GA-NN APPROACH FOR ECG FEATURE SELECTION IN RULE BASED ARRHYTHMIA CLASSIFICATION

Gözde Aslantaş, Fikret Gürgen, Albert Ali Salah[*]

Abstract: Computer-aided ECG analysis is very important for early diagnosis of heart diseases. Automated ECG analysis integrated with experts' opinions may provide more accurate and reliable results for detection of arrhythmia. In this study, a novel genetic algorithm-neural network (GA-NN) approach is proposed as a classifier, and compared with other classification methods. The GA-NN approach was shown to perform better than alternative approaches (e.g. k-nn, SVM, naïve Bayes, Bayesian networks) on the UCI Arrythmia and the novel TEPAS ECG datasets, where the GA resulted in a feature reduction of 95%. Based on the selected features, several rule extraction algorithms are applied to allow the interpretation of the classification results by the experts. In this application, the accuracy and interpretability of results are more important than processing speed. The results show that neural network based approaches benefit greatly from dimensionality reduction, and by employing GA, we can train the NN reliably.

Key words: *ECG, Genetic algorithm (GA), Neural network (NN)*

Received: March 25, 2013 **DOI:** 10.14311/NNW.2014.24.016
Revised and accepted: May 15, 2014

1. Introduction

Alterations that disrupt the regular functioning of the heart cycle may cause cardiac arrhythmia, which might be a potential reason for a heart disease. Thus, early detection of arrhythmia can save lives. Electrocardiogram (ECG) is widely used for the diagnosis of such abnormalities. ECG output must be integrated with medical assessment to provide more meaningful results, however, ECG analysis as a part of clinical assessment requires expert knowledge. Automated ECG analysis integrated with experts' opinions provides more robust and reliable results to detect abnormal patterns.

In this paper, we aim to detect arrhythmic patterns in a rule-based manner in order to aid the cardiologists. We propose a genetic algorithm-neural network

---

[*] Gözde Aslantaş – Corresponding Author, Fikret Gürgen, Albert Ali Salah, Department of Computer Engineering, Boğaziçi University, Istanbul, 34342, Turkey, E-mail: `gozdeaslantas@gmail.com, gurgen@boun.edu.tr, salah@boun.edu.tr`

approach (GA-NN) for ECG feature selection in rule-based arrhythmia classification. The GA-NN approach performs feature selection and classification simultaneously. This way, it finds relevant features that contribute most to the classification accuracy. In this approach, the neural network is the main classifier, while the genetic algorithm performs feature selection: Neural network weights are encoded in chromosomes and the optimal neural network model is selected according to its classification performance. The objective function is based on the performance of candidate networks, as determined by the classification error.

At the end of the iterative GA search, a final model is obtained, where each input feature is assigned a weight and the network topology is optimized by eliminating connections with negligible weights. Thus, we prune dimensions to obtain parsimonious representations that can be interpreted in human readable format. We are mainly interested in producing a robust and reliable arrhythmia detection system that can be used in diagnostic decision support systems. Applying rule extraction methods directly on the original, raw data will result in a large and complex rule set, which is slow to use and difficult to interpret [11]. In the proposed system, the GA-NN approach is used as a preclassifier and feature selector, and rule extraction is applied to the classifier output to produce concise rules. Since rules are only useful to the expert as long as they are clear, simple and precise, finding discriminative features is an important step for rule set construction.

The organization of the paper is as follows: Section 2 presents related work on ECG analysis. In Section 3, we derive the proposed GA-NN approach in detail. Rule extraction is explained in Section 4. Extensive experiments performed on the UCI dataset [33] and a new proprietary ECG database introduced in this paper, are presented in Section 5. In Section 6, the proposed approach is discussed and our study concludes with a summary of the empirical results and a discussion of future work.

2. Definitions and Related Work

2.1 Definitions

Electrocardiogram (ECG) is a recording of the heart rate. In clinical settings, it is typically obtained as a laboratory test result and must be integrated with a clinical assessment. Clinicians may reason differently while performing analysis during these assessments, and typically, an analysis of ECG tracing is performed under some time pressure. Consequently, an automated ECG analysis application is meant to be a supportive element for the clinicians, while they conduct their evaluations.

An ECG tracing consists of waves representing electrical activity on a scaled paper, where the time is measured on the horizontal axis, and the voltage is measured on the vertical axis. ECG features are derived from these waves, as complexes, intervals, amplitudes and regularity of the features.

Arrhythmia diagnosis is based on these ECG features. Interpretation of ECG features as a part of the analysis of ECG tracings is essential for arrhythmia recognition. Before the detection of abnormal patterns in an ECG plot, however, an understanding of the normal patterns is very important. Fig. 1 for a plot of a

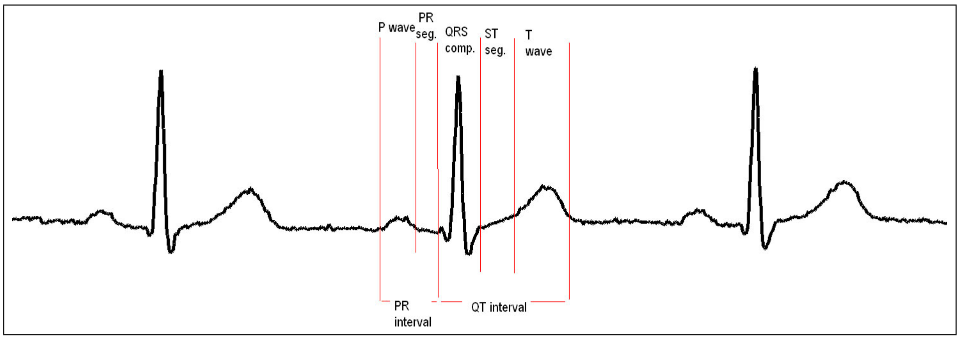

Fig. 1 *Waves on a normal ECG plot. The data have been obtained from Tepa Database [32]*

normal ECG, which also shows the convention of naming of the different intervals of the ECG plot.

