Design of a Fast Learning Classifier for Sleep Apnea Database based on Fuzzy SVM

SooYong Lee¹, Erdenebayar Urtnasan²,³, and Kyoung-Joung Lee³
¹Division of Liberal Education, Yonsei University, Wonju, Korea
²Department of Medical Engineering, Huree University, Ulaanbaatar, Mongolia
³Department of Biomedical Engineering, Yonsei University, Wonju, Korea

Abstract
In this paper, we compared the performance of support vector machine (SVM) and fuzzy SVM (FSVM) for reduction of learning time when classifying large-scale time series data into two classes. The fast learning time of the pattern classifier for large time series data is very useful in decision support systems. Considering the high interest in healthcare, including big data analysis, it is necessary to design a pattern classifier with a fast learning capability. We used large-scale time series data of 32 patients with sleep apnea (SA) for this study. The experiment was conducted by extending the parameter $n$, of the fuzzy membership function of FSVM, from 1 to 500. The result shows that the shortest learning time of FSVM is 3 s for radial base function (RBF), 17 s for a polynomial, and 35 s for a linear kernel, where the parameter $n$ of the fuzzy membership function is $n = 2$, $n = 433$, and $n = 4$, respectively. The maximum classification hit rate of FSVM is 93.23%, and the learning time is significantly faster than conventional SVM. Therefore, FSVM can be used as a good classifier for the large-scale time series SA database.

Keywords: FSVM, SVM, Fuzzy membership function, Classifier, Time series data, Learning time, Sleep apnea

1. Introduction
Support vector machine (SVM) is a high-performance classification technology that has attracted much attention in various fields. SVM includes mapping of learning patterns into a high-dimensional feature space and finding an optimal hyperplane that maximizes the margin between two classes in this space [1]. SVM has been shown to provide higher performance than traditional learning machines and has been introduced as a powerful tool for solving classification problems. However, there are some limitations, including outlier-handling, long learning times, and an increased number of support vectors (SV). To overcome these problems, SVM has been extended to fuzzy SVM (FSVM) combined with the fuzzy membership function [2]. The fuzzy membership function of FSVM is sequentially updated, and it leads to enhanced learning time. Therefore, it is well known that FSVM is useful in classification of time series data [1][2]. There are many studies on applications of FSVM in various fields.

Among those studies, Lee and his colleagues [3][6] achieved shortening of the learning time in the pattern classification of the time series data KOSPI200 in the financial market; D. Anguita et al. [7] used the Block-Toeplitz matrix to quickly calculate the gradient; and Platt [8]...
proposed an algorithm that presents the upper bound of learning time according to the number of learning patterns. However, there are challenges, including a relatively long learning time and the difficulty of handling large-scale healthcare data. To overcome these challenges, it is important to design an optimized fuzzy membership function for a given data set. In addition, considering the growing interest in big data analysis, research on improvement of learning time is continuously needed.

In this study, we present an FSVM classifier by applying a fuzzy membership function to each input point of SVM to enhance learning time and to handle large-scale time series data. The learning time performance of the FSVM is then evaluated for the large-scale time series dataset obtained from 32 patients with SA. Finally, we compare the proposed FSVM with the conventional SVM method.

2. Related Work

2.1 SVM

SVM is used to determine optimal hyperplanes that map the learning patterns into a high-dimensional feature space, and to find the maximum margin between the two classes in this space. SVM is a learning algorithm originally introduced by Vapnik [9] and successively extended by many other researchers. SVMs can work in combination with kernel techniques that automatically perform nonlinear mapping to a feature space where SVM can settle the nonlinear separation problems. In SVM, a convex quadratic programming problem is solved with optimal solutions of hyperplane by using kernel functions in a feature space. SVM can be defined as follows:

\[ \min \frac{1}{2} W \circ W + c \sum_{i=1}^{l} \xi_i \]

subject to \[ y_i (W \circ Z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \ldots, l \quad (1) \]

Detailed solution procedures are found in [9].

