Support Vector Machines

Ty Crabtree
Washington State University Vancouver
crabtreee.tyler@gmail.com

1. INTRODUCTION

Support Vector Machines utilizes machine learning to find an optimal hyperplane within a data-set. The Support Vector Machine algorithm finds the ideal hyperplane by comparing the largest minimum distance of the training examples; the algorithm uses a sample from two sets of data as training examples. The optimal separating hyperplane is computed by finding the minimum and maximum separating planes for a data-set. The margin of the training data is used to find calculate the optimal hyperplane.

1.1 Support Vector Machine Algorithm

The optimal hyperplane can be represented is found by using the algorithm in *Figure 1*. The term x represents the closest training examples to the hyperplane. The training examples that correctly predict the optimal hyperplane are the support vector.

$$\beta = weight \tag{1}$$

$$\beta_0 = bias \tag{2}$$

$$x = supportVector \tag{3}$$

$$\beta_0 + \beta^t x = 1 \tag{4}$$

Figure 1: Optimal hyperplane formula.

After the term x is calculated, the distance between a point on x and a hyperplane is calculated. Next the hyperplane margin is found, M. Finally, using the maximum value for M will be used to correctly predict the optimal hyperplane by using the training examples.

$$|(\beta_0 + \beta^t x|)/(|\beta|) = distance \tag{5}$$

$$M = margin \tag{6}$$

$$M = 2/(|\beta|) \tag{7}$$

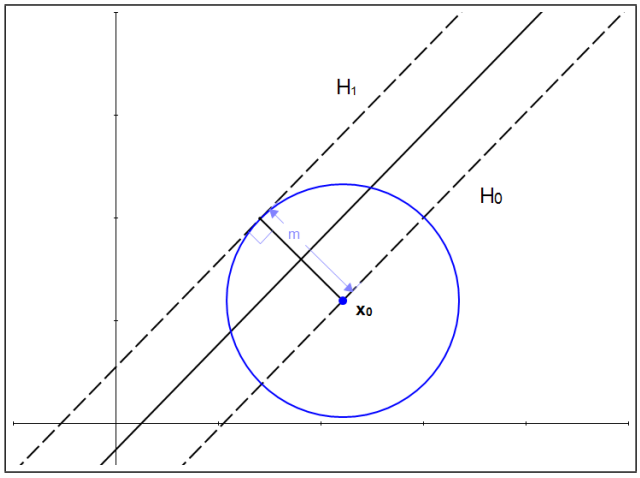

Figure 2: Formula example.

$$x_i = trainingExamples \tag{8}$$

$$M_{max} = (1/2)/\beta^2 \tag{9}$$

$$where \tag{10}$$

$$y_i(\beta_0 + \beta^t x) >= 1 \forall i \tag{11}$$

1.2 Advantages of Support Vector Machines

Support Vector Machines (SVM) are useful in predicting and separating solvent from insolvent companies. Support Vector Machines also provide prediction flexibility by not requiring linear data, and data-sets may include variable forms for all data. This algorithm is non-parametric and operates locally. Utilizing the SVM's algorithm can be used to manipulate financial ratios. The data extrapolated is linearly separable. SVMs are a robust prediction method, even when the training sample show some bias.

1.3 Machine Learning

Using machine learning, support vector machines are supervised learning models with associated learning algorithms

that analyze data used for classification analysis. By using training examples, each marked as belonging to one or the other of two categories. A typical SVM model depicts examples points. These points are mapped to allow the examples to separate the given. categories. From here, the plot point is divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

1.4 Background

Vladimir Naumovich Vapnik is one of the credited with the development of of statistical learning, and the co-inventor of the support vector machine method.

1.5 Research

I produced two programs to aid Support Vector Machine reasearch. The first, svm.py produces a model of Support Vector Machines while displaying the correct classification percentage. The typical range is usually calculated to be about 45-100 percent correct in classification. The program will randomly generate a new plot every 3 seconds and display this to the user. Also, I produced ResearchScraper.py. This program quickly parses the Internet and scrapes data for relevent information. This program then writes this information to a .txt file. The intention here to to quickly produce work and a background to a subject for future research.

2. REFERENCES

APPENDIX

A. INTRODUCTION

A.1 Algorithm

A.2 Machine Learning

A.3 Background

A.4 Research