International Journal of Fuzzy Logic and Intelligent Systems, vol. 6, no. 3, September 2006 pp. 179-183

A Comparison of the Performance of Classification for Biomedical Signal using Neural Networks

Man-Sun Kim* and Sang-Yong Lee**

***Dept. of Computer Engineering, Kongju National University**
****Division of Computer Science & Engineering, Kongju National University**

Abstract

ECG consists of various waveforms of electric signals of heat. Datamining can be used for analyzing and classifying the waveforms. Conventional studies classifying electrocardiogram have problems like extraction of distorted characteristics, overfitting, etc.
This study classifies electrocardiograms by using BP algorithm and SVM to solve the problems. As results, this study finds that SVM provides an effective prohibition of overfitting in neural networks and guarantees a sole global solution, showing excellence in generalization performance.

Key Words : ECG, Datamining, BP algorithm, SVM, Neural Networks

1. Introduction

Many methods are under research in order to analyze various biomedical signals including electrocardiogram (ECG), elec-troencephalogram (EEG) and electromyograph (EMG)[1,2,3]. Among them, electrocardiographic data are composed of diverse wave forms of the electric signals from the heart. In classifying and analyzing such wave forms, data mining may be utilized. Pattern classification, one of data mining techniques, is largely divided into future extraction, which automatically finds specific information from given data, and classification, which classifies given data into two more groups and identifies their characteristics. To solve pattern classification problems, traditional statistical pattern classification and syntactical pattern classification have been commonly used, but recently the use of neural networks is getting more extensive. Backpropagation (BP) algorithm has been applied most frequently in solving pattern classification problems using neural networks, but it has delicate problems such as the delay of convergence and the selection of kernel functions affecting approximation and convergence rate.

This study applied parameters, which were extracted from the characteristics of electrocardiographic patterns in order to classify the pattern of electrocardiographic signals, to BP algorithm and support vector machine (SVM)[4] and compared the performance of the two models in pattern classification.

2. Pattern Classification for Biomedical Signal

2.1 Biomedical Signal

Electrocardiogram (ECG) is a very important means for non-invasive diagnosis. It is one of bio-potential signals, with amplitude of several mV and frequency less than 250Hz. Research on ECG system design and signal processing began in the early 1960s in the U.S.A. Since then, there have been researches on designing hardware such as multi-channel ECG for automatic diagnosis, Holter ECG that monitors patients with cardiac disorder for 24 hours a day, stress ECG that diagnose cardiac disorders under exercise stress, and developing accurate algorithms. In Korea, research on ECG system design and signal processing began in the early 1980s. With 10 years' accumulation of fundamental technologies, the development of ECG was fully activated from the early 1990s, and currently 12-channel diagnosing ECG, Holter ECG, stress ECG, patient monitoring equipment and other heart-related diagnosing machines are being studied. Despite a great deal of research efforts, misdiagnosis is still frequent in relation to myocardial ischemia and myocardial infarction. Diagnosis of such diseases is based on the up and down of the level or the gradient of ST segment of ECG signal.

Figure 1 below shows the location of ST segment. Because ST segment has a frequency band below 1Hz, it shares the same frequency band with the baseline variation noise of low frequency and muscle artifact that exists in every frequency band. Thus inaccurate removal of noises causes signal distortion, which in turn causes misdiagnosis. Currently available pre-processing methods to remove baseline variation noise are spline interpolation technique, FIR filtering, adaptive filtering, neural network, wavelet transform technique, etc. These techniques minimize signal distortion and remove baseline variation noise. Among the methods, wavelet transform processes signals in multiple resolution, and transformed signals have high resolution in the domains of time and frequency. Thus the method is suggested as an advantageous method for analyzing non-stationary signals. Because the entire process of wavelet transform is performed through mother wavelet, even if the same wavelet transform method is used, the wrong selection of the generating function may bring about the severe

Manuscript received Jun. 17, 2006; revised Sep. 13, 2006.
Corresponding Author : Sang-Yong Lee

distortion of signals.