2.2 FSVM

FSVM was introduced by Lin and Wang [2], which combines fuzzy membership functions with a slack variable. This was proposed to enhance the SVM by reducing the effect of outliers and noise in input datasets. Additionally, FSVM is an improved pairwise classification method for SVM, in which a fuzzy membership function is introduced into the decision function, based on the pairwise classification. For the data in the classifiable regions, FSVM provides the same results as pairwise classification with the SVM method. For the data in the unclassifiable regions, FSVM generates better results than the pairwise classification with the SVM method. FSVM is defined as follows:

Suppose we are given a set \( X \) of labeled training points with an associated fuzzy membership function, \( \alpha \in (0, 1), s_i = f(t_i), \alpha \leq s_i \leq 1 \). The optimal hyperplane of FSVM is regarded as follows:

\[ \min \frac{1}{2} W \circ W + c \sum_{i=1}^{l} s_i \xi_i \]

subject to \[ y_i (W \circ Z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \ldots, l \quad (2) \]

In this study, we used the fuzzy membership function as

\[ s_i = f(t_i) = (1 - \sigma) \left( \frac{t_i - t_1}{t_1 - t_l} \right)^n + \sigma \quad (3) \]

where \( n \) is the dimension of fuzzy membership function and it is applied to the kernel functions (linear, polynomial, and RBF) in the range of \( n = 1, 2, 3, \ldots, 500 \). Equation (3) was first proposed by Lin and Wang [4], and generalized by Lee et al. [6,10].

3. Database

3.1 Participants and Data Recordings

All 32 patients with SA underwent polysomnography (PSG) using an Embla N7000 (Embla Systems Inc., USA) device. The PSG studies were performed to standard clinical guidelines at Samsung Medical Center in Seoul, Korea. All subjects provided written informed consent, and the PSG study protocol was approved by the Institutional Review Board of the Samsung Medical Center. The PSG data were annotated by a sleep specialist according to the AASM guidelines [11].

The ECG were recorded at 200 samples-per-second and stored with 16-bit resolution via a single-lead transducer during the PSG. The ECG signal was processed with a bandpass filter (0.5-32 Hz) to remove noise. The heartbeat signals were extracted from ECGs using a Pan and Tompkins algorithm [12] and manually corrected. The ECG signals were then segmented into lengths of 10 s.

www.ijfis.org  Design of a Fast Learning Classifier for Sleep Apnea Database based on Fuzzy SVM | 188
Table 1. Participants information in detail

| Measures                  | Mean   | SD [range] |
|---------------------------|--------|------------|
| Gender (male:female)      | 27:5   |            |
| Age (yr)                  | 56.31  | 10.59 [33-77] |
| BMI (kg/m²)               | 26.8   | 2.97 [21.6-32.4] |
| AHI (per hour)            | 47.96  | 13.77 [30.0-84.9] |
| Total recording time (min)| 442.08 | 37.22 [369.1-515.5] |
| Total sleep time (min)    | 340.06 | 70.53 [148.0-474.0] |
| Sleep efficiency (%)      | 76.71  | 14.20 [32.5-92.3] |

BMI, body mass index; AHI, apnea hypopnea index.

3.2 Large-Scale Time Series Database

The large-scale time series database consists of seven parameters extracted from the heartbeat signals of the SA patients. The heartbeat signals were analyzed in the time and frequency domains to extract parameters. In the time domain, three parameters were calculated as the mean of normal-to-normal intervals (NN), the standard deviation of NN (SDNN), and the root-mean-square of successive differences of NN (RMSSD). Another four parameters were extracted in the frequency domain, including very low-frequency (VLF), low-frequency (LF), high-frequency (HF) and LF/HF ratio. All selected parameters are well known in sleep studies, and have been used in various studies to detect or classify SA [13–15]. All 98,060 segments were combined from extracted parameters in the SA database (Table 2). Eighty percent of these were designated for the training set, and the rest for the test set.

