

Full Length Research Paper

Developing of a learning-based system to assist treatment process of arrhythmia patients

Adnan Fatih Kocamaz^{1*}, Fatih Aydın¹ and Erdem Uçar²

¹Technical Science Vocational Collage, Kırklareli University, Turkey.

²Computer Engineering Department, Trakya University, Turkey.

Accepted 22 November, 2011

The transformation process of the data kept in data warehouse into usable information for decision support system is very important. In this process, it is necessary to reveal the useful data that can meet the needs of the developed system. For this purpose, the study of data mining should be done on data warehouse. In this article, the phases of developing and applying a learning-based system in order to be used in the treatment process of arrhythmia patients were presented. Firstly, the data collected in the data warehouse was transformed into information through using data mining methods. In the data mining studies, the nearest neighbour algorithm (kNN) which is one of the machine learning algorithms was used. The process of selecting training data among the data stored in database was realized by a specialist through a web-based practice. In the process of composing training data, the evaluation results of the system about learning level were shown to the specialist in an online way. Thus, the generalizing power of the model made up by a classifier on training data was measured. Besides, to determine and filter the noisy data in the data warehouse, the quality of the signal taken from the patient and the evaluations of the specialist were used. In this way, the classifier used was contributed to form a suitable hypothesis on training data. By finding the proper hypothesis, the critical situations of the patient are conveyed to the doctor of the patient by the improved system within seconds. Thus, the patient is provided to be constantly kept under supervision free from location.

Key words: Cardiac arrhythmia, inductive expert systems, k-nearest neighbor, machine learning, tele medicine.

INTRODUCTION

Heart diseases are the leading cause of death worldwide (Tantimongcolwat et al., 2008). For this reason, a large number of investigations have been carried out all over the world. It is seen that the achieved researches are generally directed to the detection of heart diseases (Tantimongcolwat et al., 2008; Kukar et al., 1999); their analysis or treatment (Krstacic et al., 2001, 2006; Kukar et al., 1999; Brüser et al., 2010; Galvin et al., 2011). The analysis and treatment process of heart disease may sometimes take a very long time. During this process, the patient may be required to be constantly kept under supervision. Therefore, one way to

better manage cardiac diseases is to observe the affected patients at home over longer periods of time (Brüser et al., 2010). For the treatment of the cardiac patients who are at home or a different location, to do necessary measurements, there is a need for a monitoring device. There are various methods developed for constructing such devices (Koivistoinen et al., 2004; Alihanka et al., 1981; Brink et al., 2006; Shirouzu et al., 2001). Some of these methods were developed to solve the problems stemming from the physical situation of the patient. Because of the fact that heart diseases are among the prominent cause of death all over the world (World Health Organization, 2004), the data taken from the monitoring of a cardiac patients who is perpetually or periodically kept under supervision must be real time data. So, monitoring must be mobilized. Wireless and

*Corresponding author. E-mail: faith.kocamaz@kirklareli.edu.tr.

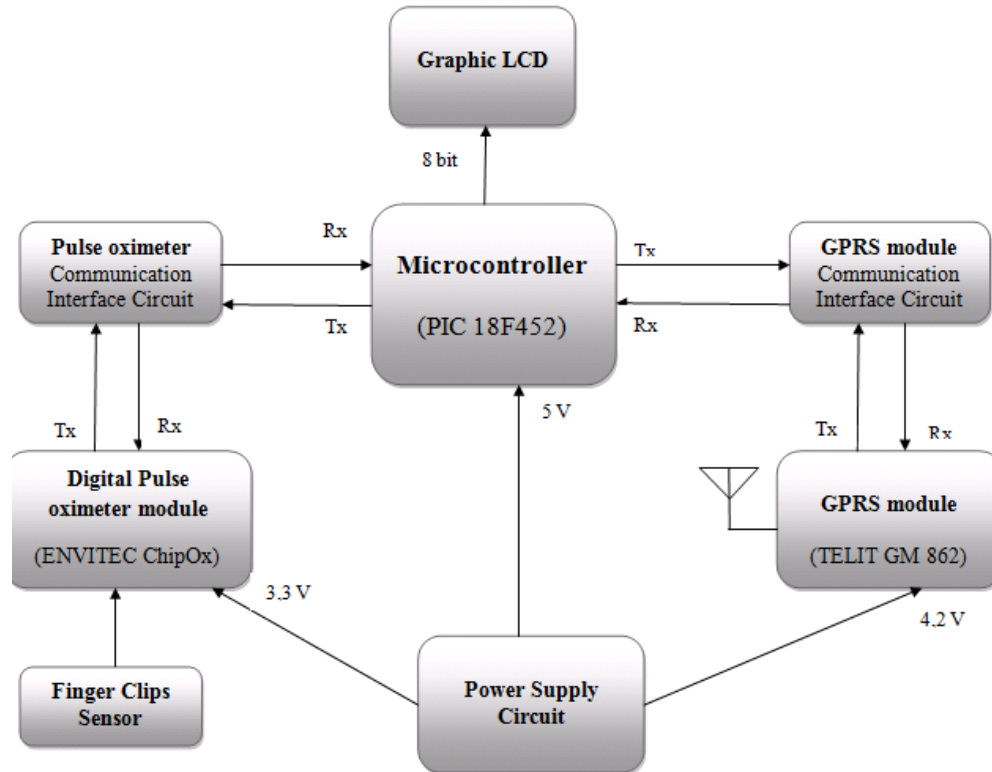


Figure 1. Mobile device blok diagram overview.

portable devices promise to support more convenient interactions (Patel et al., 2009). In this way, the collection of real time data by the computer systems in data warehouses and their evaluation are achieved. A data warehouse is a large collection of integrated data, built to assist the decision-making process (Benitez-Guerrero et al., 2004). While the data warehouses are built from operational data bases, it is required to regulate all data in such a way that they will be related to work flow process, be certain and belong to a certain period of time (Benitez-Guerrero et al., 2004). After this phase, data analysis and questionings such as OLAP (Codd et al., 1993) or data mining (Witten et al., 2005) that will assist the decision-making are made. Data mining is defined as the process of discovering patterns in data (Witten et al., 2005).