There are several methods of classifying ECG data [1,2,3], but it is necessary to compare and analyze by applying BP algorithm and SVM of high performance in reducing misclassification rate and shortening learning time.

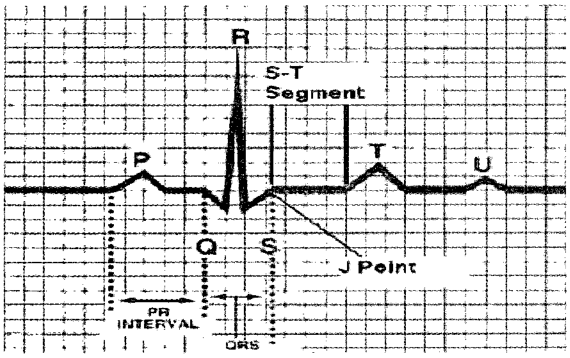


Figure 1. ST segment of ECG signal

2.2 Existing researches on classifying ECG signals

Existing researches on classifying ECG signals are divided into those based on neural networks and those based on signal processing. One of neural networks based researches is HEDEN's research using BP algorithm[5], in which the initial learning rate (μ) was set at 0.5 and training was performed so that the learning rate, a constant adjusting the variation of weight, gradually decreased. Equation below is the equation of gradually changing learning rate.

$$\mu = k\mu (k = 0.998)$$

The performance was verified using k-fold cross test, and k was set at 3 to prevent over-fitting. Because of a serious error of data in the measuring program, however, the network was not expressed appropriately.

Silipo's research compared the performance of statistical classification methods with that of neural network based classification methods[6]. Ambulatory ECG (AECG) was analyzed and European Database was used. Among statistical classification methods, the single linkage method was used and the final error was measured in RMS. It was designed so that training was stopped when RMS went below 0.6. The result is not considered accurate, however, because the error tolerance limit is too large, which should have been at least 0.001.

As one of signal processing based researches, Sternickel's technology of automatic detection of ECG time series data pattern was mentioned above[7]. It expressed multi-resolution using wavelet variation, and tested using a Holter ECG recorder to improve the reliability of time series data. In P-wave detection, we used coiflet6 as wavelet generating function and Haar as a generating function to find QRS Complex. Because the selection of wavelet generating function is an extremely important factor, a wrong selection of the function may distort the diagnosing parameters of ECG.

2.3 Extracting characteristics

In order to decide wavelet generating function that can remove baseline by minimizing the distortion of raw signals, this study removed baseline by applying various wavelet generating functions.

ST segment was set at 160ms from 60ms after R-peak if the RR interval is larger than 600ms or from 40ms if not. In order to extract the characteristics of ST segment, we took as candidate parameters representing the characteristics of ST segment ST0 [amplitude on the starting point of ST segment (R+60 or R+40 ms)], ST60, ST80 [amplitude at R+140 or R+120 ms], the gradient of ST segment, and the area of ST segment (the area surrounded by ST segment and isoelectric level in the interval of ST segment).

2.4 Classification

2.4.1 BP algorithm

BP algorithm is a multi-layered learning algorithm used in feedforward neural networks and it utilizes generalized Delta rule which is supervised learning. That is, it requires input data and desired output data for learning. Neural networks learning using BP algorithm is carried out in three steps. The first step is to enter learning input patterns into the neural networks and produce output. The second step is to obtain the difference between the output and the target. The third is to propagate the difference backward and change the connection strength of the output layer and that of the hidden layer. Although learning may take a long time, once learning has been completed, results come out quickly in the step of application.

2.4.2 SVM

In SVM, there is a training group S that is divided into two classes, and hyperplane is determined, which divides training patterns into the two classes. Here, hyperplane means a cutting plane that divides each group. Input patterns that determine the hyperplane are called support vectors. If it is possible to divide a pattern group, the hyperplane maximizes the margin between the plane and each support vector and all support vectors are positioned at the same least distance from the hyperplane. Actual pattern groups, however, are rarely linear. Thus, the two classes can obtain the optimal solution from a pattern group equation that divide non-linearly. The optimal solution has tradeoff between maximizing the margin (the distance among support vectors in each class) and minimizing the number of errors, and it is adjusted by normalized parameters.