The irregularity of the waves may be indicative of an abnormality in the heart beats. An example of such an abnormality is given in Fig. 2. Here, it is possible to observe the irregular R-R intervals, and narrow, indistinct QRS complexes. This pattern emerges when several ectopic pacemakers at faster rates emerge and the atrium can no longer respond to each stimulus. This generates multiple depolarization and the atrioventricular node is blocked [3].

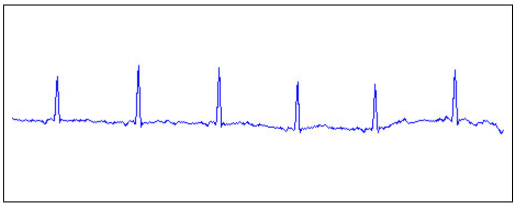

Fig. 2 *Atrial Fibrillation on ECG plot. The data have been obtained from MIT-BIH Atrial Fibrillation DB record 4126 [21] [22]*

Waves in normal ECG samples have known ranges for their parameters. Any difference in such measurements may indicate an arrhythmic pattern. Additionally, any disruption in the order of waves and missing waves might indicate abnormalities. Fig. 3 indicates a ventricular tachycardia. In this case, P waves cannot be seen in any regularity, and no waves are discernible between QRS complexes. These result in a rapid succession of ventricular premature beats and irregular ventricular rhythm.

2.2 Related Work

In the literature, there are many approaches to improve automated ECG analysis for detecting abnormal patterns. The UCI Arrhythmia Repository [33] and MIT-

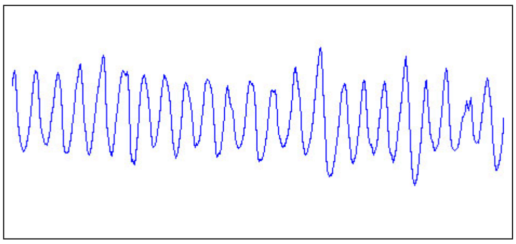

Fig. 3 *Ventricular tachycardia on ECG plot. The data have been obtained from CU Ventricular Tachycardia DB record cu01 [21] [22]*

BIH [21] are the two most commonly used databases for ECG analysis. The most important algorithmic approaches are summarized in Tab. I.

Ref.	**Year**	**Protocol**	**Algorithm**	**Results**	**Dataset**
Yeap *et al.* [36]	1990	5% tra. 95% test.	ANN BP	98.36% sens.	AHA
Hu *et al.* [15]	1993	3-fold CV	51-25-2 MLP	≈90% acc.	MIT-BIH
Chazal and Reilly [6]	2000	10-fold CV	LDA	69.3-74.7% acc.	Frank Lead ECG Data[8]
Niwas *et al.* [23]	2005	58% tra.42% test.	ANN	≈99% acc.	MIT-BIH
Zhang and Zhang [37]	2005	2/3 tra.1/3 test.	PCA-SVM	≈99% acc.	MIT-BIH
Song *et al.* [31]	2005	CV	LDA-SVM	99.35% avg. acc.	MIT-BIH
Uyar [34]	2006	10-fold CV	PCA-SVM	83.7% acc.	UCI
Kara and Okandan [18]	2007	24 of 72 normal - 28 of 52 AF signals test.	ANN BP	100% acc	MIT-BIH
Oliveira *et al.* [24]	2010	75% tra. 25% test.	Bayesian Networks	≈99% sens.	MIT-BIH and QT database
Jadhav *et al.* [16]	2011	90% tra. 10% test.	ANN BP	86.67% sens.	UCI
Froese *et al.* [10]	2005	CV	GA-NN	99.6% avg. acc.	MIT-BIH
This work	-	10-fold CV	GA-NN	89% acc. 85% acc.	UCI, TEPA

Tab. I *Previous work on arrhythmia classification.*

Typically, a preprocessing step is required to detect characteristic waves in ECG signal data. Pan and Tompkins developed an algorithm to detect QRS complexes, based on a dual threshold technique, but this approach is only useful if the heart rate is regular [26]. An optimized C language implementation of QRS detection, known as the Hamilton-Tompkins algorithm improves the dual thresholding by discrimination of the T wave [13].

Wavelet transform analysis is also widely used to detect P waves, QRS complexes and T waves in the ECG signal [27, 31]. Lin *et al.* used a Morlet wavelet decomposition to extract features [19]. Benitez *et al.* presented a robust method for QRS detection, which uses a variable threshold determined automatically from the root mean squared (RMS) value of the data [4]. By using the first differential of the ECG signal and its Hilbert transformed data, R waves were successfully differentiated from T and P waves. Kara and Okandan extracted Daubechies wavelet coefficients and trained neural networks for successful discrimination of atrial fibrillation [18]. In the study by Oliveira *et al.* [24], features such as distance between two consecutive QRS complexes and QRS complex shape are combined with a hidden Markov model based framework, developed earlier by Andreão *et al.* [2].