The main aim in this algorithm is to explore the appearance withing the data. The algorithms used in data mining are also known as machine learning algorithm (Witten et al., 2005; Alpaydin, 2004). For the operation of finding pattern on the data cluster, many techniques are benefitted on (Agrawal et al., 1993). One of these techniques is the classification method. There is a wide number of algorithms. At the same time, application of machine learning methods to large databases is called data mining (Alpaydin, 2004). In knowledge discovery, machine learning is most commonly used to mean the application of induction

algorithms which is one step in the knowledge discovery process (Kohavi et al., 1998). In this study, the phases of development and application of a learning based system directed to the treatment process of arrhythmia patients. In the same time, the data collection process necessary for the system, and which qualities were selected and why they were selected were explained. Later, the reasons for usage of the selected machine learning algorithms were examined. After this, the reasons for the use of learning strategies and statistical evaluation criteria chosen to decide whether the system learnt or not was explained.

MATERIALS AND METHODS

This study was developed for use in Trakya University department of cardiovascular surgery. The purpose of the system is to keep the patients under observation outside the hospital. Therefore, the developed system consists of two parts. The first part includes the design of a mobile device which is movable with the patient and can send the patient's data to the server. The second part consists of patient data collection and evaluation. This part is used to identify critical situations and includes patient data which is used as training data. At first step, mobile device which is carried with cardiac patient was designed. The components of mobile devices are microcontroller (PIC 18F452), ENVITEC Digital Pulse Oximeter Module ChipOx (ENVITEC, 2008), TELIT GM862 module (TELIT, 2011). Mobile device block diagram is shown in Figure 1. The second part of study consists of two software applications. One of

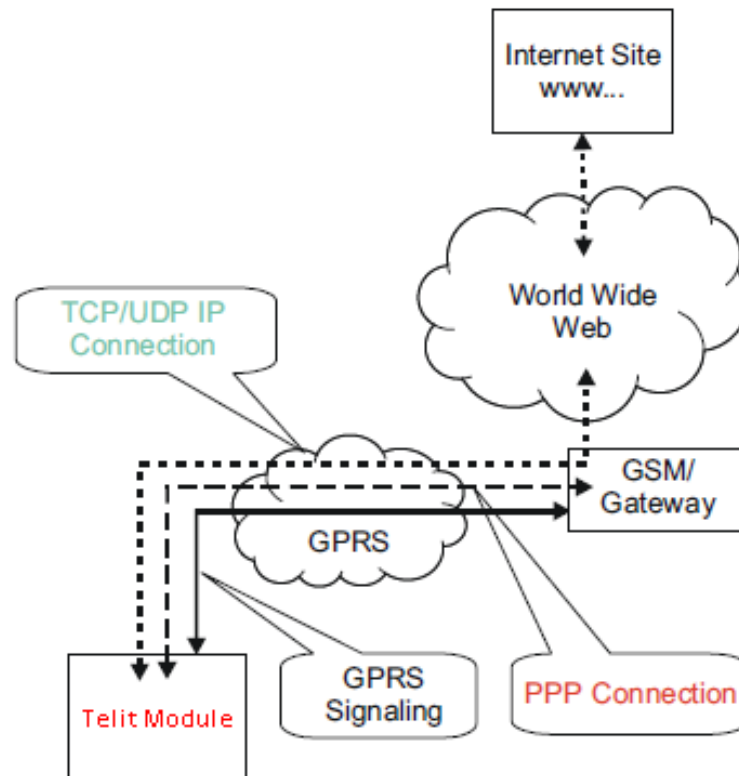


Figure 2. GPRS overview.

them is web based application labeled the data of cardiac patients by the doctor. The other is software application that the data collected on Server are learnt and evaluated. In this study, the second part has been more focused and used Weka Software Developer version 3.7.1 (WEKA, 2011).

Data acquisition

The data obtained from the patients are saved into the data base on Server by using General Packet Radio Service (GPRS) (Bettstetter et al., 1999; Brasche et al., 1997) technology. By this way, the absence of a cabled communication technology in the environment of the patient does not have importance. Additionally, the data are transferred into World Wide Web through the relevant GSM/Gateway and conveyed to the Server in TCP/IP packets. The background of this system is shown in Figure 2. It is seen that the performance of TCP is inflexible (Meyer, 1999; Othman et al., 2007). For this reason, the system used for collecting data is effective in terms of performance, flexibility and cost. Before the collected data is saved in the date base, the process of data filtration is implemented. This process includes determining the disorders over the data and cancelling these data because the data obtained from the patient via pulse oximeter (Nellcor N-200 pulse oximetry note number 6, 1998) have low quality signal. Following the data filtration procedure, the data are saved into the operational data base. Next the data determined by the expert are normalized, and trasported into date warehouse. The data in the warehouse are used as training data for learning based system. Systems, whether based on artificial intelligence in medicine (AIM) or other methods, must operate in conjunction with human practitioners (Patel et al., 2009). Therefore, in this system, learning and inductive expert

system design accompanied by the attendance of a specialist was also applied. This system is shown in Figure 3.

The number of data gathered for each patient differs. The reason for this is whether the training data occurring as a result of the expert's evaluations on the data in the data warehouse are enough or not.

Feature selection

One of the factors affecting the deciding processes of the machine learning algorithm is if the qualities are suitable or not. When selecting a good attribute subset, there are two fundamentally different approaches. One is to make an independent assessment based on general characteristics of the data; the other is to evaluate the subset using the machine learning algorithm that will ultimately be employed for learning (Witten et al., 2005). In our research, we used both approaches in the selection of quality. Before the learning stage is started, using the first approach, the qualities that can help composing a good training data so that the learning based system can take influential decisions were assessed. With the second approach, the results of the machine learning algorithms which were used were evaluated. Making use of the second approach, at the same time, the utility theory approach in the selection of the qualities was applied. Utility theory is concerned with making rational decisions when we are uncertain about the state (Alpaydın, 2004). If we express this as follows: let the quality we observed beforehand be x and let S position to be composed of detailed situations; S_k , $k = 1, \dots, n$. According to this, the probability of S_k that is known to belong x quality is counted as $P(S_k|x)$. Lets define the decision motion that designates the x quality to S_k position as a_i , and our utility function as U_{ik} . Thus, the

