RESEARCHERID
THOMSON REUTERS

[Gulia, 6(6): June 2019]
IDSTM-2019

ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

A REVIEW PAPER ON SUPPORT VECTOR MACHINE

Ankit Gulia[1], Anisha Dhaiya[2] & Anshul[3]
[1]M.Tech Scholar, Department of Computer Science & Engineering, SET (SOLDHA), INDIA
[2]Asistant Professor, Department of Computer Science & Engineering, SET (SOLDHA), INDIA
[3]Asistant Professor, Department of Computer Science & Engineering, SET (SOLDHA), INDIA

ABSTRACT

A brief introduction to Support Vector Machine (SVM) is given in this paper. It is also explained which applications can use SVM for the classification purpose. SVM is defined by a separating hyperplane which mainly classify the classes into two or more types or categories. In this the different classes are classified based on the particular category of data it belongs to. The review is done for SVM based techniques and how it will be helpful in future.

Keywords: *Gamma, Margin, Kernel, Tuning Parameters.*

I. INTRODUCTION

A Support Vector Machine (SVM) can be described as a discriminative classifier formally defined by a separating hyperplane. It is an algorithm which outputs an optimal hyperplane by categorizing new examples based on given labeled training data (supervised learning) . In two dimentional(2-D) space this hyperplane is a line dividing a plane in two parts where in each class lay in either side.

II. MATERIALS AND METHODS

Suppose we are given with a plot of two label classes on graph as shown in image (*Fig. 1*):

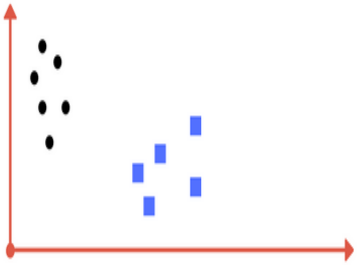

Fig. 1.: Draw a line that separates black circles and blue squares

We can also redraw the above image in a similar way as (*Fig. 2*). It fairly separates the two classes. Any point that is on the left of line falls into black circle class and on right falls into blue square class we can call it Separation of classes and this is what SVM does. SVM finds out a line or a hyper-plane (in multidimensional space that separate outs classes).

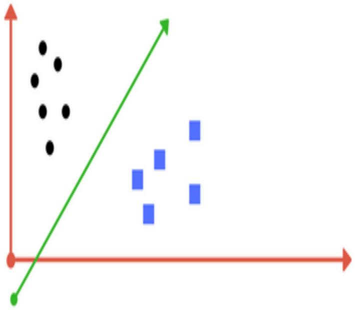

Fig. 2. Sample cut to divide into two classes

Now if we consider another example with data which is shown in *Fig.3*. Clearly, We can see that there is no line that can separate the two classes in this x-y plane. So what should we do now? We apply transformation to it and add one more dimension to it as we call it z-axis. Lets assume value of points on z plane, $z = x^2 + y^2$. In this case we can manipulate it based on the distance of point from z-origin.

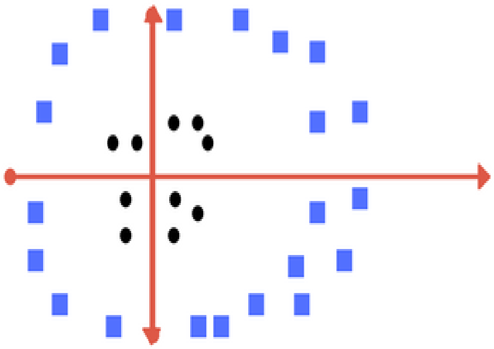

Fig. 3.Can you draw a separating line in this plane?

Now if we plot in z-axis, a clear separation is visible and a line can be drawn as shown in *Fig. 4*.

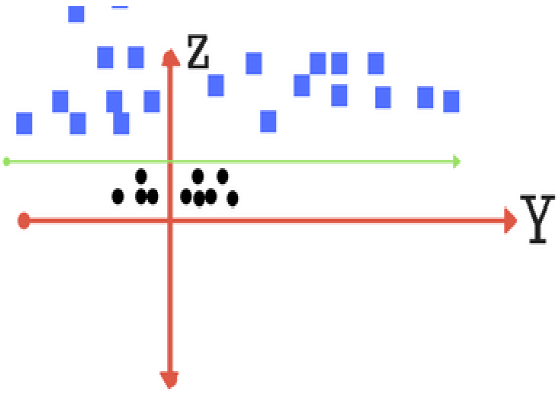

Fig. 4.plot of z-y axis. A separation can be made here

When we transform back this line to original plane, it maps to circular boundary as shown in image (*Fig. 5*). These transformations are called kernels.