3. Experiment and Result

3.1 Experiment and data

Experiment was implemented in Matlab 6.5 on a PC of Windows XP, 2.5GHz PⅣ and 256M RAM. This study used European ST-T database for experiment, and selected training data and test data as shown in Table 1 and Table 2 respectively. Training data and test data in the 1st to 5th columns are ST0, ST60, ST80, the slope of ST segments and the

Table 1 : Training Data

ST0	ST60	ST80	slope	area	normal/ abnormal
−0.1135	0.0040	2.4893	0.0385	0.0503	0
−0.4756	0.0138	3.9045	0.0393	0.0271	0
−0.0946	0.0029	1.4481	0.0458	0.0270	0
−0.0324	0.0034	3.9385	0.0515	0.0798	0
−0.5512	0.0172	5.5546	0.0641	0.0626	0
0.0116	0.0019	3.9832	0.0753	0.0782	0
0.0467	−0.0031	3.2753	0.0802	0.0833	0
0.0940	−0.0040	3.5395	0.0993	0.1002	0
−0.5187	0.0144	4.1742	0.0604	0.0336	0
−0.0096	−0.0006	0.7599	0.0226	0.0412	0
−0.0700	0.0031	1.6434	0.0055	−0.0056	0
−0.0660	0.0023	1.5885	0.0074	0.0082	0
−0.2038	0.0060	2.3681	0.0293	0.0211	0
−0.2346	0.0068	1.8981	−0.0271	−0.0022	0
−0.0628	0.0022	1.5759	−0.0052	0.0262	0
−0.0823	0.0025	1.3705	0.0006	0.0328	0
0.0489	−0.0009	1.6991	0.0265	−0.0032	0
0.0243	0.0004	1.9748	0.0566	0.0617	0
0.0727	−0.0023	1.7114	0.0093	0.0270	0
0.0299	−0.0029	1.6474	0.0578	0.0559	0
−0.0381	0.0033	3.4012	0.0632	0.0916	0
−0.4954	0.0154	5.0721	0.0684	0.0702	0
−0.2272	0.0080	4.2015	0.0571	0.0837	0
−0.3592	0.0112	4.6839	0.0691	0.0643	0
0.0363	0.0007	2.8697	0.0487	0.0814	0
0.0560	−0.0006	2.6385	0.0345	0.0508	0
−0.4880	0.0136	2.9854	0.0012	0.0119	0
−0.5596	0.0151	3.4249	−0.0142	0.0097	0
0.0485	−0.0013	2.4668	0.0663	0.0675	0
−0.5500	0.0166	5.0943	0.0583	0.0633	0
−0.0884	0.0053	3.9729	0.0698	0.0822	0
−0.5254	0.0149	5.1860	0.1368	0.0949	0
0.0001	0.0001	0.9578	0.0122	0.0223	0
−0.1113	0.0060	3.7360	0.0755	0.0909	0
0.0583	−0.0028	3.8606	0.0932	0.1189	0
0.0834	−0.0026	3.7370	0.1414	0.1477	0
−0.0261	0.0037	3.9400	0.0978	0.0922	0
0.0646	0.0001	4.2586	0.0895	0.1031	0
0.0876	−0.0021	3.0816	0.0737	0.0730	0
−0.0233	0.0020	1.4452	0.0274	0.0305	0

