

Intelligent combination of data capable of online weight update for the diagnosis of Heart disease

Roya Zare Hanji ^{1*}, Ali Mahdizade ², Amin Zare Hanji ³

¹ Graduated with a master's degree in electricity, Alborz University, Qazvin city, Iran. ² Ph.D. Student in Civil Engineering, Chalous Branch Azad University, Chalous City, Iran. ³ Graduated from IT, Faran Mehr Danesh University, Tehran, Iran.

Correspondence: Roya Zare Hanji, Graduated with a master's degree in electricity, Alborz University, Qazvin city, Iran. Email: roya.zare.alborz@gmail.com

ABSTRACT

A large number of people perish or their functions are affected by heart disease in the world every year. The death toll from heart problems is much higher than other accidents or natural disasters. This is the main reason for the development of scientific and research activities in medical sciences and in the context of heart disease. It is widely and rapidly expanded into other sciences, such as engineering, to find effective preventive methods for these diseases. In this study, such an objective has been developed with the aim of helping the development of cardiac arrhythmia detection algorithms using data from measurement systems of cardiac characteristics, such as electrocardiogram signals. The features were extracted from the pulses according to the shape of the captured signals and the changes in the process. Then, the frequency and temporal properties of the signal were studied simultaneously by examining different methods, among which the UWMV-GSA method was selected for this purpose. After feature extraction, the feature length was reduced in the next step by applying the PCA operator to reduce the feature space dimension. In the final step, the received signals were classified based on the extracted feature vectors through combining information in the classification of cardiac arrhythmias using five base classifiers of LMS, RLS, decision tree, KNN, and SVM to classify cardiac datasets. Furthermore, the weight was selected based on the gravitational search algorithm (GSA) optimization. Different hybrid classification methods are compared in this study and a new UWMV-GSA algorithm is proposed for dynamic dataset classification in which each specialist (base classifier) predicts a specific area of the dataset with the highest accuracy. Thus, system performance is increased by up to 90% by combining the results. The accuracy and weight of each classifier and the UWMV-GSA for test and training data are presented in various tables and graphs.

Keywords: Cardiac arrhythmias, ECG signal, Neural networks, Knowledge combination, SVM algorithm, Decision tree algorithm, KNN algorithm, RLS algorithm, ELS algorithm, GSA algorithm.

Introduction

The heart is a hollow, almost cone-shaped muscular organ in the chest, which is one of the main components of the circulatory system. The heart is working constantly and beats about a million times a day to circulate blood through the arteries for 113,000 km in the human body. Similar to a pump, the heart sends arterial blood to the farthest parts of the body and then returns to itself the used (venous) blood to the lungs and kidneys for purification. Each cardiac activity round lasts about 0.8 seconds and consists of systolic contraction and rest (diastole)

stages. With each heartbeat, oxygenated blood is pumped to all parts of the body by the arteries and returns into the heart by the veins.

Upon relaxation of the other bodily muscles, the electrical currents of the heart can be received and recorded by an electrocardiograph. As such, the waves resulting from the electrical activity of the heart are received by inserting electrodes in the chest surface around the heart.

Electrocardiogram is used to diagnose many cardiac and non-cardiac disorders, such as abnormal heart rhythms, coronary artery spasms, heart attacks, hypertrophy of cardiac muscle, causes of dyspnea, electrolyte disturbances, effects of drugs, etc.^[1]

The electrocardiogram records electrical events of the heart at a standard rate of 25 mm/s. A few years ago, electrocardiogram signals contained very important information about the heart condition and were one of the tools commonly used by physicians for the diagnosis of various heart diseases. Some disadvantages of this method include low spatial accuracy,

Access this article online

Website: www.japer.in

E-ISSN: 2249-3379

How to cite this article: Roya Zare Hanji, Ali Mahdizade, Amin Zare Hanji. Intelligent combination of data capable of online weight update for the diagnosis of Heart disease. *J Adv Pharm Edu Res* 2020;10(S4):144-152. Source of Support: Nil, Conflict of Interest: None declared.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

limited time accuracy, similarity of signals in some diseases to those of a healthy heart, and possible unseen information by the physician. The importance of the electrical signals of the heart and the above-mentioned disadvantages necessitates the focus on an optimal alternative method for the expression and analysis of cardiac activity ^[2].

Smartification of the diagnosis process of heart disease has long been discussed by researchers in all countries. This process involves the steps during which the electrocardiographic (ECG) signal is selected as the software input, and the software is expected to accurately detect health or disease and even the type of heart disease with acceptable accuracy. All these software receive the signal, extract and select appropriate features, and then diagnose the type of disease. There are different methods in each of the mentioned steps ^[3].

