Arrhythmia Identification from Time Series Biosignal by Neural Network Classifier with Bayesian Framework

Dayong Gao, Michael Madden, Michael Schukat,
Des Chambers, and Gerard Lyons
Department of Information Technology
National University of Ireland
Galway, Ireland

Abstract
This paper presents an ANN-based diagnostic system for arrhythmia using Neural Network Classifier with Bayesian framework by time series biosignals. The Neural Network Classifier is built by the use of logistic regression model and back propagation algorithm. The prediction performance in training and test phases is evaluated by the False Rate. The dual threshold method is applied to determine diagnosis strategy and suppress false alarm signals. The results show that more than 90% prediction accuracy could be obtained using the improved methods in the study. Hopefully, the system can be further developed and fine-tuned for practical application.

1 Introduction

Electrocardiogram (ECG) is the important time series biosignal and the major diagnostic tool for cardiologists. The ECG signal provides almost all information about the electrical activity of the heart. Continuous ECG monitoring can observe cardiac variations over an extended period of time at the bedside or in ambulatory cases. This provides more information to physicians, increases the understanding of the patients' circumstances and allows a more reliable diagnosis for cardiac abnormalities.

The ECG, however, is a nonlinear signal generated from a nonlinear system - human body. Detection of abnormal ECG rhythm is a critical step in administering aid to the subject. Often, patients are hooked up to cardiac monitors in hospital continuously. This requires continuous monitoring by the physicians as well. Due to the large number of patients in intensive care units and the need for continuous observation of such conditions, several methods for automated arrhythmia detection have been developed in the past several decades to attempt to solve this problem, including Bayesian [1] and heuristic approaches [2], expert systems [3], Markov models [4], self-organizing map [5], and neural networks [6].

In general, past approaches, according to published results, seem to suffer from common drawbacks that depend on high sensitivity to noise included in the data and unreliability in dealing with new or ambiguous patterns. In clinical domains, we had to face the problems for realizing classifiers that are able to deal even with nonlinear discrimination between classes, to accept incomplete or ambiguous input patterns, and to suppress false alarm signals for different diagnosis strategy. It is necessary to develop new detection schemes with high level of accuracy, or equivalently, low false-positive and false-negative statistics in practical application. Guvenir *et. al.* [7] developed a supervised machine learning algorithm for arrhythmia analysis based on a technique called Features Interval for the unbalanced dataset [8] with missing features and unlabeled classes to solve this problems and achieved an accuracy of 62%. However, the performance of such method is not satisfactory enough for clinical uses. So far, implementations of ANN-based ECG classification schemes have been focused on problems of narrow clinical domains.

Since artificial neural networks are inherently nonlinear model, neural network method is potentially useful in the area of ECG analysis. Various techniques for neural networks classifier and its combination with other methods have been used to improve classification results, such as Fourier-transform Neural Networks [9], Recurrent Neural Networks [10], and Back Propagation (BP) Neural Networks [11]. The Bayesian approach has been known for some time, but only recently has it been started to infiltrate different areas of science and technology systematically, with useful results [12], [13], [14], [15]. With newer knowledge available from the theory of ANNs and combination with the variety of other methods, the stage is now set for moving towards a practical approach in designing ANN-based clinical ECG interpretation systems.

This paper present an ANN-based diagnostic system for arrhythmia with ECG Signals. Such system can determine the patient's current condition in real-time. Artificial Neural Network (ANN) is used to generate a pattern recognition model based on given {input, output} sets to classify future input sets for arrhythmia diagnosis as a system module. The logistic regression model and BP algorithm are used to build the Neural Network Classifier for arrhythmia detection with ECG signals. The dual threshold method is used to suppress false alarm signals. The method used in this study aims to produce a system which perform well practically.

The rest of this paper is organized as follows. The next section outlines the proposed system and presents the methodologies for ECG diagnostic system using Neural Network Classifier in our study. The third section discusses the empirical results. The fourth section presents a brief summary of our findings and the future work.

2 Methodological Consideration

2.1 System Scheme and Data Acquisition

The proposed system consists of three basic modules: Server, Client Machine and BAN-Hub. The elements are as follows:

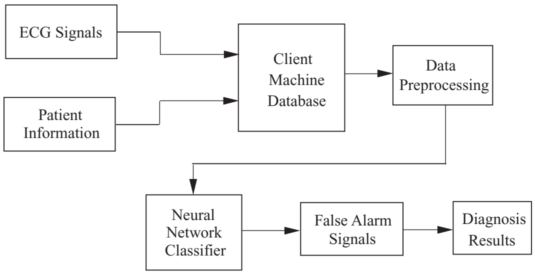

Figure 1: Block Diagram of ANN-based Diagnostic System for Arrhythmia with ECG Signals

- Server
  The server provides the backbone of the network. It gathers patient data, sends control information to the BAN-Hub and establishes the precise location of each BAN-Hub.