For the arrhythmia detection problem, neural network based methods are very common. Niwas *et al.* studied arrhythmia classification using a multilayer feedforward neural network trained with the backpropagation algorithm [23], whereas [18] used Levenberg-Marquard based backpropagation. Jadhav *et al.* performed a comparison study with multilayer perceptrons (MLP), generalized feedforward neural network, and modular neural network models using the UCI arrhythmia dataset [16]. Among the three approaches, MLP was found to have the best performance. It is also possible to use multiple classifiers in tandem. As an example of this, Hu *et al.* used MLPs in a cascade for beat classification of one normal and 12 abnormal classes [15]. The first MLP is used for classifying normal vs. abnormal beats, and the second level is used to categorize abnormal beats into one of the 12 abnormal classes. It is shown that the composite MLP classifier works better than a single MLP for the multiclass problems, but the overall performance principally relies on how well the first neural network performs. Additionally, Froese *et al.* performed a comparison study with Linear Discriminant Analysis (LDA) and feedforward MLP, using either backpropagation or genetic algorithms for the training process [10]. The study reports 99.6% average accuracy on the MIT-BIH database for distinguishing between the atrial beat types and the ventricular beat types.

Another general approach in ECG classification is the usage of support vector machines (SVM). Typically, SVM classifiers are used in conjunction with dimension reduction techniques. Uyar *et al.* studied arrhythmia classification using SVM with principal component analysis (PCA) and independent component analysis (ICA) [34]. According to experiments, PCA was a better feature extraction choice compared to the ICA. Song *et al.* used linear discriminant analysis (LDA) for dimensionality reduction before applying SVM [31]. For the discrimination of six types of arrhythmia beats, SVM with LDA was shown to perform better than SVM with PCA.

3. GA-NN Approaches

While backpropagation is applicable for training neural networks (NN), Genetic Algorithms (GA) can be used to perform training [1] or joint training and model selection [25, 17, 38]. In this study, we extend the work in this area and propose a novel GA-NN approach.

The main idea behind the GA-NN approach is to harness the global search ability of the GA in ECG feature space, while performing arrhythmia classification with the neural network. In this way, the aim is to find the most effective ECG features for determining arrhythmia patterns. The type of the combination used in this study is called *supportive combination* by Schaffer *et al.* [29], because the genetic algorithm is used to assist the neural network in feature selection and weight learning.

The number of hidden units can be determined by the experiments performed for parameter selection. In this work, we set the number of hidden units to bes half the number of input dimensions, and fix the network topology. The sigmoid function is used in the hidden layer to introduce non-linearity. Each individual in the GA population encodes one candidate solution, and is assigned a fitness value, which is an assessment that indicates how good the solution is. Fitness values are calculated based on the classification performance of the solutions. During the GA iterations, NN weights are optimized and fitness values of the solutions are computed. The possibility of the solutions to be in the next generation is inversely proportional to their fitness values. Parent solutions are selected by using roulette wheel selection, crossover is applied to produce offspring, and these offspring are mutated so as to provide variation. Additionally, a second mutation operator is applied to prune irrelevant features from the network by zeroing their weights as described by Sexton *et al.* [30]. The best offsprings produced as a result of these operations comprise 90% of the next generation. The remaining part is filled by the best solutions from the old generation in order to preserve the best individuals. The solution with the best fitness value in a generation is compared with the one from the old generation. The iterations are stopped when the best solution of the current generation is not better than the best solution of the previous generation.

As a candidate solution, a neural network is evaluated according to its performance on validation data. The performance is generally related to the difference between the actual output and the predicted output. The fitness value obtained from the fitness function is given as [30]:

$$f = \sum (O_i - \hat{O}_i)^2 + C\sqrt{\frac{\sum_{i=1}(O_i - \hat{O}_i)^2}{N}},\ O_i \in \{0,1\} \tag{1}$$

where O_i is the target class, and $\hat{O}_i$ is the estimated class of the instance i, and N is the number of instances. Here, C represents the number of nonzero weights in the network. In selection, each individual in the population is assigned a probability based on its fitness value. The probability is computed as:

$$P(X = x) = \frac{f_{cur} - f_{bad}}{\sum_{i=1} f_i - f_{bad}} \tag{2}$$

where f_i indicates the fitness value of the solution i, f_{bad} is the worst fitness value and f_{cur} is the fitness value of the current solution.

Given a neural network topology, we use the backpropagation algorithm for setting the initial weights, in order to reduce the error to a predetermined level and to start the GA from a decent initial population. Training parameters like the maximum fail count, hidden node count, mutation rate, as well as the number of neurons in the hidden layer are determined by evaluating the overall accuracy. Sensitivity analysis is performed to assess the effects of parameter changes on the overall accuracy, and the results are illustrated in Fig. 4 and Fig. 5. All networks in a population are tested using the same parameters. Our experimental results, given in Section 5, show that the GA-NN approach improves significantly over NN trained with backpropagation only.

