

Support Vector Machines for Pattern Classification

Presented by Prof. Shigeo Abe

Graduate School of Science and Technology

Kobe University

Kobe, Japan

abe@eedept.kobe-u.ac.jp

Brief Description:

Support vector machines (SVMs) have gained wide acceptance due to its solid theoretical basis and high generalization ability for wide range of applications. This tutorial emphasizes the applicability of SVMs to pattern classification problems.

Three-layer neural networks are universal classifiers in that they can classify any labeled data correctly if there are no identical data in different classes. In training multilayer neural network classifiers, usually, network weights are corrected so that the sum of squared errors between the network outputs and the desired outputs is minimized. But since the decision boundaries between classes acquired by training are not directly determined, the generalization ability depends on the training method. And it degrades greatly when the number of training data is small and the overlap among classes is scarce or non-existent.

On the other hand, in training SVMs the decision boundaries are determined directly from the training data so that the separating margins of decision boundaries are maximized in the high dimensional space called feature space. This learning strategy, based on statistical learning theory developed by Vapnik, minimizes the classification errors of the training data and the unknown data.

Therefore, the generalization abilities of SVMs and other classifiers differ significantly especially when the number of training data is small. This means that if some mechanism to maximize margins of decision boundaries is introduced to non-SVM type classifiers, their performance degradation will be prevented when the class overlap is scarce or non-existent.

In the original SVM, the n -class classification problem is converted into n two-class problems, and in the i th two-class problem we determine the optimal decision function that separates class i from the remaining classes. In classification, if one of the n decision functions classifies an unknown datum into a definite class, it is classified into that class. In this formulation, if more than one decision function classify a datum into definite classes, or no decision functions classify the datum into a definite class, the datum is unclassifiable.

To resolve unclassifiable regions for SVMs we discuss four types of SVMs: one-against-all SVMs; pairwise SVMs; ECOC (Error Correcting Output Code) SVMs; all-at-once SVMs; and their variants.

Another problem of SVMs is slow training. Since SVMs are trained by solving a quadratic programming problem with the number of variables equal to the number of training data, training is slow for a large number of training data. We discuss training of SVMs by decomposition techniques combined with a steepest ascent method.

To improve generalization ability of non-SVM type classifiers, we introduce mechanisms to maximize decision margins. Multilayer neural networks can be trained based on the training method for SVMs. But in this tutorial we discuss layer by layer training. For a fuzzy classifier with ellipsoidal regions, we maximize margins of slopes of the membership functions to improve generalization ability.

Outline of the Tutorial:

1. Introduction to SVMs
2. Direct and Indirect Decision Functions
3. Architecture of SVMs
4. Characteristics of Solutions
5. Multiclass SVMs
6. Training Methods
7. Maximum Margin Neural Networks
8. Maximum Margin Fuzzy Classifiers
9. Performance Comparisons

Brief Curriculum Vitae

Professor Shigeo Abe received the B.S. degree in Electronics Engineering, the M.S. degree in Electrical engineering, and the D. Eng. degree, all from Kyoto University, Kyoto, Japan, in 1970, 1972, and 1984, respectively. After 25 years in the industry, he was appointed as full professor of Electrical Engineering, Kobe University in April 1997. He is now a professor of Graduate School of Science and Technology, Kobe University. His research interests include pattern classification and function approximation using neural networks, fuzzy systems, and SVMs. He is the author of *Neural Networks and Fuzzy Systems: Theory and Applications* (Kluwer, 1996) and *Pattern Classification: Neuro-fuzzy Methods and their Comparison* (Springer-Verlag, 2001). Dr. Abe was awarded an outstanding paper prize from the Institute of Electrical Engineers of Japan in 1984 and 1995. He is a member of IEEE, INNS, and several Japanese Societies.