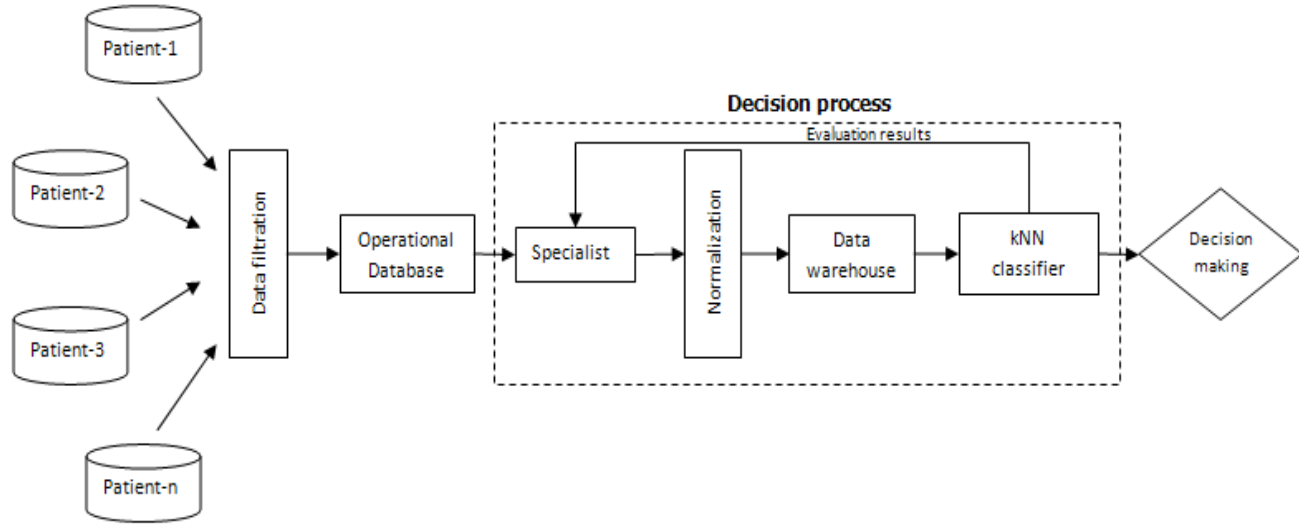


Figure 3. Learning-based system overview.

expected utility is:

$$EU(\alpha_i | x) = \sum_k U_{ik} P(S_k | x) \tag{1}$$

If we suppose the motion that maximizes the expected utility is α_i , the expected utility of α_i for x quality is shown as in Equation 2 equality:

$$EU(x) = \max_i \sum_k U_{ik} P(S_k | x) \tag{2}$$

Let our new quality added into the quality cluster be y . Hereunder, the expected utility is demonstrated as in Equation 3 equality:

$$EU(x, y) = \max_i \sum_k U_{ik} P(S_k | x, y) \tag{3}$$

According to this, if $EU(x, y) > EU(x)$, y quality can be said to be a remarkable quality. In other words, if the expected utility of the new quality added into the quality cluster is more than the expected utility of the former quality cluster, then the new quality can be said to be useful.

Features

The qualities chosen according to the criteria specified earlier are: pletismogram, saturation of peripheral oxygen (SPO_2), pulse strength ve signal quality. These qualities are described as follows:

Pletismogram: Together with pulse, the change of blood flow is trasmitted into arteries and can be measured as a pulse fluctuation. The volume change of this measured pulse fluctuation is named pletismogram (Sakane et al., 2003). In other words, the reason fort his volume change is the fluctuation in the blood amount in the artery. Using pletismogram test, the blood clots in the arms and legs can be checked or how much air is kept in the lung can be measure.

The pletismogram value range of the oximeter used in the patients is 0 LSB and 255 LSB (ENVITEC, 2008).

SPO_2 : The oxygen saturation in the blood or SPO_2 that is measured as the melted oxygen level is calculated with oximeter according to the following formula.

HbO_2 : Oksihemoglobin, **Hb:** hemoglobin:

$$SPO_2 = \frac{HbO_2}{HbO_2 + Hb} \tag{4}$$

The measured value range of the pulse oximeter used in the patients is 45 ile 100%. The accuracy range for these values is, however, determined as $70\% < SPO_2 < 100\%$ by the oximeter used (ENVITEC, 2008).

Pulsation strength: Pulse is the indication of how enough the heart works. Besides, quick or slow pulse can be helpful in the diagnosis of the cardiovascular diseases, because this case shows how the heart pumps the blood. The value range fort the used pulse oximeter is 0 and 250‰. The pulsation strength is low if the value is under 10‰ and is sufficient if this value is 15‰ (ENVITEC, 2008).

Signal quality: The signal quality value range is 0 to 100%, and if the signal quality is above 90%, the data on the signal quality are saved into the data base. Before the data stored in the data base is carried into its new warehouse, its quality can be said to be fine (ENVITEC, 2008). In the course measurements implemented with pulse oximeter, the data over 90% quality are recorded in the date base. Before the data recorded in the data base were transferred to the data warehouse, the min-max normalising method was applied to the signal quality feature. By applying min-max normalising, the pressure of the signal quality on the other qualities was decreased. By this way, the learning algorithm which was used could use its efficiency on the learning process except for signal quality in a better way. The data are transformed into the values between 0 and 1 through the min-max normalising method. This transformation procedure is demonstrated in Equation 5. For this equation; X' transformed values show X observation values, X_{min} shows the

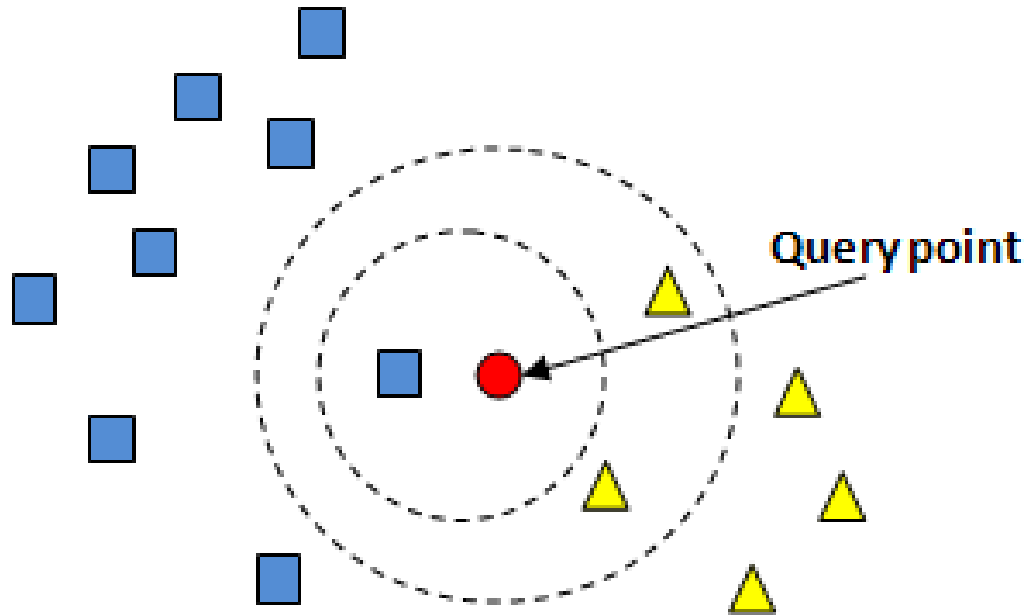


Figure 4. Squares and triangulars represent two different classes and the red circular is the query point.

minimum observation value and X_{\max} shows the maximum observation value:

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (5)$$

In the min-max normalising method applied for the signal quality feature, X_{\min} value was taken as 90 and X_{\max} as 100.