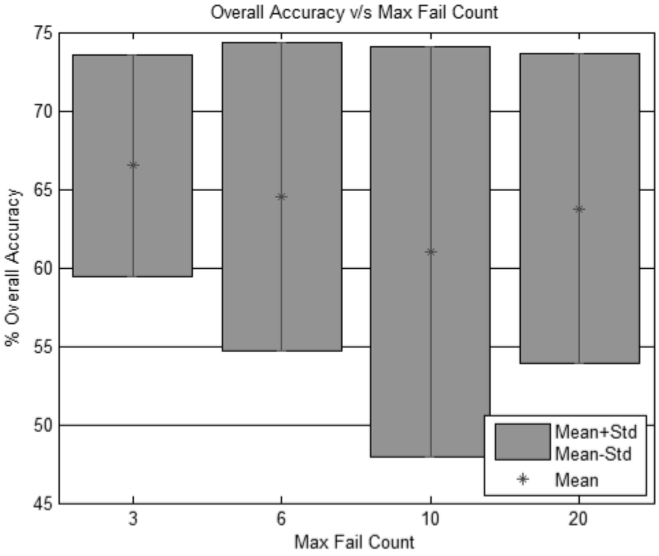

Fig. 4 *Max. fail selection versus overall accuracy.*

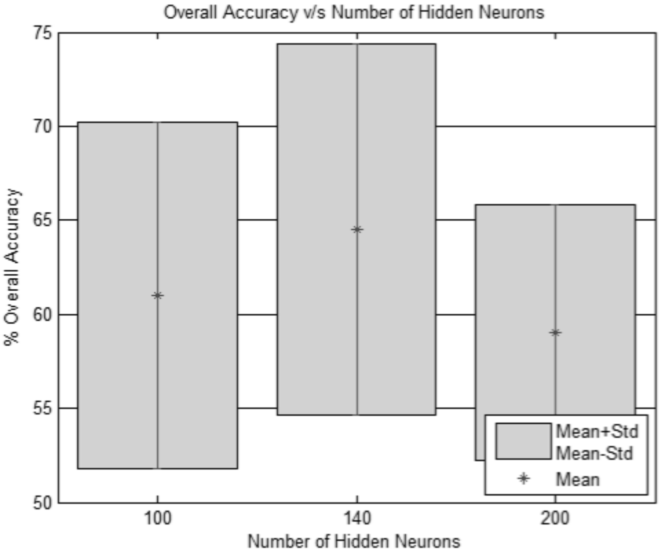

Fig. 5 *Number of neurons in hidden layer versus overall accuracy.*

To detect the presence or the absence of arrhythmia accurately, it is important to fine-tune the network parameters. The data are divided into training, test and validation sets for model selection. In order to select parameter values for max. fail count, hidden node count, initial population number etc., we build neural network populations with different parameter values. Then we apply GA steps for each NN population to obtain a final model with the optimized weights. The final model selection is performed on the validation set, whereas each model is optimized on the training set. Feature selection and classification are jointly performed for a given parameter set.

4. Rule Extraction

The last stage of the system is the rule extraction stage, since our main interest is to aid cardiologists with medical assessments in producing robust and reliable arrhythmia classification. There are several approaches in the literature that focus on rule generation for arrhythmia, and some of these use features that are known to be discriminative for this problem. We contrast several approaches with the proposed method, and test different rule extraction methods. Additionally, the output of each rule extraction method is tested on a validation set and performance metrics such as accuracy, sensitivity, f-score, precision and MCC (Matthews correlation coefficient) are calculated, in order to assess how well these rules depict the characteristic differences between arrhythmia and normal cases. For rule extraction the following methods are contrasted:

- *C4.5* is a decision tree algorithm that uses information entropy for deciding the best split at each level of the tree. Rules are generated by adding the conditions on each path from each leaf node to the root node [28].
- In *RIPPER*, the rule set is produced by repeatedly adding rules that contain features with the highest information gain. Post-pruning is performed for optimization [5].
- *PART* builds a partial C4.5 decision tree in each iteration and the leaf node having the maximum coverage is selected as the rule of that iteration. Following iterations are performed on the instances that are not covered by the previous rule [9].

The main flow of the system that uses rule extraction is summarized in Fig. 6.

5. Experimental Setup

5.1 Datasets

The first dataset used in this study is obtained from the UCI Repository [33]. The UCI Arrhythmia dataset originally contains 452 instances with 279 attributes. There are 16 arrhythmia cases associated with each instance. 15 cases indicate anomalies, and one case shows a normal rhythm. The class distribution of this dataset is very unbalanced, and instances of classes 11, 12 and 13 are lacking in

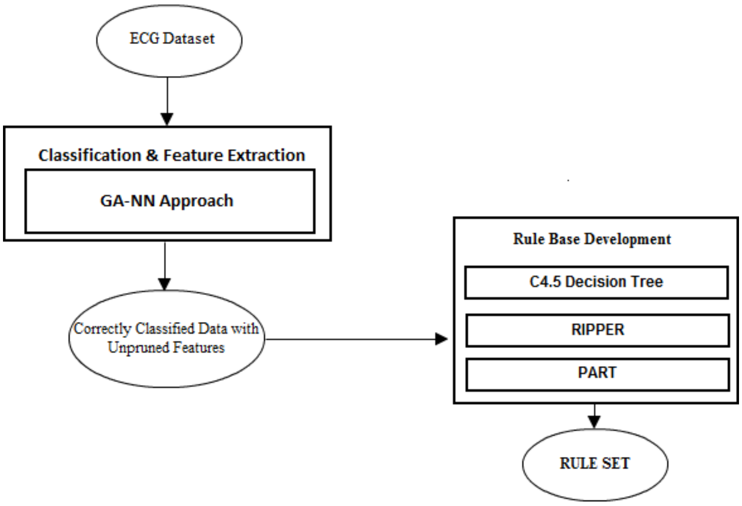

Fig. 6 *Main flow of the system.*

the dataset, while a majority of instances belong to the normal rhythm class. To have better sensitivity rates, the dataset is preprocessed. Initially, all anomalies are grouped into one class as "abnormal rhythm", and a normal-abnormal rhythm classification problem is obtained.

About 0.33% of the feature values in the dataset are missing. Samples with missing values among the features are removed. This leads to a reduction to 278 input features and 420 instances. Data are normalized to a range of −1 and +1 in order to provide robustness.

The second dataset is a novel ECG dataset, which we use for database-independent verification of the performance. The data are obtained with the Kardiosis ECG Tool of the manufacturing firm TEPA [32]. This dataset is smaller than the UCI Arrhythmia dataset. There are 20 records, 13 of which are normal and the rest abnormal.