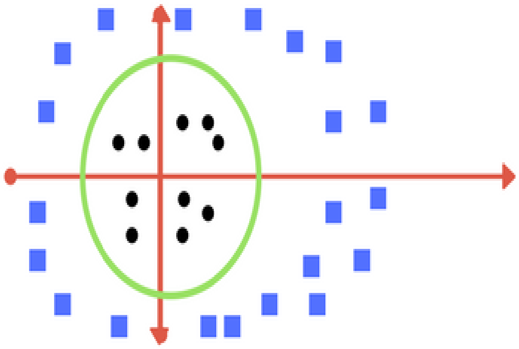

Fig. 5.Transforming back to x-y plane, a line transforms to circle

Tuning parameters for SVM: Kernel, Regularization, Gamma and Margin.

Kernel

In a linear SVM, the problem is transformed with the help of some linear algebra for the learning of the hyperplane and this is where the kernel plays role. We can calculate the equation for prediction for a new input in linear kernel using the dot product between the input *(i)* and each support vector *(i_n)* as :

$\underline{f(x) = B(0) + sum(a_n * (i,i_n))}$ ...*(1)*

The equation*(1)* above involves the calculation of the inner products of a new input vector (*i*) with all support vectors in training data. The co-efficient *B(0)* and i_n (for each input) must be estimated from the training data by the learning algorithm.The polynomial kernel with degree-*d polynomials* can be expressed as $K(i, i_n) = 1 + sum(i * i_n)^d$. The exponential kernel also known as RBF can be written as $K(i, i_n) = exp(-\gamma * (\|i—i_n\|^2)$. Polynomial and exponential kernels are responsible for the calculation of separation line in higher dimension. This is called kernel trick.

Regularization

The Regularization parameter (C) tells the SVM optimization how much misclassifying of the training example we want to avoid. A smaller-margin hyperplane is chosen by the optimization for large values of C if that hyperplane does a better job of getting all the training points classified correctly. While on the other hand , a small value of C will make the optimizer to look for a larger-margin separating hyperplane, even if that hyperplane misclassifies more points.

The images below (*Fig. 6(a)* and *Fig. 6(b)*) are example of two different regularization parameter. In *Fig. 6(a)* there are some misclassification due to lower regularization value whereas Higher value leads to results like we see in *Fig. 6(b).*

ISSN 2348 – 8034
Impact Factor- 5.070

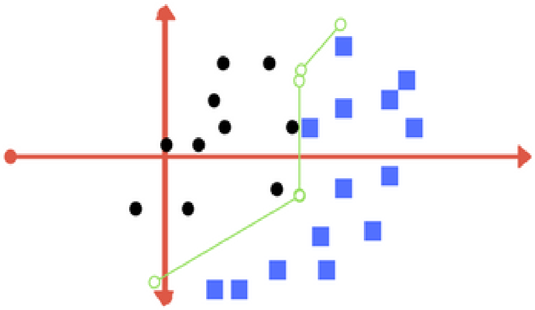

Fig. 6.(a) low regularization value

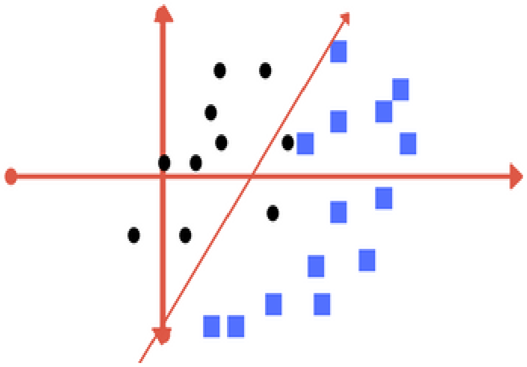

Fig. 6.(b) high regularization value

Gamma(γ)

The gamma parameter defines how far the influence of a single training example reaches, with low values meaning ‘far’ and high values meaning ‘close’. We can also say if gamma is low, points far away from plausible seperation line are considered in calculation for the seperation line but with high gamma the points close to plausible line are considered in calculation.