- Client Machine
  The Client machine provides the medical and administrative interface to the network. It provides a patient registration mechanism for each BAN-Hub, indicates alarm conditions (either unit or patient) in each BAN-Hub.

- BAN-Hub
  The BAN-Hub forms part of a wireless Body Area Network (BAN) used to monitor key patient biosignal data. The BAN-Hub is the patient interface to this network. The BAN-Hub is attached to the patient using either a shoulder strap, waist or other attachment form.

The system scheme is illustrated by the block diagram in Fig. 1. The ECG rhythm and patient data are gathered and sent to the Client Machine. The information is transmitted using the wireless network to the server. This stored digital data can then be processed to detect firstly the various complexes and then detect specified arrhythmia. An ECG complex represents the electrical events occurring in one cardiac cycle. A complex consists of five waveforms labeled with the letters P, Q, R, S, and T, as shown in Fig. 2. Because ECG is very helpful in the diagnosis of cardiac disease, it is suggested that the ECG signals are used in the initial stages of the project.

The dataset used in the study consists of 452 ECG recordings from UCI database [8]. The dataset includes about 0.33% missing attribute values and 22 unclassified classes, so the prediction accuracy of our model will be influenced. However, such condition is much closer to the real-world dynamic environment when the noisy data are acquired in our project. Therefore, The system could be easily extended to actual ECG signals, which are in progress in our project. Each record consists of a set of clinical parameters measured on ECG signals

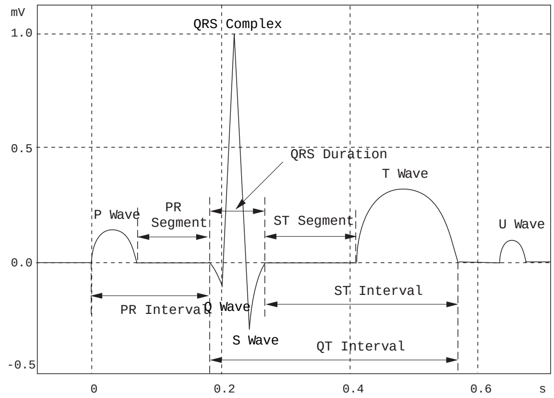

Figure 2: ECG signals. P wave: the sequential activation (depolarization) of the right and left atria. QRS complex: right and left ventricular depolarization. U wave: probably represents "after depolarizations" in the ventricles. PR interval: time interval from onset of atrial depolarization (P wave) to onset of ventricular depolarization (QRS complex). QRS interval: duration of ventricular muscle depolarization. QT interval: duration of ventricular depolarization and repolarization.

and some personal information about the subjects. The dataset is divided into two groups as Normal and Abnormal (arrhythmia). The ECG signals used as system inputs include 5 parameters: QRS duration, PR interval, QT interval, T interval and P interval. The personal information is also available for system inputs, *i.e.* age, height, weight and sex. There are 245 cases in the Normal group and 207 cases in the Abnormal (arrhythmia) group.

Succeeding data of extracted parameters, denoted by t_n, are converted to the deviation from their mean value and scaled by a factor m to normalize their values in the range of -1 to +1 by equation:

$$x_n = \frac{t_n - \bar{t}}{m}, \qquad n = 1 \sim N \tag{1}$$

where n is the number of input data samples, corresponding to training and test sets respectively, and N is the maximum number of input data samples. $\bar{t}$ is the average of t_n, m is the standard deviation of input data for training and test sets respectively.

The whole data set is divided into training and test subsets. The first subset, containing 360 ECG recordings of the collected data, is used for training; the second subset, containing 92 ECG recordings, is used for test.