The use of this technique as a complementary tool can increase the speed and accuracy of heart disease diagnosis by a physician or a user. In this method, signal information of a normal person by default (and the patient) is provided to the user as a two-dimensional image and its ease of use and detection compared to conventional processing-free images can increase the accuracy and speed of the disease diagnosis process. Signal processing automatically by software significantly reduces the error level due to mechanical recording and visual processing of signals, enabling physicians to more easily and accurately analyze one's signal. Some changes in the electrical activity of the heart may not be observed in one's signal, which can be detected in the analysis by software. It is also possible to determine the location and severity of cardiac arrest with good accuracy. In fact, the presence of such a device reduces the need for magnetic resonance imaging and prevents its harmful effects. It is also possible to diagnose a cardiac arrest by the use of such a method, even in the early stages of the disease ^[4].

In many cities, however, access to a specialist is not readily feasible, and there are children with unknown congenital heart disease until adulthood, when the disease symptoms appear gradually. Due to the importance of this issue, an intelligent system has been designed for heart disease detection in children in the country ^[5]. Data mining is one of the most important programs that can be used in healthcare management. Increasing the parameters necessary for the disease detection makes it difficult to diagnose and predict the disease even for a qualified medical professional. For this reason, computer diagnostic tools have been used to help physicians in recent decades. This important issue has partially reduced potential errors caused by fatigue or inexperience of a specialist, and the physician is provided with required medical data in less time with more detail and accuracy ^[6].

Additionally, cardiac arrhythmias are among the most common causes of death worldwide. Hence, distinction and diagnosis of cardiac arrhythmias using smart devices is one of the most important methods for rapid and accurate detection of symptoms in cardiac signals.

Arrhythmia is usually diagnosed by visual analysis of a human electrocardiograms. In a study, a computational approach was proposed to diagnose arrhythmia using the analysis of variable heart rate signals by a SONFIS (Structure Identification and Modeling with a Self-Organizing Neuro-Fuzzy Inference System) model. This research method produces a set of linguistic inference rules for pattern classification and uses artificial neural networks and support vector machines for accuracy as well as several other performance indicators ^[7].

Probabilistic neural networks (PNNs) can be used to more accurately diagnose coronary heart disease. Due to its specificity and high sensitivity, this method can prevent probable complications and damages of angiography in patients who do not need thereto. It can also identify patients who really need these diagnostic procedures in the shortest time with the highest accuracy ^[8].

A plethora of biomedical research is devoted to electrocardiogram techniques to assist in early diagnosis. However, reported analyses for ECG methods are commonly limited to PC host operations. The authors present an arrhythmia classification method performed on a digital signal processing platform intended for real-time operations to classify eight heartbeats: normal sinus rhythm (N), ear fibrillation (AF), premature atrial contraction (PAC), left bundle branch block (LBBB), left bundle branch block (RBBB), premature ventricular contraction (PVC), sinoauricular heart block (SHB), and supraventricular tachycardia (SVT).

Classification is done using a PNN. The algorithm was tested with 17 ECG files obtained from physiotherapy. Results showed that the proposed method and prototype can be suitable for use in wearable-assisting measurement programs for the online, real-time detection ^[8].

A cardiac arrhythmia classification algorithm has been introduced using the heart rate variation (HRV) signal ^[9]. The proposed algorithm is based on general document analysis (GDA) and the support vector machine (SVM) classification system. The presented cardiac arrhythmia classification algorithm is effective. The main advantage of the proposed algorithm is comparable with the approaches that use their ECG signal. In addition, using the HRV signal leads to an effective reduction in the processing time that provides the arrhythmic classification system.

A cardiac arrhythmia classification was assessed using neural network troubleshooting training in ^[10]. The aim of this study was to use learner vector quantization (LVQ) neural networks to classify arrhythmias from ECG datasets. LVQ classification algorithms do not draw density functions of class samples but directly define class boundaries based on prototypes, a nearest-neighbor rule, and a winner-takes-it-all paradigm. This superior performance over back-propagation (BP) method means minimization of the classification errors while maintaining rapid convergence. The classification performance of each algorithm is evaluated using four criteria, viz. sensitivity, specificity, classification accuracy, and the time required to build the

system. Based on experimental results, the use of the LVQ algorithm is recommended for more extensive research on this subject [10].

Heart disorders are one of the leading causes of death and, therefore, they need continuous and efficient detection techniques. ECG is one of the main tools for detecting cardiovascular disorders such as arrhythmias. A new technique for cardiac arrhythmia classification was performed using spectral correlation vector machines [11]. The proposed system is based on the stress signal analysis, which examines hidden signals in hidden periods and thus can identify hidden properties. Experimental results showed that the approach using raw spectral correlation data was superior to various artistic methods. This approach achieved sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of 99.20%, 99.70%, 98.60%, 99.90%, and 97.60%, respectively [11].