Machine learning and lazy learning

Inductive machine learning algorithms can learn patterns from labeled data, that is, cases that have a known outcome (Duda et al., 2001). There are very different machine learning approaches (Bishop, 2006) for pattern learning on data. One of these approaches is 'lazy learning'. Lazy learning describes case-based learning and reasoning approaches that scan a database of stored patterns and construct a localised model as required to respond to a query (Brownlee, 2007). One of the machine learning algorithms that uses 'lazy learning approach' is k-nearest neighbour (kNN). kNN is one of the simplest ones of the classifying machine learning algorithms.

Classification algorithm used in the experiment: k-nearest neighbor (kNN)

Let a data point collection in an m-dimension space and an inquiry be given, the process of finding the nearest data point is defined as the nearest neighbor problem (Beyer et al., 1999). K-nearest neighbor algorithm is, however, based on the principal of an inquiry point's distance to the nearest k piece of data point which is in a data point collection. In the calculation of the distances, such types of distance formulas as Euclidean, Manhattan, Chebyshev, Minkowski, etc. In Minkowski distance formula, the distances calculated for $m = 1$ and $m = 2$ give Manhattan and Euclidean

distances respectively. Taking the p and q as data points, $d(p, q)$ is the distance between two points. According to this:

$$d_{Euclidean}(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (6)$$

$$d_{Manhattan}(p, q) = \sum_{i=1}^n |p_i - q_i| \quad (7)$$

$$d_{Chebyshev}(p, q) = \max_i (|p_i - q_i|) \quad (8)$$

$$d_{Minkowski}(p, q) = \left(\sum_{i=1}^n |p_i - q_i|^m \right)^{1/m} \quad (9)$$

The distance formula we chose for kNN classifier is Manhattan distance formula. In the selection of this distance formula, that the training cluster was a quality space at a low dimension became influential. But, that the quality space is at high level does not render the use of kNN meaningless. In the high dimension quality space, the discreteness of the clusters from each other naturally provides the new query decrease for these classes (Beyer et al., 1999). For our classes which are low dimension and discrete from each other, while the distance is calculated, it is enough to take the absolute value of the difference between the points. This choice also affects the performance positively. The determination of the proper k value for k-nearest neighbor algorithm is very important; because according to k value, to which class the query point will belong may change. This is shown in Figure 4. When $k = 1$, the query point is designated for "square" class, but the increasing growth of k value will cause various problems.

One of these problems is that if there is not adequate training data for any of the classes, or if there is an imbalance in the training

Table 1. Landis and Koch Kappa values table.

Kappa	Interpretation
< 0	No agreement
0.0 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

data between the classes, in this case, test instance may be designated to a wrong class. If the distance of a query point in the training data to the data points is not so different from the average distance, then the nearest neighbourhood can be beneficial (Beyer et al., 1999). The reason for this situation is; the probability that the query point belong to a label or class apart from the classes or labels used for the classification of the training data. In other words, there may be classes which are not used during the classification of training data. Since we had a discrete and low dimension, we took k value as 1 for this application. kNN algorithm is called 1bk in Weka.

k-fold cross validation

A method for estimating the accuracy (or error) of an inducer by dividing the data into k mutually exclusive subsets (the “folds”) of approximately equal dimension. The inducer is trained and tested k times. Each time it is trained on the data set minus a fold and tested on that fold. The accuracy estimate is the average accuracy for the k folds (Kohavi et al., 1998). Using cross validation, restrictions in the data amount can be made (Witten et al., 2005). Accordingly, the elimination of the unnecessary data from training data can be provided. This prevents the learning algorithm from memorizing the data instead of learning it. On the other side, it will enable performance acquisition for memory-based learning algorithms like kNN. In the situations when the training data are not enough, since appropriate learning will not happen, the possibility of accurate prediction of the real class of a test data will decline. In the learning process achieved through using cross validation technique, 10-fold cross validation was used. Bias-variance (or, underfitting-overfitting) trade-off can be stabilized by using cross validation technique. The bias of a classifier is named as the difference between the prediction made by the classifier and the real class of the data. That this difference is quite high indicates that the classifier does not make accurate predictions. Namely, the classifier cannot realize adequate learning on the data. Nonlinear methods like kNN have low bias (Manning et al., 2008).

The variance of a classifier is stated as the change in the predictions made by the classifier. That is to say, the predictions made for a query point are different from each other.

Statistical evaluation criteria

Kappa statistic

Kappa statistics is an alternative to accuracy measure for evaluating classifiers (Kiliçaslan et al., 2009). It was first introduced as a metric that can be used in measuring the degree of agreement between two observers (Cohen, 1960) and has been used in several disciplines since (Kiliçaslan et al., 2009). In the area of machine learning, it is used as a measure to assess the improvement of a classifier’s accuracy over a predictor employing

chance as its guide (Kiliçaslan et al., 2009). Cohen’s Kappa statistic is shown in Equation 10:

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \quad (10)$$

Pr(a) shows the accuracy of the classifier and Pr(e) shows the accuracy obtained by a randomly predicting classifier on the same data set. Kappa statistics has a range between -1 and 1, where -1 is total disagreement (that is, total misclassification) and 1 is perfect agreement (that is, a 100% accurate classification). Kappa fundamentally assesses how much better a classifier is compared to the majority and class distribution based random classifiers scoring zero kapa (Kiliçaslan et al., 2009). A kappa score over 0.4 indicates a reasonable agreement beyond chance (Landis et al., 1977). Landis and Koch provided Table 1 to interpret the acquired Kappa values.

F-measure

F-measure value is expressed as the harmonical mean of precision and recall. For the patterns which are classified as positive and negative, the precision gives the rate of the positive ones classified correctly among the patterns predicted as positive. But, recall gives the rate of the positive patterns classified accurately among the patterns whose real class is positive:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$F - \text{Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

F-measure is an important criterion particularly in the preparation process of training data in order to increase the performance of the classifier. In this sense, Yilmaz Kiliçaslan’s study in which he revealed the relation between f-measure and data size in the performance analysis of the classifiers is significant. It was tried to attain a happy graphy (Russell et al., 2002) from the learning curve drawn by using the criteria presented in this study.