4. Experiments

4.1 Environment

A large-scale time series SA database built from 32 patients was used for this study. The SVM and FSVM were used as a pattern classifier, and the three types of kernel (linear, polynomial, and RBF) were selected. In particular, the parameter \( n \) of the fuzzy membership function used in FSVM is applied from \( n = 1 \) to \( 500 \). In our experiment, the parameters were \( d \) (the polynomial degree), \( c \) (the trade-off point), SV, and std (standard deviation); where \( d = 1, 2, 3, c = 100, 500, 1000, 1500, 2000, \) and std = 1.0, 1.1, 1.2, ... , 9.9, respectively. The hardware specifications used in the experiment are as follows: OS = Win7; CPU = intel-core-i5-3330-3.00 GHz; RAM = 4.00 GHz; SVM and FSVM = C++.

4.2 Results and Discussion

In our experiment, we demonstrated an optimal model of the FSVM, which combines the fuzzy membership function with SVM to enhance learning time for the SA database. The experiments were done by changing the dimension parameter \( n \) of the fuzzy membership function from 1 to 500, sequentially. The experimental results are shown in Table 3, including the experimental results of the classification hit rate and learning time of SVM and FSVM. The maximum classification hit rate of SVM and FSVM is the same: 93.23%, but there was a big difference between learning time and the number of SVs. The learning time was significantly faster: 1,304 times, 2,787 times and 2,994 times for linear, polynomial and RBF kernels, respectively.

Figures 1-3 show the classification hit rate for the parameter \( n \) of FSVMs fuzzy membership functions. The average of the classification hit rate by each kernel function is 77.5% for linear kernel, 76.26% for polynomial and 43.56% for RBF, respectively. For parameter \( n \), linear and polynomial show gradual and stable classification hit rates in the region where \( n \geq 250 \). However, the case of kernel RBF, the variability is high and the classification hit rate is low.

Table 4 shows the learning time for the parameter \( n \) of FSVMs fuzzy membership functions. The average of the classification hit rate by each kernel function is 77.5% for linear kernel, 76.26% for polynomial and 43.56% for RBF, respectively. For parameter \( n \), linear and polynomial show gradual and stable classification hit rates in the region where \( n \geq 250 \). However, in the case of kernel RBF, the variability is high and the classification hit rate is low.

Table 4 shows the learning time for the parameter \( n \) (1, ..., 500) of fuzzy member functions when the maximum classification hit rate of FSVM is 93.23%. The frequency that represents the maximum classification hit rate of 93.23% is 50 (10%) for linear, 60 (15%) for polynomial, and 37 (7%) for RBF. The shortest times of the parameter \( n \) of fuzzy membership function were RBF (3 s for \( n = 2 \)), polynomial (17 s for \( n = 433 \)), and linear (35 s for \( n = 4 \)). The longest learning times were polynomial (5,908 s for \( n = 279 \)), linear (1,951 s for \( n = 231 \)), and RBF (904 s for \( n = 48 \)).
Table 2. Example of parameters per each epoch obtained from the SA database

| # epochs | NN    | SDNN  | RMSSD  | VLF    | LF    | HF    | LF/HF ratio | Label |
|----------|-------|-------|--------|--------|-------|-------|-------------|-------|
| 1        | 111.936 | 1506.506 | 2136.395 | 2973.803 | 2.399 | 4.842 | 0.495       | 1     |
| 2        | 110.564 | 1491.715 | 2100.032 | 2988.900 | 2.393 | 4.484 | 0.533       | 1     |
| 3        | 111.307 | 1486.296 | 2107.827 | 2979.534 | 2.774 | 5.606 | 0.494       | -1    |
| 4        | 112.004 | 1495.731 | 2121.206 | 2966.919 | 2.493 | 7.196 | 0.346       | 1     |
| 5        | 111.600 | 1461.560 | 2072.975 | 2981.679 | 2.487 | 2.846 | 0.873       | 1     |
| 6        | 110.741 | 1425.734 | 2022.367 | 2985.789 | 4.080 | 3.994 | 1.021       | -1    |
| 7        | 110.581 | 1477.179 | 2094.892 | 2982.092 | 3.289 | 9.905 | 0.332       | 1     |