A study focused on cardiac arrhythmia classification with support SVMs and genetic algorithm (GA), presenting a new approach to heart arrhythmia classification. The proposed method combines both support SVMs and the GA approach. To this end, the support SVM classification design is optimized by searching for the best parameters that regulate their detection function

performance, and seeks the best subset of features that optimize the fitness classification function. Empirical findings indicate that the applied approach better classifies ECG signals. The four types of arrhythmias were different with 93% accuracy [12].

According to the above, a new method for heart disease detection capable of online weight update was presented in this study using intelligent methods of data combination.

Materials and Methods

Simulation

This research is based on the simulation of the ensemble learning method. Due to a high error rate in other studies, it was tried here to reduce the error rate with a new method to diagnose heart disease by summarization and correction of test methods. The proposed methodology consists of several sections. The cardiac database section covers attributes used to recognize diseased and healthy individuals. The database consists of 14 columns and 267 rows, of which 13 columns and one column represent attributes and the class label, respectively.

Our input data, i.e. the characteristics extracted from patients' cardiac signals, are as follows:

| | | | | | | | | | | | | |
|----------|---------------|---------|------------|--------------|--------|--------------|---------------|----|----------------|--------|--------|--------|
| Complete | | | | attribute | | | | | documentation: | | | |
| 1 | Age: | Age | in | years | , | | | | linear | | | |
| 2 | Sex: | Sex | (0 = male; | 1 = female) | , | | | | nominal | | | |
| 3 | Height: | Height | in | centimeters | , | | | | linear | | | |
| 4 | Weight: | Weight | in | kilograms | , | | | | linear | | | |
| 5 | QRS duration: | Average | of | QRS duration | in | msec., | | | linear | | | |
| 6 | P-R interval: | Average | duration | between | onset | of | P and Q waves | in | msec., | linear | | |
| 7 | Q-T interval: | Average | duration | between | onset | of | Q and offset | of | T waves | in | msec., | linear |
| 8 | T interval: | Average | duration | of | T wave | in | msec., | | linear | | | |
| 9 | P interval: | Average | duration | of | P wave | in | msec., | | linear | | | |
| Vector | angles | in | degrees | on | front | plane | of; | | linear | | | |
| 10 | | | | | | | | | QRS | | | |
| 11 | | | | | | | | | T | | | |
| 12 | | | | | | | | | P | | | |
| 13 | | | | | | | | | QRST | | | |
| 14 | | | | | | | | | J | | | |
| 15 | Heart rate: | Number | of | heart | beats | per | minute | , | linear | | | |
| Of | | | | channel | | | | | DI: | | | |
| Average | width, | | in | msec., | | of: | | | linear | | | |
| 16 | | | | Q | | | | | wave | | | |
| 17 | | | | R | | | | | wave | | | |
| 18 | | | | S | | | | | wave | | | |
| 19 | R' | wave, | small | peak | just | after | | | R | | | |
| 20 | | | | S' | | | | | wave | | | |
| 21 | Number | | of | intrinsic | | deflections, | | | linear | | | |
| 22 | Existence | | of | ragged | R | wave, | | | nominal | | | |

| | | | | | | | | |
|-----------|-----------|----|--------------|------------|-------|--------|-----------|----------|
| 23 | Existence | of | diphasic | derivation | of | R | wave, | nominal |
| 24 | Existence | | of | ragged | | P | wave, | nominal |
| 25 | Existence | of | diphasic | derivation | of | P | wave, | nominal |
| 26 | Existence | | of | ragged | | T | wave, | nominal |
| 27 | Existence | of | diphasic | derivation | of | T | wave, | nominal |
| Of | | | | channel | | | | DII: |
| 28 | .. | 39 | (similar | to | 16. | 27 | of | channel |
| Of | | | | channels | | | | DIII: |
| 40 | | | | .. | | | | 51 |
| Of | | | | channel | | | | AVR: |
| 52 | | | | .. | | | | 63 |
| Of | | | | channel | | | | AVL: |
| 64 | | | | .. | | | | 75 |
| Of | | | | channel | | | | AVF: |
| 76 | | | | .. | | | | 87 |
| Of | | | | channel | | | | V1: |
| 88 | | | | .. | | | | 99 |
| Of | | | | channel | | | | V2: |
| 100 | | | | .. | | | | 111 |
| Of | | | | channel | | | | V3: |
| 112 | | | | .. | | | | 123 |
| Of | | | | channel | | | | V4: |
| 124 | | | | .. | | | | 135 |
| Of | | | | channel | | | | V5: |
| 136 | | | | .. | | | | 147 |
| Of | | | | channel | | | | V6: |
| 148 | | | | .. | | | | 159 |
| Of | | | | channel | | | | DI: |
| Amplitude | , | | | * | 0.1 | | milivolt, | of |
| 160 | | | JJ | | | | wave, | linear |
| 161 | | | Q | | | | wave, | linear |
| 162 | | | R | | | | wave, | linear |
| 163 | | | S | | | | wave, | linear |
| 164 | | | R' | | | | wave, | linear |
| 165 | | | S' | | | | wave, | linear |
| 166 | | | P | | | | wave, | linear |
| 167 | | | T | | | | wave, | linear |
| 168 | QRSA | , | Sum | of | areas | of | all | segments |
| | | | divided | by | 10, | (| Area= | width |
| | | | | | | * | height | / |
| | | | | | | 2 |), | linear |
| 169 | QRSTA | = | QRSA | + | 0.5 | * | width | of |
| | | | T | wave | * | 0.1 | * | height |
| | | | of | T | wave. | (If | T | is |
| | | | diphasic | then | the | bigger | segment | is |
| | | | considered), | | | | | linear |
| Of | | | | | | | | DII: |
| 170 | | | | | | | | 179 |
| Of | | | | | | | | DIII: |
| 180 | | | | | | | | 189 |
| Of | | | | | | | | AVR: |
| 190 | | | | | | | | 199 |
| Of | | | | | | | | AVL: |
| 200 | | | | | | | | 209 |
| Of | | | | | | | | AVF: |
| 210 | | | | | | | | 219 |
| Of | | | | | | | | V1: |
| 220 | | | | | | | | 229 |