2.2 Neural Network Classifier

Since the detailed methodological aspects of BP neural networks can be found in related literatures [16], [17], [18], here we mainly focus on the description

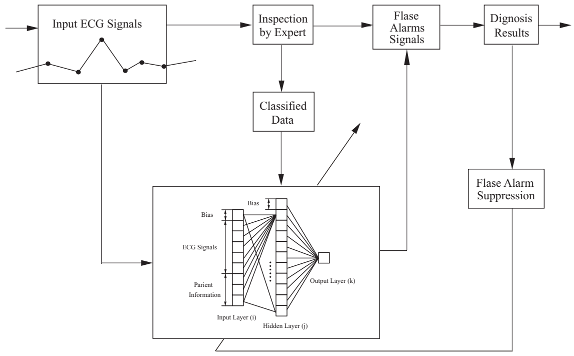

Figure 3: Block Diagram of ANN-based Arrhythmia Diagnostic System with ECG Signals and Configuration of Neural Network Classifier with input layer ($i = 1, 2, \cdots, I$), hidden layer ($j = 1, 2, \cdots, J$) and output layer ($k = 1, 2, \cdots, K$)

of the Neural Network Classifier. The general configuration of Neural Network Classifier considered here may be represented schematically as shown in Fig. 3, where $i = 1, 2, \cdots, I$ is the number of input nodes, $j = 1, 2, \cdots, J$ is the hidden nodes, $k = 1, 2, \cdots, K$ is the class labels of the output node.

Assume we have a training set $\mathcal{D}$, consisting of N input-output pairs:

$$\mathcal{D} = [(X^n, y^n) \mid n = 1, 2, \cdots N] \tag{2}$$

where X is an input vector consisting of I elements and y is the corresponding class label consisting of K classes. The objective is to use neural network model to classify the input-output relation ($y = k \mid X$). In our research, the class label is the binary form y = 1, 0, corresponding to Normal or Abnormal (arrhythmia) respectively.

The problems described above are considered to solve by the following candidate solution. An alternative target for our task is to use logistic regression model based on Bayesian method to estimate the class probability for the given input by:

$$\mathcal{P}(y = k \mid X), \qquad k = 1, 0 \tag{3}$$

Supposed that the outputs for the summation operation and sigmoidal activation function in the hidden and output neurons, denoted by S_j, S_k, respectively, can be written as follows:

Input layer

$$S_j = \tanh(\sum_i \omega_{ji} X_i + \omega_{j0}) \tag{4}$$

Output layer

$$S_k = \sum_k \omega_{kj} S_j + \omega_{k0} \tag{5}$$

where tanh is the tangent hyperbolic function, a conventional sigmoid function. ω_{ji} denotes the weight matrix in the input layer and ω_{kj} the weight matrix in the output layer. To ensure that the outputs can be interpreted as probabilities, a logistic regression model is used to model the risk (or probabilities) of the disease development for arrhythmia. Let $P\ (y = k \mid X)$ be the probabilities of the event $y = 1$, given the input vector X. This is modeled as a function of network output y by

$$\mathcal{P}(y = k \mid X) = \frac{1}{1 + \exp\left(-S_k\right)} \tag{6}$$

The logistic regression model is simply a non-linear transformation of the linear regression. The "logistic" distribution is an S-shaped distribution function which is similar to the standard-normal distribution (which results in a probit regression model) but easier to work with in most applications. The logit distribution constrains the estimated probabilities to lie between 0 and 1.

The final model investigated here is a multilayered perceptron neural network adopting the BP algorithm. The network is optimized using a log-likelihood cost function, given by

$$\mathcal{C}(\omega) = -\frac{1}{k} \sum_k \sum_i y_i(k) \ln[\mathcal{P}(y = k \mid X)] \tag{7}$$

where $\omega = [\ \omega_{ji}, \omega_{kj}\ ]$ is the vector of the network weights.

To minimize the cost function between the actual and desired outputs of the network, the BP algorithm passes information from the output neuron backwards to all hidden units to form error terms which are used by the learning procedure to update the weights of the multilayered network. A three layer fully connected network with 10 hidden units and a single output unit is used.

2.3 Evaluation methods

To examine how well the Neural Network Classifier close to the target classes and evaluate its performance in reorganization of arrhythmia with ECG signals by our method, the misclassification rate is used to verify that the Neural Network Classifier acquires underlying dynamics of the system from data, and the False Rate is used to quantify the system performance and evaluate the accuracy of the model.

The misclassification rate is defined as:

$$\mathcal{MR} = \frac{\sum i}{N}, \qquad i \subset \{K \neq O\} \tag{8}$$

where $\mathcal{MR}$ is the false rate. K and O is the target and output classification labels, respectively. N is the number of examples and classes.