5.2 Classifiers

In our proposed system, weights of neural networks are encoded in a chromosome. The structure of the neural networks is comprised of one input layer, one hidden layer and one output layer. Each network is initially trained with a scaled conjugate gradient backpropagation algorithm. Since there are 278 input features and 140 hidden units in each neural network, we have 39.200 parameters per network to optimize. As stated before, irrelevant input features are determined by their relative contributions to the classification accuracy on the training set, and they are pruned.

A population of neural networks with the same parameter set is initially created. Genetic algorithm steps are applied to the populations of 30 individuals, iteratively. Parent selection is performed by roulette wheel selection. Crossover rate is chosen as 90% and mutation rate is set as 10%. Fitness values of the offsprings are obtained from the objective function given in the previous section and based on the fitness value evaluation, probability of each solution to be in the next generation is computed. 90% of the new population is filled by the offsprings with the highest probability. The rest of the population is completed by the best solutions from the old generation.

A set of automatically computed ECG features are presented as inputs to the neural network, and the output nodes indicate class labels, which are "normal" and "abnormal". The encoding scheme is given in Fig 7.

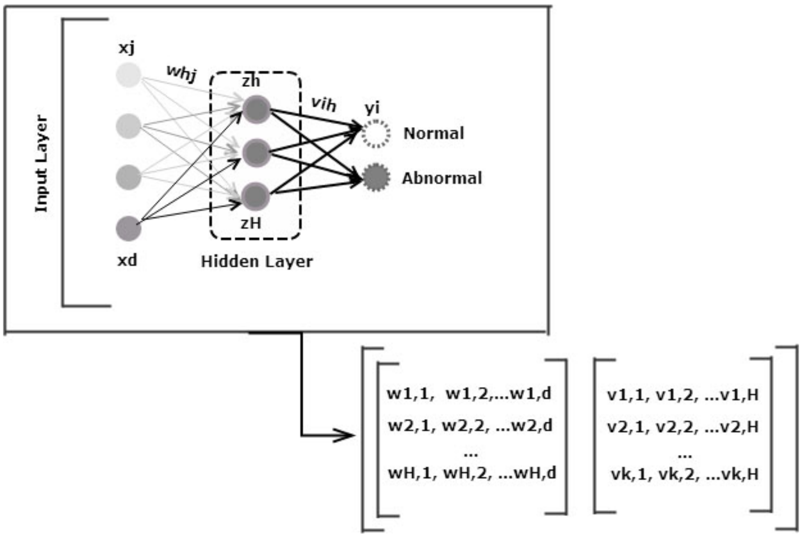

Fig. 7 *Encoding scheme of a chromosome.*

Several general classification schemes used in the literature for arrhythmia detection problem are implemented for comparison in this study, including k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), Naive Bayes, Bayesian Networks and Neural Networks. k-NN and SVM classifiers represent distance based methods for arrhythmia classification, whereas naive Bayes and Bayesian Networks are probabilistic approaches. Dimensionality reduction techniques such as recursive feature extractor (RFE-SVM), correlation based feature selection (CFS), principal component analysis (PCA) and factor analysis (FA), are used before these classifiers. The classifiers are applied to both the original dataset and the reduced dataset to observe the effect of dimensionality reduction. The flow diagram of the comparison is given in Fig. 8.

5.3 Performance Measures

In this study, the main performance measure is sensitivity. Sensitivity is also called recall, the positive rate, or hit rate. It is the proportion of correct detections of the target class (called true positives, or TP for short) to the total number of samples

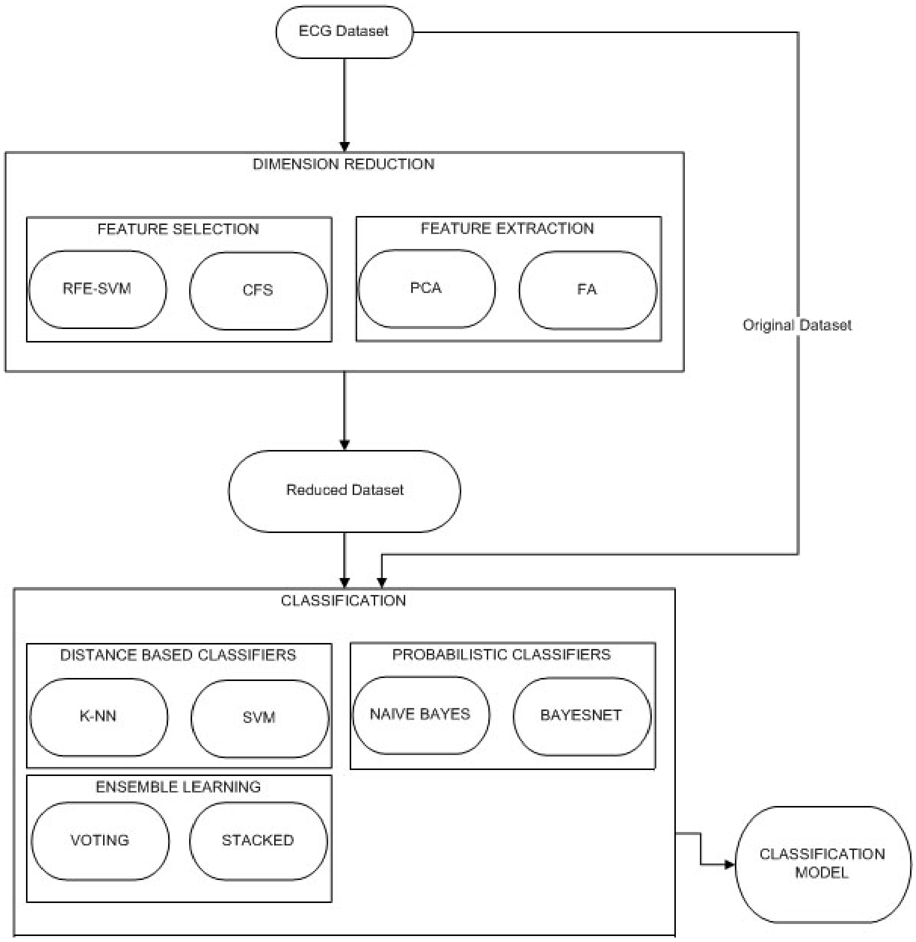

Fig. 8 *General classification schemes used in the literature for the arrhythmia detection problem.*

