

# Chapter XIV

## Classification in GIS Using Support Vector Machines

**Alina Lazar**

*Youngstown State University, USA*

**Bradley A. Shellito**

*Youngstown State University, USA*

### ABSTRACT

*Support Vector Machines (SVM) are powerful tools for classification of data. This article describes the functionality of SVM including their design and operation. SVM have been shown to provide high classification accuracies and have good generalization capabilities. SVM can classify linearly separable data as well as nonlinearly separable data through the use of the kernel function. The advantages of using SVM are discussed along with the standard types of kernel functions. Furthermore, the effectiveness of applying SVM to large, spatial datasets derived from Geographic Information Systems (GIS) is also described. Future trends and applications are also discussed – the described extracted dataset contains seven independent variables related to urban development plus a class label which denotes the urban areas versus the rural areas. This large dataset, with over a million instances really proves the generalization capabilities of the SVM methods. Also, the spatial property allows experts to analyze the error signal.*

### INTRODUCTION

This entry addresses the usage of Support Vector Machines (SVM) for classification of remotely sensed data and other spatial data created from Geographic Information Systems (GIS). Variability, noise, and the nonlinear separability property

are problems that must be confronted when dealing with spatial data, and SVM have become popular tools for classification and regression as they address most of these problems.

SVM are widely recognized as classification tools by computer scientists, data mining and machine learning researchers. To date, GIS

researchers have used older, more developed techniques such as the logistic model for regression and artificial neural networks for classification and regression. However, we consider that the SVM method has much potential and the goal of our chapter is to bring it to the attention of the GIS community.

SVM translate the input data into a larger feature space, using a nonlinear mapping, where the instances are linearly separable. A straight model line in the new feature space corresponds to a nonlinear model in the original space. In order to build the straight delimitation line called “*the maximum margin hyperplane*” in the new space, a quadratic optimization learning algorithm is applied. The support vectors are the instances with the minimum distance to the hyperplane. The new space is the result of the dot product of the data points in the feature space.

## **CLASSIFICATION PROBLEMS ON SPATIAL DATA**

Machine learning and data mining methodologies (such as artificial neural networks and agent-based modeling) have been adapted for the classification of geospatial data in numerous studies such as Fischer (1994), Pijanowski and Alexandridis (2002), and Brown et al (2005). Proposed by Vapnik (1999) in 1992 and improved in 1995, the SVM algorithm was inspired by statistical learning theory. SVM became a very popular classification tool after it was successfully implemented and applied to handwritten digit recognition. Lately, SVM have been adapted and utilized in the classification of remotely sensed data (e.g., see Foody and Mathur 2004, Mantero et al 2005, Marcal et al 2005, Niishi and Eguchi 2005) for land cover mapping. SVMs have also been applied to the classification of hyperspectral data by Melgani and Bruzzone (2004). Other studies have utilized SVMs as a classification tool with remotely sensed data alongside data derived from traditional GIS

sources (Song et al 2004, Watanachaturaporn et al 2005). SVM have been less commonly utilized as a classification tool in GIS modeling. Guo et al (2005) have utilized SVM together with GIS tools for studying the ecology of oak death in California. Other studies (e.g., see Shellito and Lazar 2005, Lazar and Shellito 2005) use SVM with GIS data to examine distributions of urbanization in Ohio with regard to predictor variables (in raster format) of urban development. The SVM methods provide not only higher accuracy compared to other classifiers using fewer parameters, but also are particularly robust to instances of missing data, errors, and redundancy, all of which typically characterize real datasets.

## **THE SVM APPROACH AND CRITICAL ISSUES**

The SVM algorithms of Cristianini and Shawe-Taylor (2000) and Schölkopf and Smola (2002) inspired from the statistical learning theory (Vapnik, 1999) combine together kernel methodology and convex function optimization to solve classification and regression problems. With numerous advantages (Table 1) and several available computer implementations (Table 2), SVM has become a viable and efficient tool.

SVM builds a decision boundary by choosing the separating hyperplane, which maximizes the margin (Figure 1) between positive and negative examples. Decision boundaries with large margins usually generalize better than those with small margins and they will not be affected by using a big training set. The training points that lie closest to the separating hyperplane are called support vectors. The algorithm used to build the hyperplane is a quadratic optimization algorithm.

Let  $x_i$ ,  $i = 1, \dots, l$ , be a training set of examples for the problem of interest. For a binary classification task, each input instance  $x_i \in R^N$  in the attribute space will have associated a decision  $y_i \in \{-1, 1\}$ .

5 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/classification-gis-using-support-vector/20393](http://www.igi-global.com/chapter/classification-gis-using-support-vector/20393)

## Related Content

---

### Utilizing BCF and IFC for Automated Dimensional Compliance Control

Neil Hyland, Shawn E. O'Keeffe and Shane Brodie (2017). *International Journal of 3-D Information Modeling* (pp. 16-36).

[www.irma-international.org/article/utilizing-bcf-and-ifc-for-automated-dimensional-compliance-control/216386](http://www.irma-international.org/article/utilizing-bcf-and-ifc-for-automated-dimensional-compliance-control/216386)

### Understanding Health Disparities through Geographic Information Systems

Samuel Soret, Karl J. McCleary, Patrick A. Rivers and Susanne B. Montgomery (2003). *Geographic Information Systems and Health Applications* (pp. 12-42).

[www.irma-international.org/chapter/understanding-health-disparities-through-geographic/18833](http://www.irma-international.org/chapter/understanding-health-disparities-through-geographic/18833)

### Application of a GIS-Based Statistical Method to Assess Spatio-Temporal Changes in Breast Cancer Clustering in the Northeastern United States

Daikwon Han and Peter A. Rogerson (2003). *Geographic Information Systems and Health Applications* (pp. 114-138).

[www.irma-international.org/chapter/application-gis-based-statistical-method/18838](http://www.irma-international.org/chapter/application-gis-based-statistical-method/18838)

### Optimizing School Bus Stop Placement in Howard County, Maryland: A GIS-Based Heuristic Approach

Michael Galdi and Paporn Thebpanya (2016). *International Journal of Applied Geospatial Research* (pp. 30-44).

[www.irma-international.org/article/optimizing-school-bus-stop-placement-in-howard-county-maryland/143075](http://www.irma-international.org/article/optimizing-school-bus-stop-placement-in-howard-county-maryland/143075)

### Target Evaluation and Correlation Method (TECM) as an Assessment Approach to Global Earth Observation System of Systems (GEOSS)

Samuel Epelbaum, Mo Mansouri, Alex Gorod, Brian Sauser and Alexander Fridman (2013). *Emerging Methods and Multidisciplinary Applications in Geospatial Research* (pp. 301-327).

[www.irma-international.org/chapter/target-evaluation-correlation-method-tecm/68266](http://www.irma-international.org/chapter/target-evaluation-correlation-method-tecm/68266)