Root mean squared error

c is the unknown value of a parameter of a distribution, and X is an

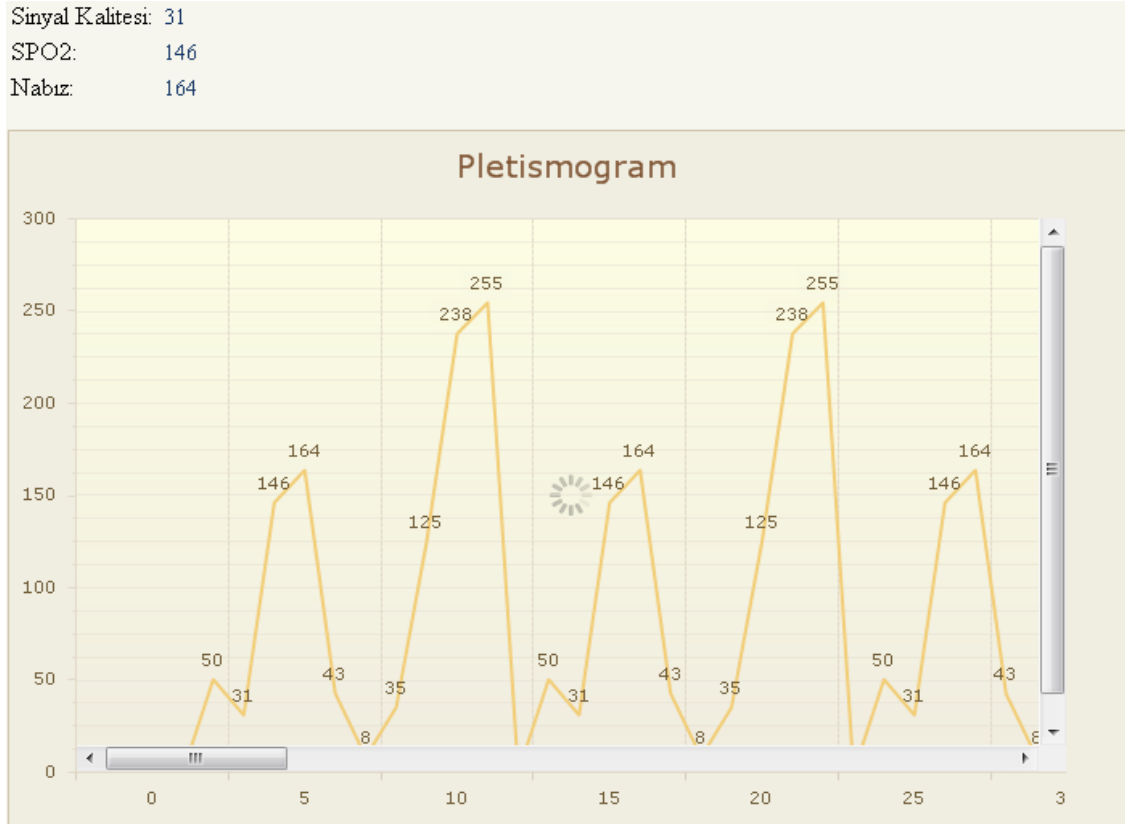


Figure 5. Overview of patient data stored database.

estimator of this parameter. In this context, it is usual to denote the estimator by θ^* (instead of X) and the value of the parameter by θ_0 (instead of c). ϵ is called the "estimation error" of θ^* . A good estimator is close to θ_0 on the average. Just how close is usually measured by the mean of the squared estimation error ϵ . This quantity is called the mean square error (MSE) of the estimator θ^* (AI ACCESS, 2011). Root mean-squared error is the square root of mean squared error:

$$MSE = E\left[(\theta^* - \theta_0)^2\right] \quad (14)$$

$$MSE = Var(\theta^*) + Bias(\theta^*)^2 \quad (15)$$

As seen in Equation 15, high rate of bias and variance raise up the value of MSE. For this purpose, in the preparation of training data, RMSE presented by WEKA is quite important. Using this criterion, the balance between bias-variance (underfitting-overfitting) can be provided. For this, the relationship between RMSE and data size is analyzed. For this relation, the change between data size and RMSE is examined, and for the lowest value of RMSE, it can be said that the value of data size at that point is appropriate for bias-variance balance.

RESULTS

In this study, two software applications, one of them is

web based and other PC based was developed. Software applications was developed using Visual Studio 2010 tool on .NET platform with C#. Web based application uses .NET 2.0, PC based application uses .NET 4.0 versions. Applications use MS SQL Server 2005 as database server. At the same time the applications serve together on the same server. The web based application has got an interface on which patient data stored in the database is analysed by an expert. This interface is shown in Figure 5. All arrhythmia patients is related own devices ID and, this relation is stored in the database. At the same time, every patient is related with an expert doctor. Software informs the statistical information about the stored data that is enough or not to learn, to the expert person. Interface related this subject is shown in Figure 6. If the expert person decide the data is enough considering this information, stored data for patient is moved to the data warehouse and, the training data is prepared. If the data is not enough, expert person continue labeling coming data and, this loop continues. The advantage of this application is to provide opportunity that the expert person can access the patient data at everywhere where there is internet. Figure 7 shows the role of expert person at the decision process.

Principally, patient data stored in the database is formatted as appropriately ARFF file format at WEKA.

Arama Kriterlerini Göster

Hastanın Adı Soyadı : fatih aydın

Sistem Değerlendirmesi	
Yeterli tanınm konulamadığı sınıflar	Kritik Çok Kritik
Sistemin Öğrenme Durumu	
Sistem mükemmel derecede öğrenme gerçekleştirmektedir. Kappa değeri : 1	

	<u>Cihaz No</u>	<u>Doktor Adı Soyadı</u>	<u>Kayıt Zamanı</u>	<u>Değerlendirme</u>
✖	1		02.11.2010 19:04:02	Normal
✖	1		02.11.2010 19:05:00	Normal
✖	1		02.01.2011 01:58:47	Normal
✖	2		02.01.2011 01:58:55	Normal
✖	2		02.01.2011 01:58:59	Normal
✖	1		02.01.2011 01:59:03	Normal
✖	1		02.01.2011 01:59:07	Normal
✖	1		02.01.2011 01:59:10	Normal
✖	1		02.01.2011 01:59:13	Normal

Figure 6. Evaluation of classifier according to all patient data.