of the target class. The latter is composed of detected (TP) and missed (false negatives, or FN for short) samples.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

The second important metric is the f-score, which is the harmonic mean of Recall and Precision. Precision represents the ability of producing the same results under changing conditions:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$F - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

In addition to these, the Matthews Correlation Coefficient (MCC) measure is calculated, which indicates the quality of a classifier for two-class problems, taking both positive and negative measures into account.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{6}$$

5.4 Results

We have used 10-fold cross-validation in our experiments. On the UCI Arrhythmia dataset, 86.75% accuracy is obtained with the proposed GA-NN scheme. There is also approximately 95% decrease in the number of features. Only 12 input features remain from the original 278 after pruning.

For comparison purposes, we have implemented k-NN, SVM, Naive Bayes and BayesNet methods, together with the four dimensionality reduction approaches mentioned in the previous section. For the k-NN classifier, 5-15 neighbors were tested. We have also tested stacked and voted combinations of k-NN and SVM to assess classifier combination. However, the amount of training data is limited, and the complexity induced by classifier combination results in overlearning.

Each classifier is evaluated based on the performance metrics for accuracy, sensitivity, f-score, precision and MCC. The results of classifiers for the reduced data sets and the original dataset are given in Tab. II (in terms of sensitivity), Tab. III (in terms of f-score), Tab. IV (in terms of MCC), and Tab. V (in terms of accuracy). In each table, the last column is the unreduced dataset, and the first four columns are dimensionality-reduced datasets.

According to these results, the proposed GA-NN approach provides the highest hit rate for the target class. Naive Bayes and Bayesian methods have similar performance. As a proof of the efficiency of feature selection, GA-NN outperforms Naive Bayes on the original dataset.

The GA-NN method produces higher f-scores than the competing methods. In terms of taking both posivite and negative classes into account, GA-NN provides the highest results. Probabilistic approaches perform better than k-NN and SVM.

	RFE-SVM	**CFS**	**FA**	**PCA**	**Original**
k-NN$_5$	**0.7179**	0.5897	0.6410	0.5128	0.3718
SVM	0.3163	0.3571	0.3061	0.3163	0.3061
NaiveBayes	0.5102	0.6186	0.7143	0.5714	0.7158
BayesNet	0.6596	0.7021	0.7128	0.5638	0.6702
Neural Network	0.536	0.530	0.612	0.541	0.560
GA-NN	–	–	–	–	0.9646

Tab. II *Comparison of classifiers in sensitivity.*

	RFE-SVM	CFS	FA	PCA	Original
k-NN$_5$	0.7044	0.6715	0.6329	0.5333	0.4328
SVM	0.3875	0.3933	0.3750	0.3713	0.3681
NaiveBayes	0.6667	0.7317	0.7778	0.6667	0.7083
BayesNet	0.7425	**0.7904**	0.7701	0.7067	0.7636
Neural Network	0.639	0.655	0.707	0.643	0.65
GA-NN	-	-	-	-	0.8916

Tab. III *Comparison of classifiers in f-score.*

	RFE-SVM	CFS	FA	PCA	Original
k-NN$_5$	0.5247	0.5281	0.4117	0.2753	0.1827
SVM	0.0432	-0.0459	0.0223	-0.0244	-0.0069
NaiveBayes	0.5691	0.5942	0.6209	0.4725	0.4562
BayesNet	0.5897	0.6701	0.6151	0.6049	0.6320
Neural Network	0.5759	0.5980	0.6596	0.5275	0.461
GA-NN	-	-	-	-	0.7375

Tab. IV *Comparison of classifiers in MCC.*

	RFE-SVM	CFS	FA	PCA	Original
k-NN$_5$	0.7762	0.7857	0.7238	0.6667	0.6381
SVM	0.5333	0.4857	0.5238	0.5000	0.5095
NaiveBayes	0.7619	0.7895	0.8095	0.7333	0.7295
BayesNet	0.7952	**0.8333**	0.8095	0.7905	0.8143
Neural Network	0.790	0.757	0.779	0.738	0.736
GA-NN	-	-	-	-	0.8675

Tab. V *Comparison of classifiers in accuracy.*

When we consider the ratio of correctly classified instances, GA-NN significantly outperforms competing methods with 86.75% accuracy. Finally, the GA-NN approach is tested on the ECG dataset obtained from TEPA [32]. An accuracy of 85% is obtained.

According to the experiments, the proposed approach outperforms others in terms of all performance metrics. The results show that dimensionality reduction has a beneficial effect on classification. An increasing trend in sensitivity is seen among classifiers when dimensionality-reduced data sets are used.

5.5 Rule Extraction

Based on the selected features, rule sets are extracted using the C4.5, RIPPER, and PART algorithms. Performance metrics of accuracy, sensitivity, f-score, precision and MCC are also computed for the rule extraction methods. The comparison of the rule extraction methods is given in Tab. VI.