| | | |
|------------|---------|-----|
| Of | channel | V2: |
| 230 | .. | 239 |
| Of | channel | V3: |
| 240 | .. | 249 |
| Of | channel | V4: |
| 250 | .. | 259 |
| Of | channel | V5: |
| 260 | .. | 269 |
| Of | channel | V6: |
| 270 .. 279 | | |

The general written program was compiled in the min pr function that has several functions. Four base learners were used initially, but the results were not very satisfactory. Thus, it was tried to make a stronger learner as each of the base learnings alone does not yield interesting results. Finally, the majority voting method was used for this purpose. In the weighted voting method, a model with the highest accuracy will have a higher weight. In this method, there are some base learners who have been weighted by the GSA method. Here, two steps were used to determine the weighting of base learners.

After identifying the labels using the base learners in the first step, the weight of the base learners is determined using the GSA method.

Lbest contains the best weighted results.

Best chart is a convergence chart containing the best fit answers.

Mean chart is the average fitness of members in the GSA calculated per repetition.

The next step is to update the model weight while running the model on the test data, for which the test data are applied to the created base models.

At the same time in this topic, the label is determined for a new sample (test I) to specify the weight of these models. In fact, the weights are updated using the GSA. The initial value of new weight is the output of the same model at the train moment. Lbest is the train output set to a new weight, which is updated in each repetition in the test data loop. This update is based on the GSA, in which the new sample tag is incorporated into the predicted label and weight decisions are made based on the new sample. This process is done for individual test data and three graphs are printed in the end. The results can be viewed and analyzed using these graphs.

In this research, base models were used in a loop called model number one. Four base models were used initially, which did not yield interesting results. Thus, a combination of decision tree method, KNN, and SVM with the other two mentioned base models was used because these methods have parameter estimation. At any time, the weights were updated by the GSA as presented below:

The initialization of the GSA parameters is followed recalling each factor equal to the number (5) of base learners, and N is equal to the number of particles.

Initialization and population were considered in the program, and the space boundary was used to examine the upper and lower limits, each considered 1 and 0.1, respectively.

In fact, the fitness = evaluate function was considered here, which itself predicts the model label extract using its own pattern, and extracts the labels using a weighted voter. Since the goal is to minimize error, this method is then used to calculate the best fit and the best member of the population.

Afterward, the GSA process can be calculated to update objects, the constant coefficient, and particle motion in the gravitational system. This process is repeated in its max number (100) until the end of GSA.

After the weight calculation by the GSA, it is important to know about our data labels. Using the weight extracted by the GSA, the best data labels can be determined using Lbest. The observed updated weight can be compared to the previous weight to examine the amount of changes.

Lbest = the amount of weight in the proposed method (Fig. 2) applied to the most events.

New weight = Applying the model to test data with updated weights.

DW = weight deviations.

The weights of the first model increased by 0.04, the second model decreased by 0.3, the third model increased by 0.04, and the fourth model had no considerable changes. Finally, the final test label determines the labels predicted for the test data.

In the program, the label of each base classifier learner can be seen for the test data, which is a 3×17 matrix.

Test label is each of the base classifier learner in our model on the test data.

Accuracy is the output of models on test data.

Test label and test data are the output of our model on test data.

Train accuracy of our model is the best data with an accuracy of 1, indicating that the model had a pre-fit as well.