Table 2 : Test Data

ST0	ST60	ST80	slope	area
0.1330	−0.0008	6.6396	0.1716	0.1659
0.0688	0.0017	7.4188	0.1679	0.1610
0.1722	−0.0013	7.5799	0.1850	0.1901
−0.3290	0.0128	8.2698	0.2126	0.1470
0.1702	−0.0033	4.6080	0.1286	0.1256
0.1879	−0.0066	6.3977	0.1246	0.1866
0.2113	−0.0048	7.0558	0.1913	0.2211
0.1762	−0.0064	5.3308	0.1798	0.1268
0.1062	0.0005	7.6900	0.1535	0.1800
0.0365	0.0019	6.2079	0.1375	0.1337
0.1484	−0.0039	5.0433	0.1432	0.1621
0.0786	0.0010	6.0625	0.1504	0.1785
0.1559	−0.0038	5.8166	0.1666	0.1707
0.1271	0.0004	6.7219	0.1731	0.1664
0.1855	−0.0075	4.2403	0.1281	0.1074
0.1509	−0.0039	5.5711	0.1364	0.1425
0.0890	−0.0044	3.9069	0.1249	0.1250
0.1213	−0.0058	3.9128	0.1217	0.0889
−0.0129	0.0038	6.8925	0.1479	0.1825
0.1849	−0.0024	6.2268	0.1588	0.1255
−0.3306	0.0120	6.3880	0.1558	0.1317
0.1668	−0.0052	2.9486	0.0748	0.0000
0.1473	−0.0045	4.6849	0.1511	0.1309
−0.3090	0.0102	5.4607	0.1375	0.0839
0.0556	0.0020	6.6297	0.1628	0.1680
0.1924	−0.0052	6.1906	0.1733	0.1692
−0.2612	0.0106	6.9414	0.1955	0.1905
0.1925	−0.0028	6.5225	0.1639	0.1679
0.1334	−0.0024	7.0215	0.1868	0.1784
0.1284	−0.0030	3.6770	0.1174	0.1065
0.1302	−0.0048	3.5794	0.0855	0.0862
0.1864	−0.0065	4.6877	0.1037	0.1365
0.0277	0.0010	5.8279	0.1330	0.1409
0.1097	−0.0026	4.8000	0.1294	0.1082
0.1466	−0.0065	5.0105	0.1434	0.1253
0.1316	−0.0055	4.8869	0.1517	0.1490
0.1254	0.0040	3.1376	−0.0063	−0.0586
−0.4069	0.0124	5.7989	0.1458	0.1509
−0.0587	0.0035	5.6214	0.1269	0.1328
−0.2794	0.0082	5.5125	0.1707	0.1623

area of ST segments respectively, and training data in the 6th column indicates if the result of the data is normal (0) or abnormal (1). In addition, the test data do not include data of the 6th column.

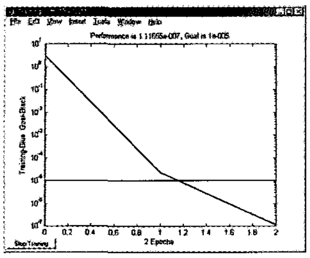
3.2 Extracting characteristics

In order to decide wavelet generating function that can remove baseline by minimizing the distortion of raw signals, this study removed baseline by applying various wavelet generating functions. According to the result of the experiment, the best wavelet generating functions were db8 (diff.: 27.12), coif5 (diff: 25.32) and sym7 (diff: 25.13), and it was found that diff (meanSNR - meanRSE) less than 23 is unusable because it distorts even the diagnosing parameters of ECG.

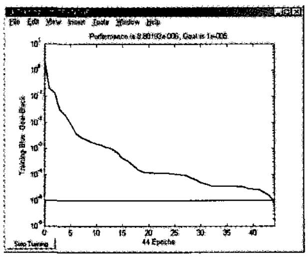
3.3 BP algorithm

(1) BP algorithm training methods and the result of convergence

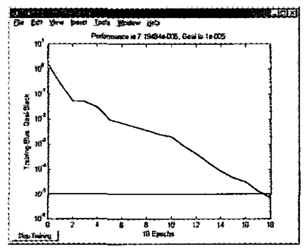
Comparing speed and memory of BP algorithm training methods provided by MATLAB6.5, this study selected the best four methods and applied them to ECG data. Here, traincgf is Fletcher Powell Congugate Gradient, trainlm is Levenberg Marquaedt, trainbfg is BFGS Quasi Newton, and trainrp is Resilient Backpropagation. Figure 2 shows the result of convergence. In the figure, (b) shows fine convergence at 10^{-5} as (a) but (b) and (d) are under-fitting and over-fitting respectively, so they are not fit for problems to which they are to be applied. According to the result of experiment, the application of traincgf was found to produce the most appropriate convergence among BP algorithm training methods.