Method	**Accuracy**	**Sensitivity**	**F-score**	**Precision**	**MCC**
C4.5	87.8%	81.0%	85.2%	89.8%	75.0%
RIPPER	88.5%	85.6 %	86.6%	87.6%	76.5%
PART	94.2%	87.9%	93.0%	98.7%	88.6%

Tab. VI *Performance results of the rule sets.*

According to these measures, PART provides the highest values in all metrics. C4.5 and RIPPER have similarities, but RIPPER outperforms C4.5 especially in terms of sensitivity, which is related to its ability to identify the target class. The resulting rules produced by PART are given in Tab. VII. (For an explanation about the individual features, please check the UCI database manual).

1	If V1_Avg_QRSA $>$ 1 then CLASS = ABNORMAL
2	If AVL_Amplitude_T_wave $\leq$ -0.8 and V1_N_intrinsic_deflections $\leq$ 8 then CLASS = ABNORMAL
3	If heartrate $\leq$ 57 and Tinterval $\leq$ 165 then CLASS = ABNORMAL
4	If Tinterval $>$ 221 and DI_Avg_QRSTA $\leq$ 25.5 then CLASS = ABNORMAL
5	If heartrate $>$ 94 and Tinterval $>$ 148 then CLASS = ABNORMAL
6	If V3_Avg_width_S_wave $\leq$ 28 and V3_Amplitude_S_wave $>$ -6.8 and DI_Avg_QRSTA $>$ 17.7 then CLASS = ABNORMAL
7	If V1_N_intrinsic_deflections $>$ 24 and V3_Avg_width_S_wave $>$ 40 then CLASS = ABNORMAL
8	If QRSduration $>$ 107 and V1_Avg_QRSA $\leq$ -25 then CLASS = ABNORMAL
9	If V2_Avg_width_S_wave $>$ 44 and V1_N_intrinsic_deflections $>$ 4 and heartrate $>$ 64 and V3_Avg_width_S_wave $\leq$ 56 then CLASS = ABNORMAL

Tab. VII *Rules selected by the PART algorithm.*

The resulting rule set consists of nine rules. The selected features are the QRSA of the V1 Lead, amplitudes of S and T waves, the heart rate, duration of the T wave and the QRS complex, the width of the S wave, the QRSTA of the D1 Lead and the number of intrinsic deflections. The rules are not shown in an order of importance, because PART produces rules from repeatedly generated partial decision trees. We can assume that these rules contain complementary information.

6. Conclusions

In this work, we proposed a GA-NN approach for ECG-based arrythmia detection, followed by rule extraction. We have obtained a small set of interpretable rules with high predictive performance. With the proposed GA-NN approach, we performed feature reduction and classification simultaneously. By doing so, features that contribute most to classification are determined. We observed that utilizing the global search ability of the genetic algorithm with the neural network as a classifier improves the performance over using neural networks directly, even in conjunction with dimensionality reduction techniques. The GA-NN approach considerably reduces the dimensionality of the data, and trains the NN reliably.

We performed comparisons with state of the art classification schemes, and dimensionality reduction methods. The proposed approach outperforms others according to the experimental results.

Additionally, a novel ECG dataset is introduced and used for verifying the proposed approach. We have obtained high accuracy rates on this new dataset. Several rule extraction algorithms were applied to interpret the classification results in a human readable format. The resulting rule set is concise and useful. Obviously, clinical assessment is indispensable for this problem, but our results indicate that automated analysis can give real-time assessment help to the clinicians.

References

[1] Amma N. G. B.: Cardiovascular Disease Prediction System Using Genetic Algorithm and Neural Network, 2012 International Conference on Computing, Communication and Applications (ICCCA), 2012, pp. 1-5.

[2] Andreão R.V., Dorizzi B., Boudy J.: ECG Signal Analysis through Hidden Markov Models, IEEE Transactions on Biomedical Engineering, 53, 8, 2006, pp. 1541-1549.

[3] Becker E. B.: Fundamentals of Electrocardiography Interpretation, Anesthesia Progress, 53, 2, 2006, pp. 53-64.

[4] Benitez D., Gaydecki P. A., Zaidi A., Fitzpatrick A. P.: The Use of the Hilbert Transform in ECG Signal Analysis, Computers in Biology and Medicine, 31, 5, 2001, pp. 399-406.

[5] Cohen W. W.: Fast Effective Rule Induction, In Proceedings of the Twelfth International Conference on Machine Learning, Morgan Kaufmann, 1995, pp. 115-123.

[6] de Chazal P., Reilly R.B.: A Comparison of the ECG Classification Performance of Different Feature Sets, Computers in Cardiology 2000, 2000, pp. 327-330.

[7] Downey T. J. Jr., Meyer D. J., Price R. K., Spitznagel E. L.: Using the Receiver Operating Characteristic to Asses the Performance of Neural Classifiers, International Joint Conference on Neural Networks (IJCNN '99), 5, 1999, pp. 3642-3646.

[8] Frank E.: An Accurate, Clinically Practical System for Spatial Vectorcardiography, Circulation, 13, 5, 1956, pp. 737-749.

[9] Frank E., Witten I. H.: Generating Accurate Rule Sets Without Global Optimization, Proceedings of the Fifteenth International Conference on Machine Learning, ICML '98, Morgan Kaufmann Publishers Inc., 1998, pp. 144-151.

[10] Froese T., Hadjiloucas S., Galvão R. K. H., Becerra V. M., Coelho C. J. : Comparison of extrasystolic {ECG} signal classifiers using discrete wavelet transforms, Pattern Recognition Letters, 2006, 27, 5, pp. 393-407.

