.
51. Association of American Medical Colleges. (2015). Careers in medicine.
52. Peters, B., & OBI Consortium (2009). Ontology for biomedical investigations.
53. Gene Ontology Consortium. (2006). The gene ontology (go) project in 2006. *Nucleic Acids Research*, 34(suppl 1), D322–D326.
54. Rector, A., & Rogers, J. (2006). Ontological and practical issues in using a description logic to represent medical concept systems: Experience from galen. *Reasoning Web* (pp. 197–231).
55. Fujita, H., Hakura, J., & Kurematsu, M. (2010). Multiviews ontologies alignment for medical based reasoning ontology based reasoning for VDS. In *2010 11th International Symposium on Computational Intelligence and Informatics (CINTI)* (pp. 15–22). IEEE.
56. Cornet, R., & de Keizer, N. (2008). Forty years of snomed: A literature review. *BMC Medical Informatics and Decision Making*, 8(1), S2.
57. Resnik, P. (2011). Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language. *arXiv preprint arXiv: 1105.5444*.
58. Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man and Cybernetics*, 19(1), 17–30.
59. Jiang, J. J., & Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. In *Proceedings of the international conference on research in computational linguistics* (pp. 19–33).
60. Wu, Z., & Palmer, M. (1994). Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics* (pp. 133–138). Association for Computational Linguistics.
61. Sinnott, R. W. (1984). Virtues of the haversine. *Sky and Telescope*, 68, 158.