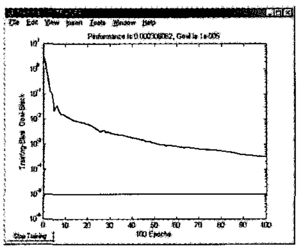
(a) traincgf



(b) trainlm



(c) trainbfg



(d) trainrp

Figure 2. the result of convergence for four methods

(2) The effect of variation in the number of training times on MSE

With the increase of the number of training times, learning takes a longer time. Thus it is necessary to set the number of training times so that learning stops when error does not decrease any more. In Figure 3, MSE decreases exponentially with the increase of the number of training times and the rate of MSE decrease becomes slow after 10,000 times of training.

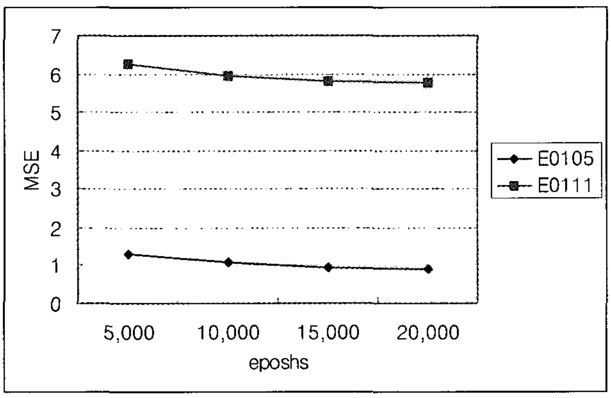


Figure 3. the number of training times and MSE

(3) The effect of variation in learning rate on MSE

Learning rate is usually between 0.01 ~ 1. Because the result of learning is different according to the characteristic of problems to which the neural network is applied, however, it is necessary to perform an experiment to find the optimal learning rate for the characteristic of ECG signals. Thus, this study carried out an experiment by varying learning rate from 0.1 to 0.7 by increasing by 0.2. According to the result, the distribution of MSE was smallest at around 0.1. Thus we again tested with learning rate ranging from 0.01 to 0.1 by increasing by 0.01. Table 3 below is the results of experimenting on data e0105 and e0111 with different learning rates and 5,000, 10,000 and 15,000 training times.

3.4 SVM

The kernel function of SVM classifier used a polynomial, the degree of which was 3. The value of C, which plays the role of finding a compromising point between the margin and classification error and is a error penalty variable for inseparable data, was set at 1. The experiment classified ECG signals, the characteristic of which was extracted for two methods as in Table 4, and compared the performance of the methods.

Table 3. learning rate and MSE

learning rate	E0105 epochs			E0111 epochs		
	5,000	10,000	15,000	5,000	10,000	15,000
0.01	1.69	1.68	1.68	8.69	8.55	8.24
0.02	1.68	1.66	1.60	8.55	7.62	6.42
0.03	1.65	1.49	1.17	8.10	6.20	4.93
0.04	1.57	1.21	**0.89**	7.37	6.09	4.50
0.05	1.47	**0.97**	0.90	6.65	4.89	**4.30**
0.06	1.35	1.00	0.90	6.18	4.84	**4.30**
0.07	1.27	1.00	0.91	5.98	**4.67**	4.33
0.08	1.23	1.10	0.92	**5.88**	4.71	4.36
0.09	1.22	1.04	0.95	6.02	4.80	4.52
0.10	1.21	1.06	0.97	6.28	4.98	4.56
0.15	3.25	1.41	1.18	6.79	6.91	6.80
0.20	7.45	6.03	5.79	7.49	7.09	6.81
0.25	11.44	9.24	9.12	10.09	7.78	7.04
0.30	19.27	11.74	10.99	13.00	13.0	14.26