Let

FP: false positive;
FN: false negative;
TP: true positive;

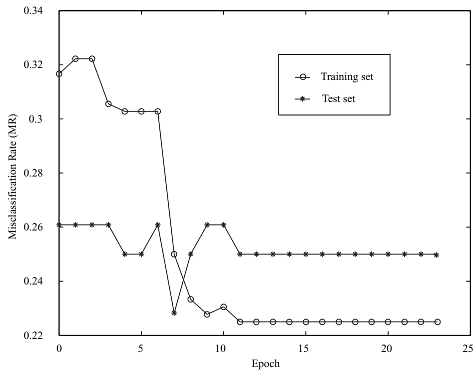

Figure 4: Classification Error. $\mathcal{MR}$: Misclassification Rate.

TN: true negative; and
UN: uncertain class, if applicable.

Then, the False Rate is defined as:

$$\mathcal{FR} = \frac{FP}{FP + TP + UN} \tag{9}$$

$$\mathcal{FR} = \frac{FN}{FN + TN + UN} \tag{10}$$

$$\mathcal{FR} = \frac{FP + FN}{FP + TP + FN + TN + UN} \tag{11}$$

where $\mathcal{FR}$ is the False Rate for Positive, Negative and overall, respectively.

3 Empirical Results

3.1 Classification Results

To verify that the Neural Network Classifier acquires underlying dynamics of the system from data, the misclassification rate for both training and test phases are estimated.

Fig. 4 shows the misclassification rate in training and test phases respectively. Before training, synaptic connections ω_{ij} have random values and the networks yield large classification error. However, as the training processes, the misclassification rate decreases rapidly from initial values and remain almost constant after several epochs of iterations, where one iteration corresponds to one round of training using all training data. After training, the misclassification rate reduces to 0.225. The corresponding misclassification rate for test is also small enough to meet the requirement of the study.

The classification results for training and test phases are shown in Fig. 5 (a) and (b), respectively. The output probability from the Neural Network

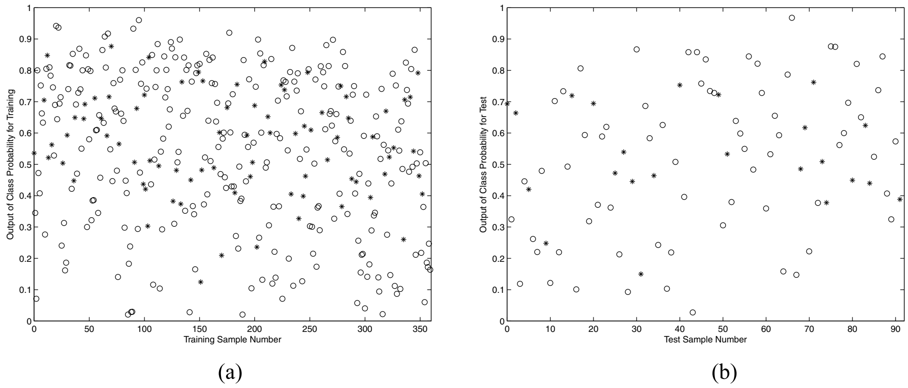

Figure 5: Dispersal Plot of Classification Results for Training (a) and Test (b), respectively. The true and false classification results are marked by o and * respectively.

Classifier will be rounded to 1 or 0 as classification results. Fig. 5 also shows the distribution condition for true and false classification in both phases if the output probability from the Neural Network Classifier is rounded to 1 or 0 by 0.5 as classification results.

3.2 Evaluation for Classification Results

Table 1 shows arrhythmia diagnosis results. The False Rate for training and test phases is 22.50% and 25.00%, respectively. Therefore, more than 75% prediction accuracy is obtained in both phases.

Fig. 6 shows the histogram of misclassification distribution for training and test phases, respectively. From Fig. 8, we found that the distribution of misclassification mainly concentrates on the area between 0.4 and 0.7. Considering that the probability of network outputs is rounded for classification results, we can conclude that the classification results around 0.5 have higher risk for misclassification. Such condition also hints that if the misclassification between 0.4 and 0.7 could be ignored, the False Rate will be suppressed notably.

3.3 False Alarm Suppression

A well-known difficulty with medical monitoring devices that raise alarms to alert the physician of patient problems is the "cry-wolf" dilemma: these devices follow a conservative strategy of raising false alarms rather than risking a situation arising where a patient requires attention but no alarm is raised, but physician may end up ignoring all of the alarms if so many of them are false.

The dual threshold method is used here to solve above problems and suppress False Alarm Signals. The basic idea is to adopt a dual threshold, *i.e.*

Table 1: Arrhythmia Diagnosis Results for ANN-based Diagnostic System with ECG Signals. $\mathcal{FR}$: False Rate.