Due to a high volume of data in this program, the data size reduction method was used to improve the model function and increase the speed of the model, which reduced the data size by deleting unnecessary data, thereby accelerating the model running.

Line 8 of the program separates and thereby selects the features. Using hold our, data were divided into two categories of test and best, each comprising 70% and 30%, respectively.

In the program output, there is a set of i , which is the output of model number one, which updates the weights on the test data by each application of the test data. The evaluateF, Gconstant, Gfield, massCalculation, move, initialization, and space_bound belong to the GSA function. Entropy, JointEntropy, and MutualInformation are used to calculate the information feature. Model Label Extract is used to determine data labels in GSA and min pr based on the incorporated inputs. W denotes the dominant label identity. In fact, W is the GSA output and L_{best} is the output of the base modes.

Evaluation criteria

The evaluation criteria were explained in the previous chapters. This section represents the criteria used to measure the performance of the algorithms. In this research, the following two criteria were used to evaluate the algorithms. The first criterion was accuracy used to measure the efficiency of algorithms. This criterion is calculated based on the perturbation matrix. The recall and accuracy are calculated as follows:

$$Accuracy = \frac{tp+tn}{tp+tn+fp+fn} \tag{1}$$

In Formula 1, TP is the number of samples with a positive class detected positive, FP is the number of samples with a negative class detected positive, FN is the number of samples with a positive class detected negative, and TN is the number of samples with a negative class detected negative.

The second criterion for calculating the performance of algorithms was the Q-statistics, which is calculated as follows:

$$Q_{i,j} = \frac{(ad-bc)}{(ad+bc)} \tag{2}$$

In criterion 2, a is the number of samples predicted by both classifications i and j correctly, b is the number of samples predicted by class i correctly, but class j was not able to predict, c is the number of samples predicted by class j correctly, but class i was not able to predict, and d is the number of samples predicted incorrectly by both classifications.

This criterion shows the behavior of classifiers regarding the data. If this value is equal to one, it means that both classifiers had similar predictions. If this value is -1 , it means that the two classifiers (predictions) behaved quite contrarily. And if this value is zero, it means that the classifiers had maximum diversity.

Model implementation software

In this research, the proposed model was implemented by the MATLAB software.

Results

Results of base classifiers

Based on the results (Tables 1), the accuracy of the base classifiers was relatively good, but not acceptable. Based on values of Q-statistics, on the other hand, the classifiers had a relatively similar performance. This indicates that classifiers with less behavioral similarity to these classifiers should be used to increase the final accuracy of the model. For this purpose, other base classifiers were used here.

Table 1- Results of base classifiers in training and test data

| Classifier | Accuracy | Mean Q-statistics |
|-------------|----------|-------------------|
| | Training | |
| RLS | 66.03 | 76.0 |
| ELS | 61.87 | |
| RLS-voltera | 65.45 | 0.87 |
| ELS-voltera | 62.87 | |
| Test | | |
| RLS | 63.03 | 88.-0 |
| ELS | 60.56 | |
| RLS-voltera | 62.84 | -0.65 |
| ELS-voltera | 61.43 | |

Results of combining classifiers

The SVM base classifier was used due to the low accuracy of the algorithm in classifying the obtained composition, and the results are shown in Table 2.

Table 2- Results of combining classifiers on training and test data

| | | Accuracy |
|-----------------------|--------------------------|----------|
| Training | | |
| Base-classifiers | RLS-Volterra | 62.84 |
| | ELS-Volterra | 61.43 |
| | RLS | 63.03 |
| | ELS | 60.56 |
| | CART | 89.90 |
| | KNN | 100 |
| | SVM | 100 |
| Combining classifiers | Majority voting | 83.13 |
| | Weighted majority voting | 96.6 |
| | UWMV-GSA | 100 |
| Test | | |
| Base-classifiers | RLS-Volterra | 62.84 |
| | ELS-Volterra | 61.43 |
| | RLS | 63.03 |
| | ELS | 60.56 |
| | CART | 68.96 |
| | KNN | 73.02 |
| Combining classifiers | SVM | 66.66 |
| | Majority voting | 65.66 |
| | Weighted majority voting | 70.037 |
| | UWMV-GSA | 83.06 |

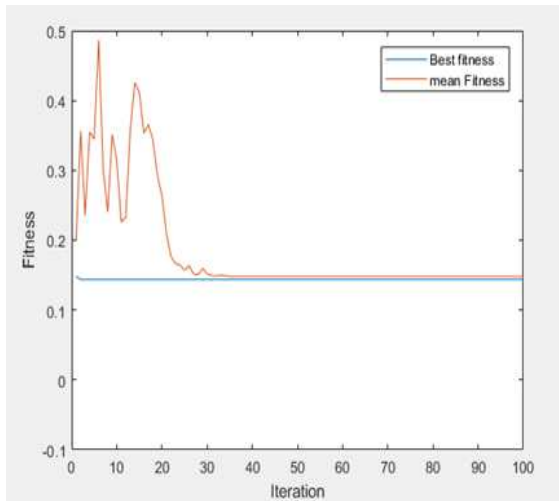
The results in Table 2 indicate that a combination of classifiers can be a good idea for data classification. Given that the best results of a classification cannot be predicted based on the data, the advantages of classes should be used by the use of some methods, for which the best option is to combine the classifiers. According to the results, the combination of classifiers shifts the results towards the dominant classifier. It can, therefore, be

concluded that the combination of classifiers is a reliable approach to classify data with unknown structures.