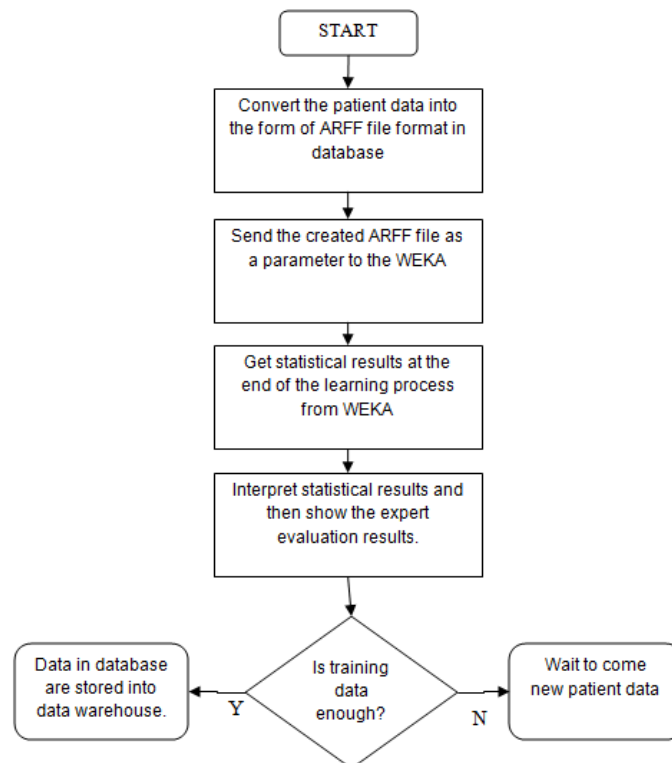


Figure 7. Decision process of the expert adequacy of data for training.

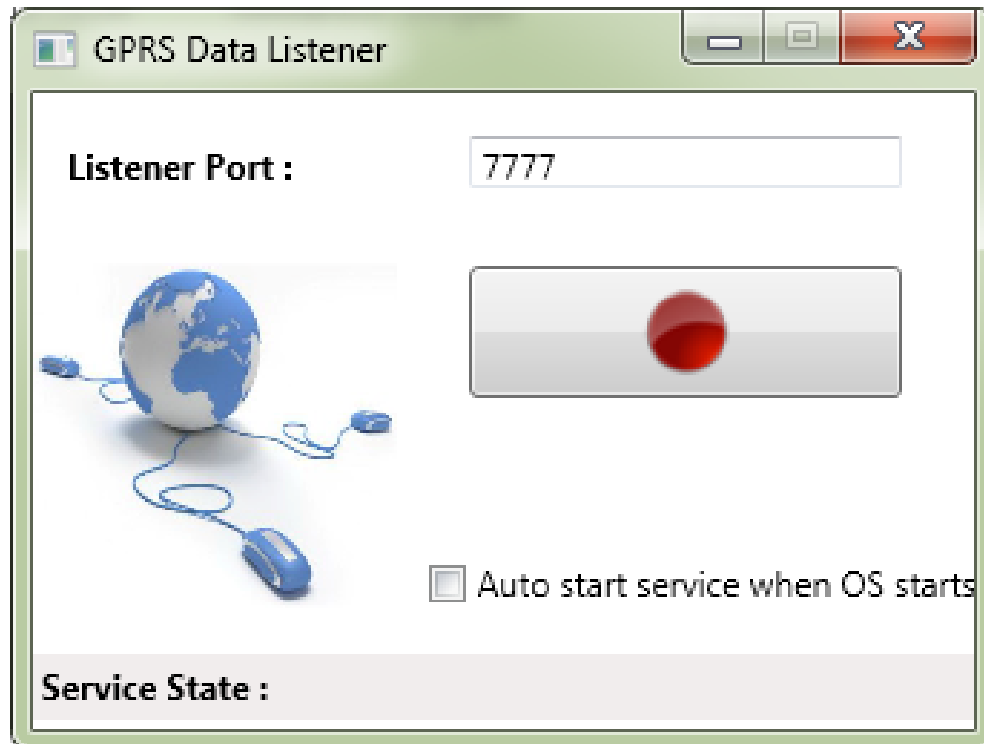


Figure 8. Developed application overview.

Then ARFF file is sent as training data to WEKA. In addition to this information, 'cross validation value', which classifier is used, etc. informations is sent to WEKA as parameter. After the learning process is completed, statistical results are taken from WEKA. Then these statistical results are evaluated by the system and the evaluation results are shown to expert person. The evaluation results contain information about which is enough to learn the data. If they are convinced that the data is not enough, then they wait for new patient data. If he is convinced that the data is enough, the data is carried to data warehouse from database. But the data in database are not deleted. This process repeats for each patient. The developed application for datas collection and classification stores arrhythmia patient data which comes over internet to the database. If the patient data is enough to learn according to expert person, the data are used as training data and this training data is classified by kNN classifier. If there is a critical condition for coming patient data which comes over internet, this condition is sent to patient's doctor by e-mail. If the patient data is not enough to be training data, coming data is continued to be collected. The interface of developed application is shown in Figure 8. Flow diagram at Figure 9 shows evaluation phase of patient data which comes over GPRS; firstly, it is determined that whichever data belongs to whichever patient. This determination is made by matching up the ID of device given patient and patients in the database. First coming data is ID of

device. After patient is identified and it is determined whether preparation step of the patient training data is completed. If the patient data is found enough as training data by the expert person, it is accepted that patient training data is ready. This situation informs us whether the training data is ready. If the training data is not ready, the data from patient is stored in the database and they remain as training data candidate. If the training data is ready, the data are used as test data. The test data, training data prepared as ARFF file from data warehouse and other data are sent WEKA as parameter. WEKA turns prediction result, about class of the test data.

According to this prediction result, if the test data class is not "normal", data about patient is sent to patient doctor determined in the system via e-mail and data are stored in the database. If the test data class is labeled as "normal", the data are stored only in the database. The capability and performance of the system developed are also directly proportionate to the performance of the classifier used in the learning procedure. For this reason, the performance analysis of the classifier and the methods to be used in the analysis are notable. For the basic classifiers like kNN, the size of the training data becomes disadvantageous. For this purpose, the most suitable data size should be selected. This problem reminds us the bias-variance balance. With the provision of this balance, the classifier fulfils an effective learning. For this purpose, f-measure, RMSE and kapa were used.