	Case of Classes	Predicted Results Normal	Abnormal	$\mathcal{FR}$ (%)
Training				
Normal	195	167	28	14.36
Abnormal	165	53	112	32.12
Total	360	220	140	22.50
Test				
Normal	50	39	11	22.00
Abnormal	42	12	30	28.57
Total	92	51	41	25.00

Table 2: Arrhythmia Diagnosis Results for ANN-based Diagnostic System with ECG Signals. $\mathcal{FR}$: False Rate.

	Case of Classes	Predicted Results Normal	Abnormal	Uncertain	$\mathcal{FR}$ (%)
Training					
Normal	195	94	9	92	4.62
Abnormal	165	22	86	57	13.33
Total	360	116	95	149	8.61
Test					
Normal	50	21	5	24	10.00
Abnormal	42	3	24	15	7.14
Total	92	24	29	39	8.70

Threshold 1 and Threshold 2.

IF:

Output probability > Threshold 1

THEN:

y = Classification 1 (Normal)

IF:

Output probability < Threshold 2

THEN:

y = Classification 2 (Abnormal)

In addition above two conditions, an uncertainty criterion is introduced for the high risk questionable classification outputs.

IF:

Output probability $\leq$ Threshold 1, and

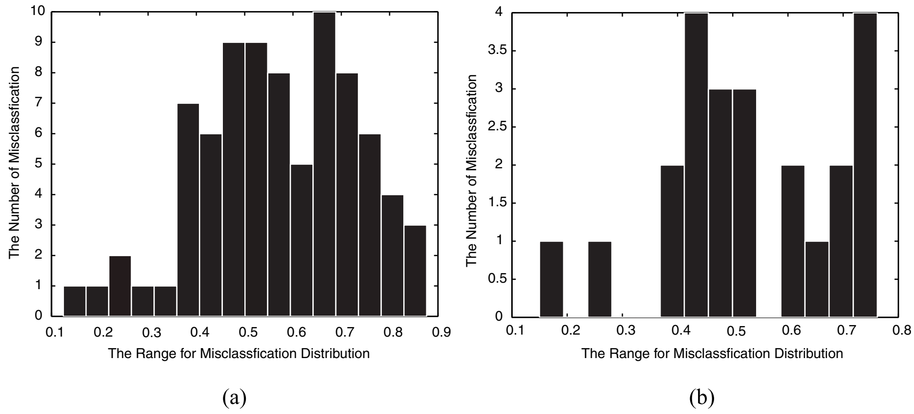

Figure 6: Histogram of Misclassification Distribution for Training (a) and Test (b), respectively.

Output probability $\geq$ Threshold 2

THEN:

y = uncertain class

The two thresholds, *i.e.* 0.4 and 0.7, divide the outputs into three ranges unevenly instead of two ranges with low range negative, middle range questionable and high range positive. When the output is higher than 0.7, it is rounded to one and classified to positive class. The higher this threshold, the more conservative to raise false alarm, which means higher prediction accuracy and bigger uncertainty. When the output is lower than 0.4, it is set to zero. The outputs will be classified as uncertain class between 0.4 and 0.7.

Table 2 shows diagnosis results using dual threshold 0.4 and 0.7. The False Rate for training and test phases is reduced to as small as 8.61% and 8.70%, respectively. Therefore, more than 90% prediction accuracy could be obtained in both phases. However, we got an uncertain class between 0.4 and 0.7 with medium risk for arrhythmia, which should be reviewed by physicians for determination.

When the thresholds are varied, the different classification accuracy and uncertainty can be obtained. This property can be used to determine different diagnosis strategy. When the suitable thresholds are chosen to build the Neural Network Classifier, the optimum False Rate and uncertainty are expected to get from such model. The dual threshold method presented here could be applied to suppress False Alarm Signals and generate general rules about the suppression of alarms. These rules can also be reviewed by a physician and potentially incorporated into the training phase of the system.

4 Conclusion

In this paper, an ANN-based diagnostic system for arrhythmia with ECG signals using Neural Network Classifier with Bayesian framework is present. The prediction performance of the system is evaluated by the False Rate. The logistic regression model and back propagation algorithm is used to build our Neural Network Classifier. The Neural Network Classifier acquires arrhythmia properties from underlying dynamics of the system with ECG signals, even when the dataset includes the uncompleted information, such as missing feature values and unclassified classes. Artificial Neural Network (ANN) is potentially useful to generate a pattern recognition model based on given {input, output} sets to classify future input sets for arrhythmia diagnosis. The capability of uncertainty management with the dual threshold method could be used to determine diagnosis strategy and suppress false alarm signals.