Table 3. Learning time and classification hit ratios for SVM and FSVM

| Model | Kernel | Parameters | SV | Learning time (s) | Classification hit rate (%) |
|-------|--------|------------|----|-------------------|-----------------------------|
| SVM   | Linear | 100        |    | 7,452             | 45,645                      | 93.23                      |
|       | Polynomial | 100 | d = 1 | 7,439             | 47,389                      | 93.23                      |
|       | RBF    | 100        | 1.7 | 22,702            | 8,983                       | 93.20                      |
| FSVM  | Linear | 1000       |    | 9                 | 35                          | 93.23                      |
|       | Polynomial | 100 | d = 1 | 2,437             | 17                          | 93.23                      |
|       | RBF    | 500        | 9.9 | 151               | 3                           | 93.23                      |

Table 4. Learning times for parameter $n$ of the fuzzy membership function in FSVM

| Model | Kernel | Maximum hit rate (n=500) | Learning time (s) | Fuzzy membership index n |
|-------|--------|--------------------------|-------------------|-------------------------|
|       |        | min_time | max_time | index_n | index_n |
| FSVM  | Linear | 50 (10%) | 35       | 4        | 1,951 | 231 |
|       | Polynomial | 60 (15%) | 17       | 483      | 5,908 | 279 |
|       | RBF    | 37 (7%)   | 3        | 2        | 904   | 48  |

Figure 2. Classification hit ratio of FSVM - polynomial classifier (%).

Figure 3. Classification hit ratio of FSVM - RBF classifier (%).
Figures 4-6 show highly ranked learning times for the parameter $n$ of the fuzzy membership function in the FSVM, depending on kernels. Among the parameters $n$ obtained with the maximum classification hit rate of 93.23%, only 25 highly ranked numbers, in terms of learning time (No.1-No.25), are shown in the graph. The interval of the parameter $n$ for fuzzy membership function is $n = 1$ to 500. That is, the maximum classification hit rate (93.23%) is obtained in a given interval. Thus, a classifier implementing the shortest learning time can be designed. On the other hand, for $n$ greater than 500, repeated results are obtained where the learning period is longer but the interval of $n$ is reduced.

5. Conclusions

In this study, we proposed an optimal FSVM classifier model suitable for shortening the learning time by combining the fuzzy membership function to handle a large-scale time series database. The usefulness of FSVM in reducing learning time was verified by using a large-scale SA database. The results show that the proposed FSVM classifier model can reduce the learning time by about 15,215 times, compared to the existing SVM classifier model at the 93.23% classification hit rate. Finally, an optimal FSVM classifier was designed, reducing the learning time to 3 s for the RBF kernel, with parameter $n = 2$ ($c = 500$, std = 9.9, and SV = 151). Generally, it is important to select a fuzzy membership function in which the most recent pattern influences the next pattern when solving the classification problem in the large-scale SA database with time series characteristics. Through the proposed FSVM, we confirm that the performance of the optimal classifier that reduces the influence of outliers and shortens the learning time is useful. In the future, a real-time automatic classifier implementation using FSVM in SA database may be possible.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgements

This work was supported by the Human Resource Training Program for Regional Innovation and Creativity through the Ministry of Education and National Research Foundation of Korea (NRF-2014H1C1A1063845).
References

[1] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines. Cambridge: Cambridge University Press, 2000.

[2] C. F. Lin and S. D. Wang, “Fuzzy support vector machines,” IEEE Transactions on Neural Networks, vol. 13, no. 2, pp. 464-471, 2002. http://doi.org/10.1109/72.991432

[3] S. Y. Lee, S. Y. Sohn, C. E. Kim, and Y. B. Lee, “Fuzzy support vector machine for pattern classification of time series data of KOSPI200 Index,” Journal of Fuzzy Logic and Intelligent Systems, vol. 14, no. 1, pp. 52-56, 2004. http://doi.org/10.5391/JFKIS.2004.14.1.052

[4] S. Y. Lee, “A comparative study of model design and performance test in financial index pattern classification system,” Ph.D. dissertation, Yonsei University, Seoul, Korea, 2004.