Table 4. comparison of classification performance

Data \ Classifier	BP algorithm	SVM
e0103	66%	55%
e0105	35%	75%
e0111	53%	76%

BP algorithm was slightly better than SVM for e0103 but SVM might not learn parameters for global optimum and support vectors obtained through learning might not sufficient for perfect classification of unknown test data. For 0105, BP algorithm and SVM showed classification rate of 35% and 75% respectively, and for e0111 53% and 76% respectively. This suggests that BP algorithm is inferior in generalization as a result of over-fitting. As SVM effectively prevents over-fitting in neural networks and guarantees a single global solution, it appears to be superior in generalization.

4. Conclusions

In order to classify the pattern of ECG signals, this study used parameters extracted from the characteristics of ECG patterns as input for machine learning. In an experiment with BP algorithm, the use of traincgf resulted in appropriate convergence into a desired point. In addition, the final error, namely, MSE decreased exponentially with the increase of the number of training times and the rate of MSE decrease became slow when the number of training times exceeded 10,000. As for the effects of learning rate and MSE, the optimal learning rate was different according to data but it was roughly between 0.04 and 0.1.

According to the result of comparing the classification performance of SVM and BP algorithm, SVM was superior, and particularly simple and efficient in binary classification. To reach classification rate over 90%, however, it required a larger volume of training data. In order to overcome the limitations of SVM in speed and memory, we plan research on multiple SVM rather than single SVM.

References

[1] Z. Dokur and T.Olmez, "ECG beat classification by a novel hybrid neural network," Computer Methods and Programs in Biomedicine, Volume 66, Issues 2-3, pp. 167-181, 2001.

[2] P. M. Rautaharju, S. H. Zhou, et al., "Comparability of 12-lead ECGs derived from EASI leads with standard 12-lead ECGS in the classification of acute myocardial ischemia and old myocardial infarction," Journal of Electrocardiology, Volume 35, Issue 4, Part 2, pp. 35-39, 2002.

[3] U. Rajendra Acharya, P. Subbanna Bhat, et al., "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," Pattern Recognition", Volume 36, Issue 1, pp. 61-68, 2003.

[4] www.support-vector.ws/html/downloads.html.

[5] B. Heden, "Agreement Between Artificial Neural Networks and Experienced Electro-cardiographer on Electrocardiographic Diagnosis of Healed Myocardial Infarction," JACC, Vol.28, No.4, pp. 1012-1016, 1996.

[6] R. Silipo, M. Goru, et. al., "Classification of Arrhythmic Events in Ambulatory Electrocardiogram, Using Artificial Neural Networks", Computers and Biomedical research Vol.28, pp. 305-318, 1995.

[7] K. Sternickel, "Automatic patttern recognition in ECG time series", Computer Methods and Programs in Biomedicine, Vol.68, pp. 109-115, 2002.

Man-Sun Kim

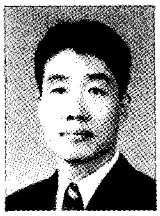
2000 : Dept. of Computer Engineering, Hongik University (B.S)

2002 : Dept. of Computer Engineering, Kongju University (M.S)

2005 : Dept. of Computer Engineering, Kongju University (Ph.D)

Research Interests : AI, NN, Datamining, Pattern Recognition

E-mail : mansun@kongju.ac.kr



Sang-Yong Lee

1984 : Dept. of Computer Science, Chungang University (B.S)

1988 : Dept. of System Science, Tokyo Institute of Technology (M.S)

1988～1989 : Researcher of NEC Research Center in Japan

1993 : Dept. of Computer sciences, Chungang University (Ph.D)

1996～1997 : Visiting Professor of University of Central Florida in U.S.A.

1993～ : Professor of Division of Computer Science & Engineering, Kongju National University

Research Interests : Artificial Intelligent, Agent System, Ubiquitous Computing, Context-Awareness

E-mai : sylee@kongju.ac.kr