[11] Halford G. S., Wilson W. H., Phillips S.: Processing Capacity Defined by Relational Complexity: Implications for Comparative, Developmental, and Cognitive Psychology, Behavioral and Brain Sciences, 21, 6, 1998, pp. 803-831.

[12] Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. H.: The WEKA Data Mining Software: an Update, ACM SIGKDD Explorations Newsletter, 11, 1, 2009, pp. 10-18.

[13] Hamilton P. S., Tompkins W. J.: Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database, IEEE Transactions on Biomedical Engineering, 12, 1986, pp. 1157-1165.

[14] Homaeinezhad M. R., Atyabi S. A., Tavakkoli E., Toosi H. N., Ghaffari A., Ebrahimpour R.: ECG Arrhythmia Recognition via a Neuro-SVM-KNN Hybrid Classifier with Virtual QRS Image-Based Geometrical Features, Expert Systems with Applications, 39, 2, 2012, pp. 2047-2058.

[15] Hu Y. H., Tompkins W. J., Urrusti J. L., Afonso V. X.: Applications of Artificial Neural Networks for ECG Signal Detection and Classification, Journal of Electrocardiology, 26, 1993, pp. 66-73.

[16] Jadhav S.M., Nalbalwar S. L., Ghatol A. A.: Artificial Neural Network Based Cardiac Arrhythmia Disease Diagnosis, Process Automation, Control and Computing (PACC), 2011 International Conference on, 2011, pp. 1-6.

[17] Jiang W., Kong S. G., Peterson G. D.: ECG Signal Classification Using Block-Based Neural Networks, 2005 IEEE International Joint Conference on Neural Networks (IJCNN '05), 2005, pp. 326-331.

[18] Kara S., Okandan M.: Atrial Fibrillation Classification with Artificial Neural Networks, Pattern Recognition, 40, 11, 2007, pp. 2967-2973.

[19] Lin C. H., Du Y. C., Chen T.: Adaptive Wavelet Network for Multiple Cardiac Arrhythmias Recognition, Expert Systems with Applications, 34, 4, 2008, pp. 2601-2611.

[20] Mathworks, http://www.mathworks.com, 2011.

[21] MIT-BIH Arrhythmia Database, http://physionet.org/physiobank/database/html/mitdbdir/, 1997.

[22] Goldberger A. L., Amaral L. A. N., Glass L., Hausdorff J. M., Ivanov P. Ch., Mark R. G., Mietus J. E., Moody G. B., Peng C.-K., Stanley H. E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, Circulation 2000, 101(23), e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]

[23] Niwas S. I., Kumari R. S. S., Sadasivam V.: Artificial Neural Network Based Automatic Cardiac Abnormalities Classification, Proceedings of the Sixth International Conference on Computational Intelligence and Multimedia Applications, 2005, pp. 41-46

[24] Oliveira L. S. C., Andreão R. V., Sarcinelli-Filho M.: The Use of Bayesian Networks for Heart Beat Classification, Brain Inspired Cognitive Systems 2008, 657, 2010, pp. 217-231.

[25] Ölmez T.: Classification of ECG Waveforms by Using RCE Neural Network and Genetic Algorithms, Electronics Letters, 33, 18, 1997, pp. 1561-1562.

[26] Pan J., Tompkins W. J.: A Real-Time QRS Detection Algorithm, IEEE Transactions on Biomedical Engineering, 3, 1985, pp. 230-236.

[27] Park K., Lee K., Yoon H.: Application of a Wavelet Adaptive Filter to Minimise Distortion of the ST-Segment, Medical and Biological Engineering and Computing, 36, 5, 1998, pp. 581-586.

[28] Quinlan J. R.: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc., 1993.

[29] Schaffer J.D., Whitley D., Eshelman L. J.: Combinations of Genetic Algorithms and Neural Networks: a Survey of the State of the Art, 1992 International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN-92), 1992, pp. 1-37.

[30] Sexton R. S., Dorsey R. E., Sikander N. A.: Simultaneous Optimization of Neural Network Function and Architecture Algorithm, Decision Support Systems, 36, 3, 2004, pp. 283-296.

[31] Song M. H., Lee J., Cho S. P., Lee K. J., Yoo S. K.: Support Vector Machine Based Arrhythmia Classification Using Reduced Features, International Journal of Control, Automation and Systems, 3, 4, 2005, pp. 571-579.

[32] TEPA Inc http://www.tepa.com.tr/, 2007.

[33] UCI Machine Learning Repository, http://archive.ics.uci.edu/ml, 2010.

[34] Uyar A.: Intelligent Arrhythmia Classification Based on Support Vector Machines, unpublished M.S. thesis, Boğaziçi University, 2006.

[35] Welch P.: The Use of Fast Fourier Transform for the Estimation of Power Spectra: a Method Based on Time Averaging over Short, Modified Periodograms, IEEE Transactions on Audio and Electroacoustics, 15, 2, 1967, pp. 70-73.

[36] Yeap T. H., Johnson F., Rachniowski M.: ECG Beat Classification by a Neural Network, Engineering in Medicine and Biology Society, 1990., Proceedings of the Twelfth Annual International Conference of the IEEE, 1990, pp. 1457-1458.

[37] Zhang H., Zhang L.: ECG Analysis Based on PCA and Support Vector Machines, International Conference on Neural Networks and Brain, 2005. ICNN&B '05., 2, 2005, pp. 743-747.

[38] Zhou J., Li L.: Using Genetic Algorithm Trained Perceptrons with Adaptive Structure for the Detection of Premature Ventricular Contraction, Computers in Cardiology 2004, 2004, pp. 353-356.