Primarily, evaluating f-measure and RMSFE criteria

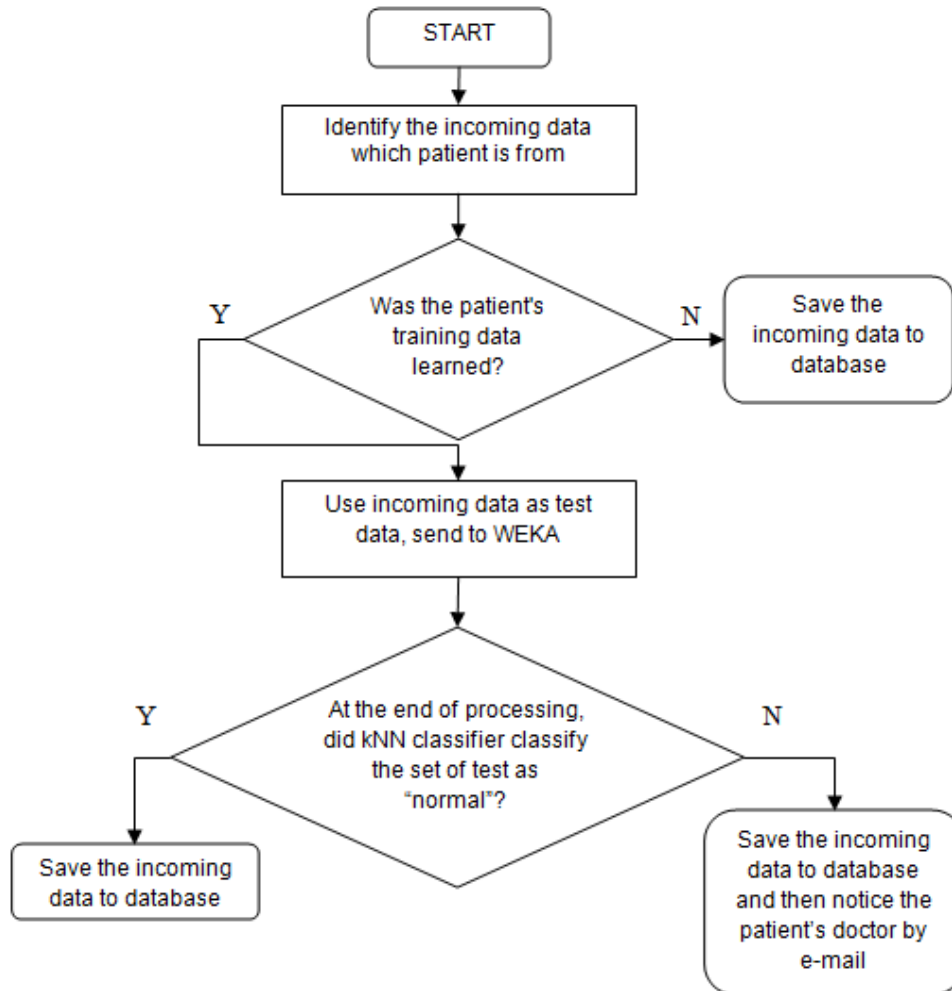


Figure 9. Decision process according to patient data.

together, the performance of the classifier was tried to be examined. Then, using kapa criterion, how well the classifier fulfilled learning was observed.

Conclusion

In this study, a system based on learning developed for monitorization of the treatment processes of the arrhythmia patients keeping them free from location was introduced. Firstly, the necessary equipment and GPRS technology were introduced to be able to gather patient data free from location and the reasons for their selection were explained. Then, the structural preparation process of training data which are necessary for a supervised learning method was described. Within this process, feature choice was made. In this feature choice, utility theory was taken into consideration. Henceforth, the features positively affecting the learning result were chosen. After the structure of the data was determined,

the data coming through internet were subjected to filtration procedure. Thusly, the data spoilt because of any reason were registered into the operational data base. The data recorded in the data base were later evaluated by an expert and it was decided if they could be used as training data or not. The data used as training data are passed through the normalization process and carried to date warehouse, and from now on, no process can be applied to these data. The resulting training data are trained by making use of kNN classifier which is a machine learning method. The training process of the data is implemented seperately for each patient and after the training procedure is completed, the new data coming from the patient is now used as query data. After the query data coming from the patient is labeled by kNN classifier, it is decided if the situation of the patient is critical or not. In consequence of this decision, if the situation of the patient is critical, the doctor of the patient is informed via sms or mail. Accordingly, the doctor can decide whether to make a change or not in the treatment

process of the patient. In the process of taking decision, the experience of a specialist and machine learning were used together in a system. Thus, it was gone out of a classical expert system design. While a human expert was fulfilling the choice of training data and labeling them, machine learning algorithm was used in order to develop a hypothesis from the training data. Following all of these processes, the learning based system got ready.

To assess the hypothesis that the kNN classifier composed on the training data, statistical methods were employed. These methods are kapa statistics, f-measure and root mean squared error (RMSE). With kapa statistics, the learning situation of the classifier was decided. Through F-measure and RMSE value, the performance of the classifier or the data amount that is enough for learning was determined. RMSE value was also used to provide the bias-variance balance. Except from this, in the determination of the k parameter for kNN classifier, the number of training data and how these classes are discrete from each other have importance. In the classification of the training data, three classes were used. Besides, in the analyses implemented, these classes were determined to be discrete from each other. The low number of class and the discrete situation of the classes became influential in the determination of k value as 1. In conclusion, in the determination and labelling training data, the influence of a human expert is very significant. In the process of learning training data and decision, machine learning techniques are seen to be the best approaches. In addition, statistical criteria in the assessment of the machine learning performance are of high importance. As a result of all of these, by enabling the integration of a human expert and machine learning methods, an inductive expert system design was made.

ACKNOWLEDGEMENTS

We thank Trakya University Faculty of Medicine Dean and Department of Cardiovascular Surgery faculty member Prof. Dr. Murat Dikmengil for assistance in the project.