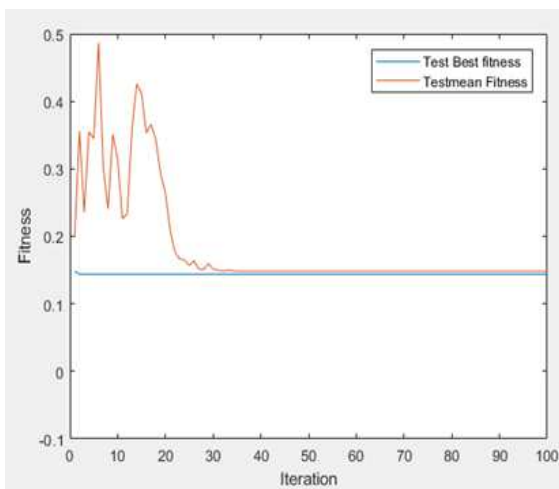
Considering that the weighting the classifiers was based on training data in this study, the higher the accuracy of the classification in the training data, the greater the weight of that classifier in the final decision.

Results of GSA

In this research, the GSA was used in two modes of without and with (the proposed method) the ability of online weight update. In this section, the results of the convergence diagrams and the GSA weight are presented for both modes.



(a)



(b)

Figure 1- GSA convergence diagram without (a) and with (b) online weight update ability

Table 3 shows the weight of each classifier in both modes of using the GSA.

Table 3- Weights of classifiers in the combination of classifiers

| Classifier | KNN | CART | RLS | ELS | RLS-Volterra | ELS-Volterra |
|------------|------|-------|------|------|--------------|--------------|
| GSA1 | 0.87 | 0.4 | 0.32 | 0.15 | 0.27 | .0.17 |
| GSA2 | 0.84 | 0.54 | 0.24 | 0.14 | 0.23 | 0.19 |
| Difference | 0.3 | -0.14 | 0.08 | 0.01 | 0.04 | -0.02 |

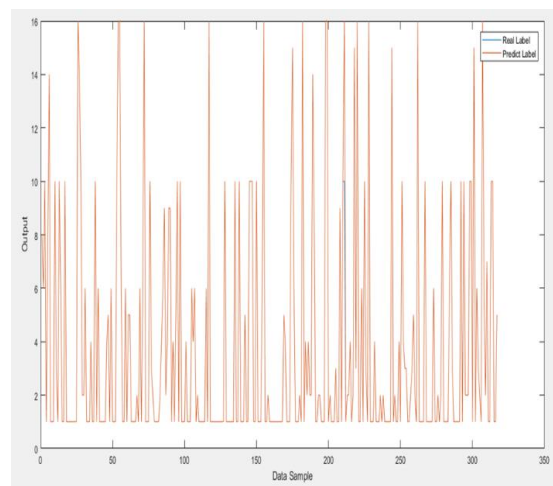
The weights of the classifiers in the GSA2 method is the same as the proposed method.

Figures A and B show that the GSA examines various answers at the beginning of running and the model output is random independent of the results of the classifiers, but then the convergent algorithm returns to its optimal answer.

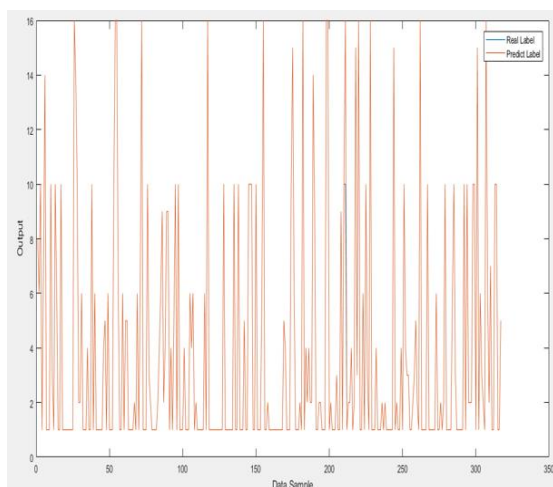
Table 3 also shows that the weights of the classifiers become more realistic by updating the weights online, and this finding is supported by the results of the models on the test and training data. As the results become more realistic in online weight update, the base classifier models become closer to their actual weights. Therefore, it can be concluded that online weight update increases the accuracy of hybrid classifier models because base models show the actual behavior on the data over time.