The future works can focus on the real-world data, which are in progress in our project. The implementation to design low-cost, high-performance, simple to use, and portable equipment with ECG signals, and offer a combination of diagnostic features, seems to be a goal that is highly worthwhile. Hopefully, the system can be further developed and fine-tuned for practical application.

5 Acknowledgements

This research is supported by Enterprise Ireland, Advanced Technologies Research Programme, Strategic Research Project number *ARTP 2002 INF 210A*.

References

[1] J. L. Willems and E. Lesaffre. Comparison of multigroup logisitic and linear discriminant ECG and VCG classification. **J. Electrocardiol.** 1987; 20: 83–92.

[2] J. L. Talmon, Pattern Recognition of the ECG. Berlin, Germany: Akademisch Proefscrift, 1983.

[3] Y. H. Hu, S. Palreddy, and W. J. Tompkins. A patient-adaptable ECG beat classifier using a mixture of experts approach. **IEEE Trans. Biomed. Eng.** 1997; 44: 891–900.

[4] A. D. Coast, R. M. Stern, G. G. Cano, and S. A. Briller. An approach to cardiac arrhythmia analysis using hidden Markov models. **IEEE Trans. Biomed. Eng.** 1990; 37: 826 - 835.

[5] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo. Clustering ECG complexes using hermite functions and self-organizing maps. **IEEE Trans. Biomed. Eng.** 2000; 47: 838–848.

[6] R. Silipo and C. Marchesi, Artificail Neural Networks for Automatic ECG Analysis. **IEEE Trans. Signal Processing** 1998; 46 (5): 1417–1425.

[7] H. A. Guvenir, B. Acar, G. Demiroz, A. Cekin. A Supervised Machine Learning Algorithm for Arrhythima Analysis. **In: Proc. Computers in Cardiology Conference**. Lund, Sweden. 1997; 24: 433–436.

[8] C. L. Blake, C. J. Merz. UCI Repository of machine learning databases. [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA: University of California, Department of Information and Computer Science. 1998.

[9] K. I. Minami, H. Nakajima, and T. Toyoshima. Real-time discrimination of ventricular tachycardia with Fourier-transform neural network. **IEEE Trans. Biomed. Eng.** 1999; 46: 179–185.

[10] P. Frasconi, M. Gori, M. Maggini, and G. Soda. United integration of explicit knowledge and learning by example in recurrent networks. **IEEE Trans. Knowl. Data Eng.** 1995; 7: 340–346.

[11] Yang Wang, Yi-Sheng Zhu, Nitish V. Thakor, and Yu-Hong Xu. A Short-Time Multifractal Approach for Arrhythmia Detection Based on Fuzzy Neural Network. **IEEE Trans Biomed. Eng.** 2001; 48 (9): 989–995.

[12] D. MacKay. The Evidence Framework Applied to Classification Networks. **Neural Computation** 1992; 4 (5): 720–736.

[13] D. MacKay. A practical Bayesian framework for backpropagation networks. **Neural Computation** 1992; 4: 448–472.

[14] M. Hintz-Madsen, L. K. Hansen, J. Larsen, K. Drzewiecki. A Probabilistic Neural Network Framework for Detection of Malignant Melanoma. **Artificial Neural Networks in Cancer Diagnosis, Prognosis and Patient Management**. 141-183, CRC Press, 2001.

[15] Pierre Baldi and Soren Brunak, **Bioinformatics: The Machine Learning Approach**, 2nd Ed., ISBN: 0-262-02506-X, A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England, 2001.

[16] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representation by error propagation. **In: Rumelhart, D. E., McClelland, J.L., and the PDP research group (Eds), parallel distributed processing**, MIT press, Cambridge, MA 318–362, 1996.

[17] D. Gao, Y. Kinouchi, K. Ito, X. Zhao, Neural Networks for Event Extraction from Time Series: A Back Propagation Algorithm Approach. **Future Generation Computer Systems (FGCS)**. (In press)

[18] D. Gao, Y. Kinouchi, K. Ito, X. Zhao. Time Series Identifying and Modeling with Neural Networks. **In: Proc. IEEE Int. Joint Conference on Neural Networks**, Portland, USA. 2003; 2454–59.