REFERENCES

- Agrawal R, Imielinski T, Swami A (1993). Database Mining: A Performance Perspective. *IEEE Trans. Knowl. Data Eng.*, 5(6): 914-925.
- AI ACCESS (2011). Mean Squared Error, Available: http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_mean_sq_uare_error.htm [Accessed: 11 March 2011].
- Alihanka J, Vaahtoranta K, Saarikivi I (1981). A new method for longterm monitoring of the ballistocardiogram, heart rate, and respiration. *Am. J. Physiol.*, 240: R384-R392.
- Alpaydm E (2004). Introduction to Machine Learning. The MIT Press, Printed and bound in the United States of America. ISBN 0-262-01211-1.
- Benitez-Guerrero E, Collet C, Adiba M (2004). The Whes Approach to Data Warehouse Evolution. *e-Gnosis[online]*, 2(11).
- Bettstetter C, Vögel HJ, Eberspacher J (1999). GSM Phase 2+ General Packet Radio Service GPRS: Architecture. Protocols and Air Interface. *IEEE Commun. Surv.*, 2(3): 2-14.
- Beyer K, Goldstein J, Ramakrishnan R, Shaft U (1999). When Is 'Nearest Neighbor' Meaningful?. *Proceedings of the 7th International Conference on Database Theory*, pp. 217-235.
- Bishop CM (2006). *Pattern Recognition and Machine Learning*. Springer Science+Business Media, Printed in Singapore. ISBN: 0-387-31073-8.
- Brasche G, Walke B (1997). Concepts, Services and Protocols of the New GSM Phase 2+ General Packet Radio Service. *IEEE Commun. Magazine*, 35(8): 94-104.
- Brink M, Müller CH, Schierz C (2006). Contact-free measurement of heart rate, respiration rate, and body movements during sleep. *Behav. Res. Methods*, 38(3): 511-21.
- Brownlee J (2007). *Lazy and Competitive Learning*, Technical Report 070508A, Complex Intelligent Systems Laboratory, Centre for Information Technology Research, Faculty of Information and Communication Technologies, Swinburne University of Technology, Melbourne, Australia.
- Brüser C, Stadthanner K, Brauers A, Leonhardt S (2010). Applying machine learning to detect individual heart beats in ballistocardiograms. *Engineering in Medicine and Biology Society (EMBC), 2010 Ann. Int. Conf. IEEE*, pp. 1926-1929.
- Codd EF, Codd SB, Salley CT (1993). *Providing OLAP (On-Line Analytical Processing) to User-Analysis: An IT Mandate*. Technical report.
- Cohen J (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Measur.*, 20(1): 37-46.
- Duda RO, Hart PE, Stork DG (2001). *Pattern classification - 2nd edition*, New York: Wiley-Interscience.
- ENVITEC (2008). *Digital Pulse Oximeter Module ChipOx User Manual* [online], Available: <http://www.envitec.com/download/media/61.pdf/47-07-00330002-h.pdf;jsessionid=676D89A877E2D3FA24BFFBA31FFA1659> [Accessed: 15 February 2011].
- Galvin SD, Celi LAG, Galvin IF, Bunton RW (2011). Use of Machine Learning Algorithms to Assess Outcomes in Patients Over 80-years of Age Undergoing Cardiac Surgery. *Heart, Lung and Circulation*, 20(1): 48.
- Kılıçaslan Y, Güner ES, Yıldırım S (2009). Learning-based pronoun resolution for Turkish with a comparative evaluation. *Comput. Speech Language*, 23(3): 311-331.
- Kohavi R, Provost F (1998). *Glossary of Terms*. Springer Netherlands, 30(2-3): 271-274.
- Koivistoinen T, Junnila S, Varri A, Koobi T (2004). A new method for measuring the ballistocardiogram using EMFi sensors in a normal chair. *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Ann. Int. Conf. IEEE*, pp. 2026-2029.
- Krstacic G, Gamberger D, Martinis M, Krstacic A, Smalcelj A, Jembrek-Gostovic M (2001). Inductive machine learning results from a coronary heart disease database. *Atherosclerosis Supplements*, 2(2): 127.
- Krstacic G, Krstacic A, Gamberger D, Jembrek-Gostovic M (2006). M-P1:5 Risk factors analysis by inductive machine learning in coronary heart disease patients. *XIV International Symposium on Atherosclerosis, Atherosclerosis Supplements*, 7(3): 47.
- Kukar M, Kononenko I, Grošelj C, Kralj K, Fettich J (1999). Analysing and improving the diagnosis of ischaemic heart disease with machine learning. *Artif. Intell. Med.*, 16(1): 25-50.
- Landis JR, Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33: 159-174.
- Manning CD, Raghavan P, Schütze H (2008). *Introduction to Information Retrieval*. Cambridge University Press. ISBN: 0521865719.
- Meyer M (1999). TCP performance over GPRS. *WCNC 1999 - IEEE Wireless Commun. Networking Conf.*, 1: 1248-1252.
- Nellcor N-200 pulse oximetry note number 6 (1988). *C-LOCK ECG synchronization principles of operation*. Calif, Nellcor: Hayward.
- Othman AK, Zakaria M, Ab Hamid K (2007). TCP performance measurement in different GPRS network scenarios. *Applied Electromagnetics, APACE 2007, Asia-Pacific Conference*, pp.1-5.
- Patel VL, Shortliffe EH, Stefanelli M, Solovits P, Berthold MR, Bellazzi

- R, Abu-Hanna A (2009). The coming of age of artificial intelligence in medicine. *Artif. Intell. Med.*, 46(1): 5-17.
- Russell SJ, Norvig P (2002). *Artificial Intelligence: A Modern Approach* - 2nd ed, Prentice Hall, Printed in the United States of America. ISBN 0-13-790395-2.
- Sakane A, Tsuji T, Tanaka Y, Saeki N, Kawamoto M (2003). Estimating arterial wall impedance using a plethysmogram. *IECON03 29th Annual Conference of the IEEE Industrial Electronics Society IEEE*, 1: 580-585.
- Shirouzu S, Sugano H, Hara K (2001). A new method for unconstrained heart rate monitoring. *Engineering in Medicine and Biology Society, Proceedings of the 23rd Ann. Int. Conf. IEEE*, 4: 3960-3961.
- Tantimongcolwat T, Naenna T, Isarankura-Na-Ayudhya C, Embrechts MJ, Prachayasittikul V (2008). Identification of ischemic heart disease via machine learning analysis on magnetocardiograms. *Comput. Biol. Med.*, 38(7): 817-825.
- TELIT (2011). Telit GM862 GSM Datasheet, Available: <http://www.telit.com/module/infopool/download.php?id=165> [Accessed: 13 March 2011].
- WEKA (2011). Available: <http://www.cs.waikato.ac.nz/ml/weka/> [Accessed: 08 March 2011].
- Witten IH, Frank E (2005). *Data mining: practical machine learning tools and techniques*– 2nd ed. p. cm. Morgan Kaufmann series in data management systems. ISBN: 0-12-088407-0.
- World Health Organization (2004). *The world health report 2004*. [Online]. Available: http://www.who.int/whr/2004/annex/topic/en/annex_2_en.pdf [Accessed: 02 February 2011].