Results of indicator outputs

In this section, in the output diagrams of the proposed method are represented on the test and training data. For this purpose, the output of the MATLAB software diagram is used directly. Figure 2 (a and b) shows the output of the proposed method on training data.



(a)

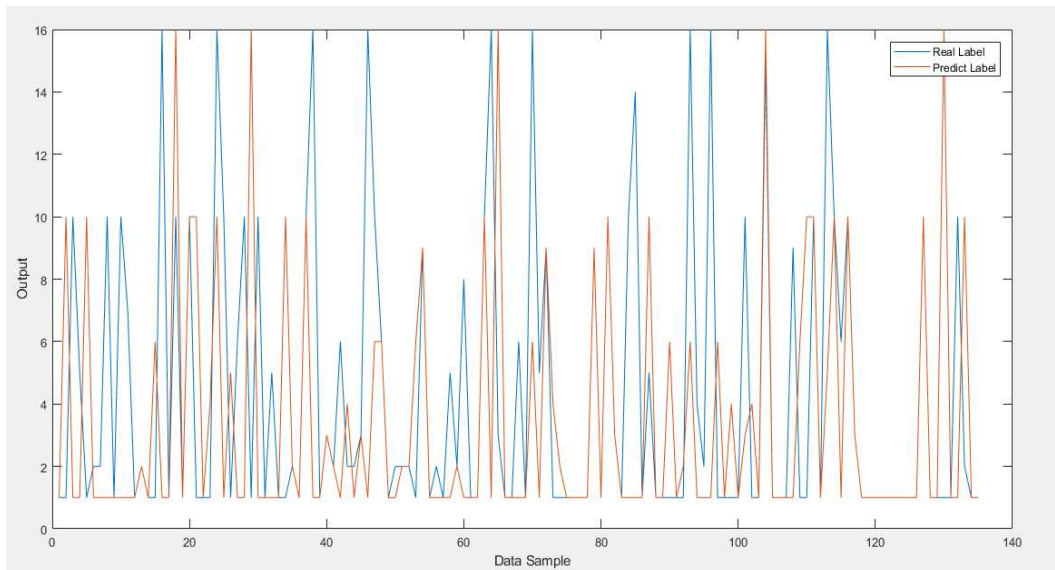


(b)

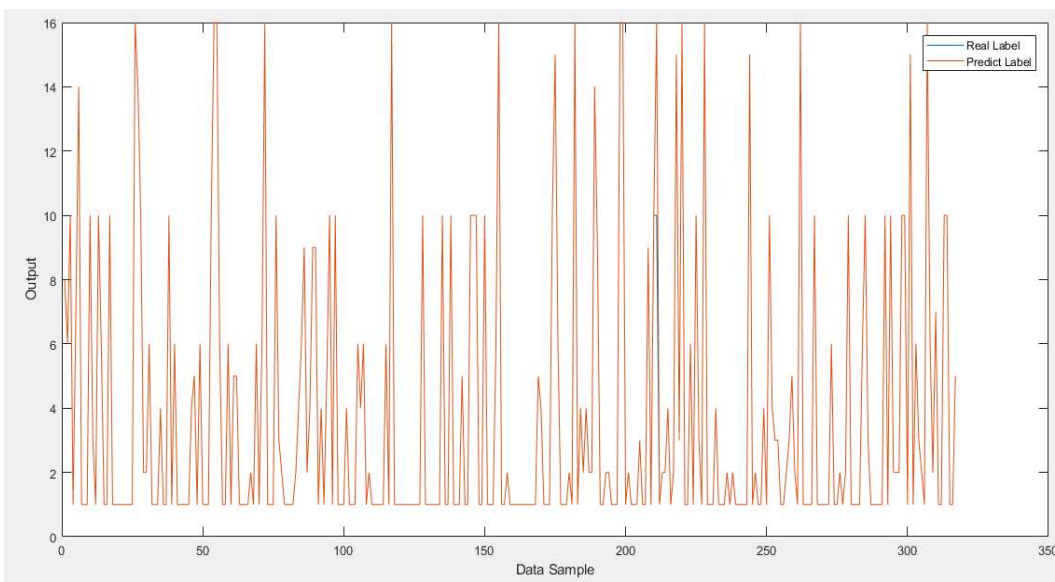
Figure 2- The outputs of the proposed method on training data of (a) Algorithm 1 and (b) Algorithm 2

In this figure, the blue curve is the actual label of the data and the red curve is the predicted output of the model.

Figure 3 (a and b) depicts the output of the proposed method on the test data. In this figure, the blue curve is the actual label of the data and the red curve is the predicted output of the model.



