[1]Time Series Data Pattern Classification using Fuzzy Membership Functions and Support Vector Machines

－KOSPI 200 : Korea Composite Stock Price Index 200－

Sooyong Lee † Euree Paik ‡ and Yillbyng Lee ‡

† Dept. of Computer Science, Yonsei University, Wonju, 220-70, Korea
‡ Dept. of Economics, London School of Economics and Political Science, UK
‡ Division of and Information Engineering, and of BERC, Yonsei University, Seoul, Korea

E-mail: † hosu@hosu.yonsei.ac.kr, ‡ e.paik@lse.ac.uk, yblee@csai.yonsei.ac.kr

Abstract SVM (Support Vector Machine) is a binary classifier proposed by Vapnik. SVM has proven performance in various application fields by minimizing misclassification based on mathematical theories. Recently, FSVM that applies Fuzzy membership function to SVM has been proposed. In this study, it is proven that FSVM (polynomial kernel) has reduced learning time better than SVM when fuzzy membership functions of FSVM have been expanded from 2-dimension to bigger than 3 dimension.

Keyword Fuzzy membership function, SVM, Fuzzy Support Vector Machine, KOSPI 200, Pattern Classification, Stock Index prediction, Learning Time Reduction.

1. Introduction

In todays, worldwide stock markets have experienced dramatic volatility in their returns. Traditionally, two main approaches – time series analysis and fundamental analysis – exit in predicting stock price. But the results of statistical analysis are not quite satisfactory to meet our expectation. Besides, they have some limitation of applications according to the data characteristics and also require comparatively strict assumptions on the distribution.

As a result, artificial intelligence technologies are introduced in this area. Especially, SVM (Support Vector Machines) and FSVM (Fuzzy Support Vector Machines) have newly received special interests.

SVM (Support Vector Machine) model that attracts attention in pattern classification field is a learning theory developed by Vapnik [6] in 1998. It is to estimate decision-making membership functions using probability distribution during learning process and to classify in binary new data in binary classification. SVM is particularly widely used in many fields because it has high generalization feature in classification.

While existing learning algorithms embody ERM (Empirical Risk Minimization), SVM applies SRM (Structural Risk Minimization). The principle of SVM is to calculate optimal hyperplane that maximizes the width of the margin and minimizes misclassification after mapping given non-linear data to higher feature space.

2. Fuzzy Support Vector Machine

Although SVM has effective performance in a pattern classification or function estimation effectively just like a neural net, it is desirable to be improved in following respects. In classification there can be some data that have bigger influence on classification. Therefore data shall be classified properly in the first hand. On the other hand outliers and biases shall be minimized.

For example, if sequential data composed a trend and

[1] "This work was supported in part by BERC/KOSEF and in part by Brain Neuroinformatics Program sponsored by KMST."

the most recent data would influence pattern the most, sequential fuzzy membership functions should be defined. Therefore, all training data could be influenced by time series without dealt with uniformly in the learning process.

Accordingly, Lin proposed FSVM (Fuzzy Support Vector Machine) where fuzzy membership functions are combined with slack variables of SVM [3]. The characteristic of FSVM is that when the gradient of hyperplane is adjusted, slack variables, the measurement of misclassification, influence learning by combining with fuzzy membership functions.

In other words, $S=\{(y_1,x_1,s_1),\ldots,(y_l,x_l,s_l)\}$, $\sigma \le s_i \le 1$,

Where in, $x_i \in R^n$ is training data, $y_i=\{-1,+1\}$ is labeled data, $s_i=\{s_i \in R \| \sigma \le s_i \le 1\}$ is fuzzy membership. Fuzzy membership s_i shows the level that vector x_i belongs to a certain class. As ξ_i is the measurement of misclassification in SVM, $s_i\xi_i$ is the measurement of misclassification with different weight. The 2-dimensional fuzzy membership function that Lin proposed is as follow;

$$s_i = f(t_i) = (1-\sigma)[\frac{t_i - t_1}{t_l - t_1}]^2 + \sigma$$

In this article, the 2-dimensional membership function was expanded to $n \ge 3$ dimension and it was applied to binary classification on the rise and fall of KOSPI 200 index. According to experiments, misclassification was minimized and learning time was reduced significantly when n value is over $n \ge 3$ rather than $n=1,2$.

In the study on accelerated learning of SVM, Anguita[1] used block-Toeplitz matrix to calculate a gradient quickly and Yang[5] presented an algorithm to provide upper bound of learning time according to the number of data.

3. Experiment

Database was composed of 24 variables of daily data (domestic financial index: 14, overseas financial index: 10) from January 2000 to August 2002. Daily data were preprocessed to follow standard regular distribution by converting them to 137 weekly data by simple movement average method. The system specification was as follows; operating system - Windows 2000/sever, CPU - Pentium IV (Dual CPU : 1+1 GHz), RAM - 1,024 Mb. SVM and FSVM written in C++ were used on the system.

In the experiment to measure learning time by data size (Table-1), the fastest kernel function was RBF and Polynomial took the longest time. Particularly, when the number of data was 1,810 and kernel function was Polynomial, SVM took 179,957 second (=49.988 hours) and FSVM took 14,949 second (=4.1525 hours). Therefore, when FSVM was used, 45.8355 hours could be saved.[8]

3.1. Learning time comparison of the machines

The results of investigating the learning times according to the size of the data are shown in Table-1. In this experiment, the number of learning patterns were organized into 92, 127, 158, 200, 249, 293, 629, 1,810 . The fastest Kernel function was RBF.

