

Early Work on Polynomial nets that leads to SVM Quality Solutions

The basic ideas behind support vector machines have been explored in the early sixties. The idea of finding robust (generalizing) separating surfaces by placing the separating surface in a position in the input space that is at a maximal distance from all support vectors is not new. Cover refers to these support vectors as "extreme points" (in the context of polynomial threshold gates. In fact, support vectors are a subset of extreme points) in his Ph.D. Dissertation (1964) at Stanford [see also T. M. Cover(1965) "Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition," IEEE Trans. Elec. Comp, EC-14, 326-334. Also, see M. Hassoun, Fundamentals of Artificial Neural Networks, pp. 27-29 (MIT Press, 1995)]. Hassoun and Clark (1988) and Hassoun and Song (1992) have proposed a family of simple incremental learning rules based on the Ho-Kashyap method of solving linear inequalities. These rules automatically generate the SVM separating hyperplane for linearly separable problems without explicitly solving for the support vectors/extreme points. Another rule proposed by Hassoun and Song is the AHKIII rule which generates a hyperplane for linearly nonseparable patterns which tend to minimize the rate of misclassifications by automatically detecting extreme points and positioning the hyperplane accordingly. Please refer to the paper M. Hassoun and J. Song (1992) "Adaptive Ho-Kashyap Rules for Perceptron Training," IEEE Trans. Neural Networks, Vol. 3(1), pp. 51-61. Also, M. Hassoun and D. Clark (1988), "An Adaptive Attentive Learning Algorithm for Single-Layer Neural Networks," Proc. IEEE Annual Conf. on Neural Networks, Vol. I, pp. 431-440. Finally, these results may also be conveniently accessed from the textbook M. Hassoun "Fundamentals of Artificial Neural Networks," pp. 78-83 (MIT Press, 1995).

Here, we note that in our work, maximizing the margin between training examples is equivalent to minimizing (via incremental steepest gradient descent) the "mean square error criterion," simultaneously in weight and margin spaces.

The .gif file [HassounSVM84](#) shows a simulation result where a polynomial net (in this case a 2-input second order polynomial net with a perceptron in the output layer) is trained using the Ho-Kashyap method to generate the optimal quadratic decision surface for a simple linearly nonseparable problem. As can be seen from the figure, the decision boundary is optimal in the sense of maximizing the margins between the extreme patterns (support vectors) and the separating boundary. THIS, to my knowledge, IS THE FIRST PUBLISHED APPLICATION OF A POLYNOMIAL SVM!! [We like to call it Extreme Point Machine (EPM) for historical reasons].

This result and problem formulation has been published in M. H. Hassoun (1986) Optical Threshold Gates and Logical Signal Processing, Ph.D. Dissertation, Dept. ECE, Wayne State Univ., Detroit, MI (Chapter III). The dissertation also describes a hand-written digit recognition application employing a single perceptron acting as an SVM. The error rate was 14% (ASET OF 273 SAMPLES WERE USED). (Similar results were reported relatively recently in Boser et al., 1992. The ACM paper).

[Link to Extreme Point Machine \(EPM\).Java Applet](#)



[Click here to go back to Java applets page](#)