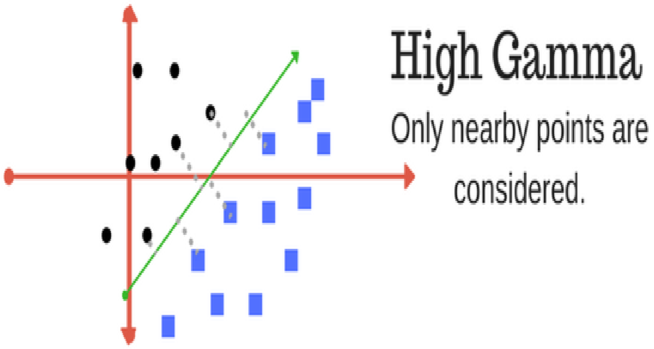

Fig. 7.High Gamma

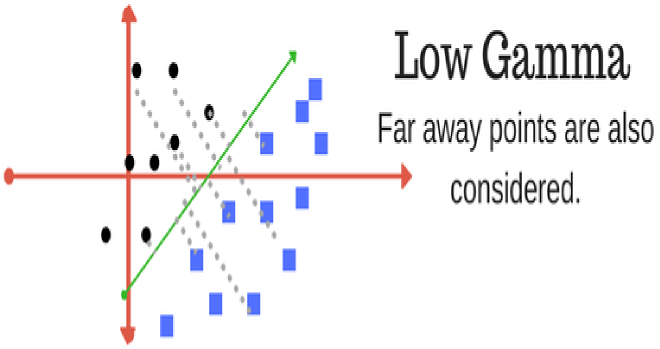

Fig. 8.Low Gamma

Margin

Margin is the most important characteristic of SVM classifier. SVM always tries to achieve a good margin. A margin is defined as the separation of line to the closest class points. A margin is considered good if the separation is larger for both the classes. We can see example of a good and a bad margin in *Fig. 9*. A good margin allows the points to be in their respective classes without crossing to other class.

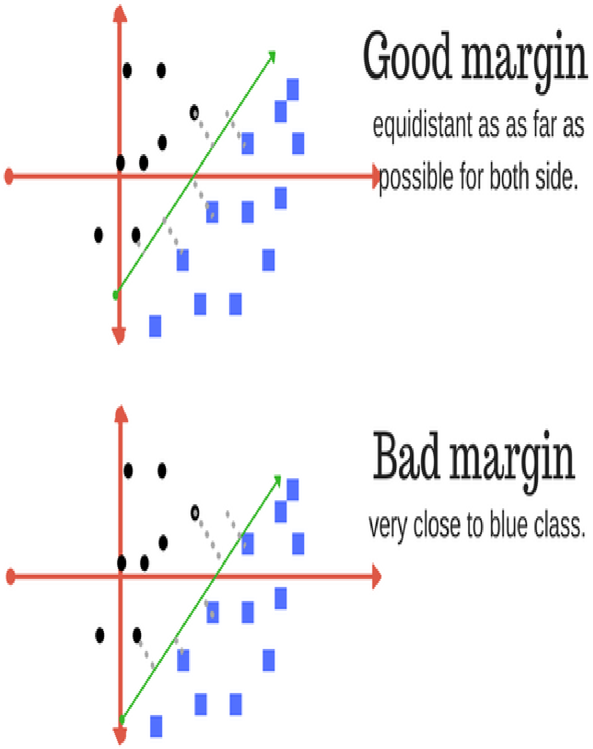
Good margin
equidistant as as far as possible for both side.
Bad margin
very close to blue class.

Fig.9. Good and Bad Margins

III. CONCLUSION

This paper is quite helpful in understanding the working behind SVM classifier. In the conventional method, support vector machine (SVM) algorithms are trained over pre-configured intranet/internet environments to find out an optimal classifier. If the data-sets are large, SVM models become complicated and costly.

REFERENCES

1. *Chih-Chung Chang and Chih-Jen Lin, LIBSVM: A Library for Support Vector Machines. Department of National Taiwan University, March 2013.*
2. *Sujay Raghavendra. N, Paresh Chandra Deka, Support vector machine applications in the field of hydrology: A review, Elsevier March 2014.*
3. *Chih-Wei Hsu et al., A Practical Guide to Support Vector Classification. Department of National Taiwan University, March 2013.*
4. *Fabian Pedregosa, Scikit-learn: Machine Learning in Python. The Journal of Machine Learning Research , October 2011.*
5. *Tom De Smedt , Walter Daelemans, Pattern for Python. The Journal of Machine Learning Research , January 2012.*
6. *Davide Albanese et al., mlpy: Machine Learning Python. arXiv, March 2012.*