Table-1. kernel learning time (sec)

#(data)	RBF		Polynomial	
	SVM	FSVM	SVM	FSVM
92	1	1	6	6
127	1	1	20	15
158	1	1	13	11
200	1	1	22	20
249	1	1	30	20
293	1	1	31	6
629	2	1	117	2
1,810	2,632	2,118	179,957	14,949

In Table 2, d stands for the Polynomial degree, c for the trade-off point and sv for the number of support vectors. Table 1 shows that the learning time of SVM takes the longest when the Kernel function is Polynomial. The results (Table 2) from the experiment on the parameters of SVM and FSVM of the Polynomial function, confirmed the existence of a significant difference in the learning times. When keeping both the accuracy rate of SVM and FSVM same (d=2, c=1,000 ; sv=98:60), the learning time rate changed by 6.8 . When SVM showed a higher accuracy rate (d=2, c=1,000 ; sv=98:60), the learning time rate changed by 8.5 . And when FSVM showed a higher accuracy rate (d=2, c=500 ; sv=98:66), the learning time rate changed by 5.3. FSVM is a combination of Fuzzy membership functions and Slack variables. Therefore FSVM has less support vectors compared to SVM. This leads to an expectation for a decrease in learning time of the machine in the experiment.

Table-2 learning time of Polynomial kernel

d	c	Time / sec		Support Vector	
		SVM	FSVM	SVM	FSVM
1	100	0.4	0.2	111	111
	500	2.5	0.6	113	111
	1000	4.9	1.2	113	111
	1500	7.3	1.8	113	111
	2000	6.6	2.3	113	111
2	100	3.5	2.6	94	82
	500	23.9	4.5	98	66
	1000	43.3	5.1	98	60
	1500	70.3	10.3	98	59
	2000	89.7	10.9	98	58
Average		25.2	4	104.9	88

Compared to the BP algorithm of Neural Network, SVM showed 12% and FSVM 40% decrease in the learning times. Particularly, when Kernel function was Polynomial, SVM took 49.988 hours (=179,957 seconds) while FSVM took 4.1525 hours (=14,949 seconds). By using FSVM, therefore 45.8255 hours could be saved.

In the investigation on the speed of learning ability of SVM, Anguita[2] used Block-Toeplitz matrix to calculate the gradient with speed.

3.2. Experiment to verify the usefulness

In order to verify the usefulness, the method of testing on the learning period and shifting the period on the test data was used. Through simulations, RBF Kernel function was proven to have an accuracy rate of 70% with BP algorithm of Neural Network, 77.78% with SVM and 88.89% with FSVM. It was also investigated if the turning points of the trend were properly classified. The accuracy rates of classifications were 64% with BP algorithm of Neural Network, 66.67% with SVM and 73.33% with FSVM.

Table-3 turning point

classifier	BPN	SVM	FSVM
Hit Rate (%)	64.00	66.67	73.33

Table-4 simulation

classifier	BPN	SVM	FSVM
Hit Rate (%)	70.00	77.78	88.89

4. Conclusion

FSVM pattern classifier where SVM is combined with Fuzzy membership functions could reduce learning time better than existing SVM when the degree of the membership functions is bigger than 3. Especially when kernel membership function was Polynomial, the learning time was reduced significantly. Conclusions on the investigation to confirm the performance of the FSVM pattern classifier are summarized below. This was investigated by applying the system on KOSPI 200 index.

Firstly, C.F Lin's FSVM pattern classifier, which is a combination of SVM (Support Vector Machine: SVM) and Fuzzy membership functions, minimizes the misclassification when the degree of the membership functions are greater than 3 rather than when it is less than 2. Secondly, FSVM pattern classifier provides a superior capability in classifications compared to previously established SVM and Neural Network systems. Also, FSVM pattern classifier reduces learning time in great deal. In particular, when Kernel function is Polynomial, a significant reduction in learning time can be observed. Thirdly, FSVM pattern classifier maintains a more stable structure compared to previously developed Neural Network, by keeping consistent results on each parameters. Therefore when constructing a portfolio in future markets and option markets, which derive from KOSPI 200 index, it would be desirable to apply FSVM system in making financial decisions.

References

[1] D. Anguita, Fast Training of Support Vector Machines for Regression, www.kernel-machines.org

[2] N. Cristianini,An Introduction to Support Vector Machines, Cambridge University Press, 2000.

[3] C. F. Lin A. D. Wang, Fuzzy Support Vector Machines, IEEE Transactions on Neural Networks, Vol.13, No.2, pp 464-471, 2002.

[4] D. Anguita, Fast Training of Support Vector Machine for Regression, www.kernel-machines.org

[5] D. Yang, Provably Fast Training Algorithms for SVM, www.kernel-machines.org

[6] V. Vapnik, The Nature of Statistical Learning Theory, Springer NY, 1995.

[7] D. S. Kim, Neural Networks Theory and Applications, Public. High-Tech, 1993.

[8] S.Y. Lee, Y.B. Lee, Learning Time Reduction of Fuzzy SVM in Classification, The 7th International Conference of te Decision Science Institute (DSI) combined with the 8th Annual Meeting of the Asia-Pacific Region, 2003.7

[9] S.Y. Lee Y.B. Lee, Pattern Classifier utilizing Fuzzy Theory and SVM, Proceeding of Korean Securities Association, pp789-809, 2002. 10 .