[5] S. Y. Lee, “On the fuzzy membership function of time series data,” Journal of Fuzzy Logic and Intelligent Systems, vol. 17, no. 6, pp. 799-803, 2007.

[6] S. Lee, E. Paik, and Y. Lee, “Time series data pattern classification using fuzzy membership functions and support vector machines,” ICS, vol. 90, pp. 111-114, 2003.

[7] D. Anguita, A. Boni, and S. Pace, “Fast training of support vector machines for regression,” in Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, 2000, pp. 210-214. http://doi.org/10.1109/IJCNN.2000.861459

[8] J. C. Platt, “Sequential minimal optimization: a fast algorithm for training support vector machines,” Microsoft, Technical Report MSR-TR-98-14, 1998.

[9] V. N. Vapnik, Statistical Learning Theory. New York, NY: John Wiley & Sons, 1998.

[10] S. Y. Lee, D. Y. Ahn, M. H. Song, and K. Y. Lee, “The classification of electrocardiograph arrhythmia patterns using fuzzy support vector machines,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 11, no. 3, pp. 204-210, 2011. http://doi.org/10.5391/IJFIS.2011.11.3.204

[11] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. Marcus, and B. Vaughn, The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Darien, IL: American Academy of Sleep Medicine, 2012.

[12] J. Pan and W. J. Tompkins, “A real-time QRS detection algorithm,” IEEE Transactions on Biomedical Engineering, vol. 32, no. 3, pp. 230-236, 1985. https://doi.org/10.1109/TBME.1985.325532

[13] E. Urtnasan, J. U. Park, S. Lee, and K. J. Lee, “Optimal classifier for detection of obstructive sleep apnea using a heartbeat signal,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 17, no. 3, pp. 76-81, 2017. https://doi.org/10.5391/IJFIS.2017.17.2.76

[14] E. Urtnasan, H. K. Lee, H. Kim, and K. J. Lee, “Sleep apnea detection using a piezo snoring sensor: a pilot study,” Journal of Biomedical Engineering Research, vol. 35, no. 4, pp. 75-80, 2014. https://doi.org/10.9718/JBER.2014.35.4.75

[15] H. D. Nguyen, B. Wilkins, Q. Cheng, and B. Benjamin, “An online sleep apnea detection method based on recurrence quantification analysis,” IEEE Journal of Biomedical Health Informatics, vol. 18, no. 4, pp. 1285-1293, 2014. https://doi.org/10.1109/JBHI.2013.2292928

SooYong Lee received his Ph.D. in Computer Science from the Yonsei University, Seoul, Korea, in 2004. He also received his Ph.D. in Mathematics from the Kyunghee University, Seoul, Korea, in 1992. He is a faculty member at Yonsei University, since 2004. He has been working on research related to artificial intelligence, machine learning and data mining.

Erdenebayar Urtnasan received his B.S. in computer science from Huree University, Ulaanbaatar, Mongolia, in 2007 and M.S. in electronic engineering from Inha University, Incheon, Korea, in 2010, respectively. He is a faculty member at Huree University, since 2010. He is currently Ph.D. candidate at Department of Biomedical Engineering, Yonsei University. His current research interests are sleep signal...
analysis and machine learning.

Kyoung-Joung Lee received his B.S., M.S., and Ph.D. in electric engineering from the Yonsei University, Seoul, Korea, in 1981, 1983 and 1988, respectively. He was an international fellow at Case Western Reserve University, USA, in 1993. He joined Yonsei University, Wonju, Korea, as a faculty member in 1989. His research interests include medical instrument, biosignal processing and biosystem modelling.