(a)



(b)

Figure 3- The outputs of the proposed method on the test data of (a) Algorithm 1 and (b) Algorithm 2

It should be mentioned that the more similar the blue curve is to the red one, the better the model performance. As shown in the tables, the model performed well on the training data, but it failed to maintain the same performance on the test data in Algorithm 1. In Algorithm 2, therefore, it seems that the final output of the model is more desirable. The diagrams in Figure 3 show that data mining methods can be used in the analysis of heart disease.

Conclusion

The heart and circulatory system are the most important organs in the human body. Disorders in these organs have a very

significant effect on the functioning of the whole body, because the circulatory system is responsible for providing the energy needed by all organs of the body, including the heart itself. The present study was performed to help develop cardiac arrhythmia detection algorithms using data from cardiac profile measurement systems, such as electrocardiogram signal, on the MIT/BIH database. Then, the frequency and temporal properties of the signal were studied simultaneously by examining different methods, among a commonly used updating algorithm, the GSA, was offered for this purpose. The UWMV-GSA is used to classify and predict cardiac arrhythmias. Several hybrid classification methods were compared for cardiac datasets, and each base classifier was designed as specific to each part of the dataset in

order to improve the overall accuracy of the UWMV-GSA. The proposed classification algorithm is used as the base classifications by the use of LMS, RLS, decision tree, and KNN predictors. Simulations of the results show that the UWMV-GSA improves the accuracy of classification and prediction of heart disease.

Based on the obtained results, it can be concluded that the proposed method is able to have a good performance using the potential of each base classifier. The reason for this is the identification and focus of the proposed method on a base classifier with a good performance. Based on the performance and influence of the base models in the final decision, it can be deduced that it is better to use hybrid learners to classify data in such applications where the distribution of data is not clear and it is not possible to specify conclusively the model that will yield good results on data.

Another finding is that the GSA can be used online to combine classifiers and prevent over-fitting of the final model by accurate update of the weights. This is because the behavior of base classifiers is determined in the face of new data.

According to the research results, the following propositions are recommended to further research in this field in the future:

1. The hybrid classification algorithm is an accurate method for detecting cardiac arrhythmias. Therefore, it is recommended to use this algorithm in other contexts, such as coronary arteries, heart failure, myocardial infarction, cardiac arrhythmia, cardiomyopathy, etc., which can be detected from ECG diagrams.
2. It is recommended to use other base classifiers to combine classifications to compare the obtained results with those of the base classifiers used for prediction in this research.
3. It is recommended to employ this research in the diagnosis of heart diseases in hospitals for faster diagnosis, greater accuracy, and less errors.

References

1. Najjar A, Heidari L, Parhan A. Principles of Nursing Care in CCU. Tehran, Boshra Publications. 2006.

2. Guyton AC, Hall JE. Textbook of Medical Physiology. Elsevier Saunders. 2006:103-145.
3. Zhihai Y, Weihong Y, Yan C, Taoying L. Mobile healthcare research based on Jade Agent and Android platform. In 2014 IEEE 5th International Conference on Software Engineering and Service Science 2014 Jun 27 (pp. 723-726). IEEE.
4. Freedman SB, Isner JM. Therapeutic angiogenesis for coronary artery disease. *Annals of internal medicine*. 2002 Jan 1;136(1):54-71.
5. Sepehri AA, Zakeri Moghadam M, Aliasgharpour M. Special Nursing Care in ICU, CCU, and Dialysis Wards. Tehran, Andisheh Rafi Publications. 2002.
6. Ordóñez C, Omiecinski E, De Braal L. Mining Constrained Association Rules to Predict Heart Disease. 2010.
7. Inan OT, Giovangrandi L, Kovacs GT. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. *IEEE transactions on Biomedical Engineering*. 2006 Nov 20;53(12):2507-15.
8. Gutiérrez-Gnecchi JA, Morfin-Magaña R, Lorias-Espinoza D, del Carmen Tellez-Anguiano A, Reyes-Archundia E, Méndez-Patiño A, Castañeda-Miranda R. DSP-based arrhythmia classification using wavelet transform and probabilistic neural network. *Biomedical Signal Processing and Control*. 2017 Feb 1;32:44-56.
9. Terek A, Soydan I. Clinical Arrhythmias and Preventives. Ege University Faculty of Medicine Press, @Turkish. 2012.
10. Variability Computational Method Progressing Biomed. 2015: 95-108
11. Khalaf AF, Owis MI, Yassine IA. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. *Expert Systems with Applications*. 2015 Nov 30;42(21):8361-8.
12. Raut RD, Dudul SV. Arrhythmias classification with MLP neural network and statistical analysis. In 2008 First International Conference on Emerging Trends in Engineering and Technology 2008 Jul 16 (pp. 553-558). IEEE.