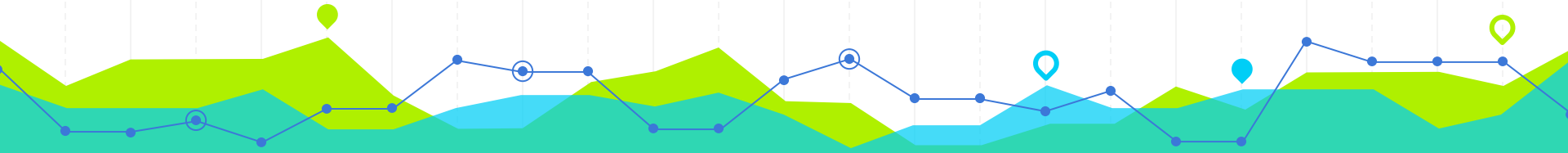
ANDREA VIVIANI



SUPPORT VECTOR MACHINE

Support Vector Machine (SVM)

It’s a powerful supervised machine learning algorithm that works best on smaller datasets but on complex ones. SVM can be used for both classification and regression problems.

Types of Support Vector Machine

Linear SVM

We have linear Support Vector Machine case when the data is perfectly linearly separable only then we can use Linear SVM

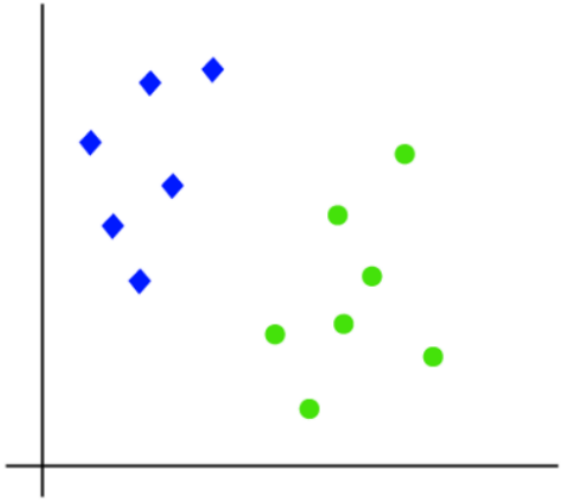
Non-Linear SVM

We have Non-Linear Support Vector Machine when the data is not linearly, that means when the data points cannot be separated into two different classes by using a straight line. In this case we can use Kernel tricks

How does SVM work?

1st step

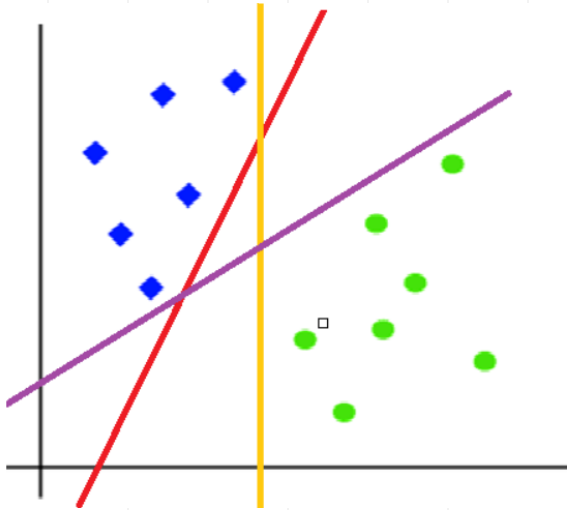
Suppose we have a dataset that has two classes (green and blue). We want to classify that the new data point as either blue or green.



How does SVM work?

2nd step

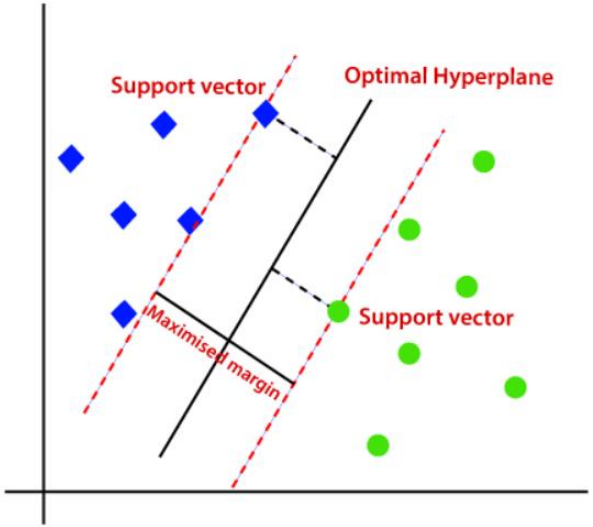
Since we are plotting the data points in a 2-dimensional graph we call this decision boundary a straight line but if we have more dimensions, we call this decision boundary a **“hyperplane”**



How does SVM work?

3rd step

The best hyperplane is that plane that has the maximum distance from both the classes. This is done by finding different hyperplanes which classify the labels in the best way.

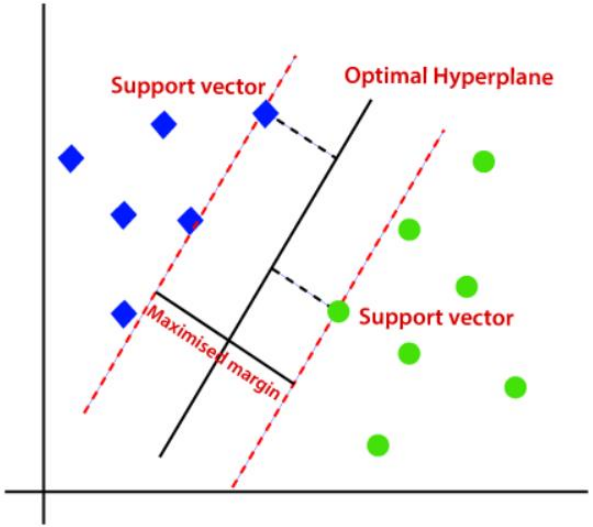


How does SVM work?

3rd step

Support Vectors: These are the points that are closest to the hyperplane.

Margin: it is the distance between the hyperplane and the observations closest to the hyperplane.



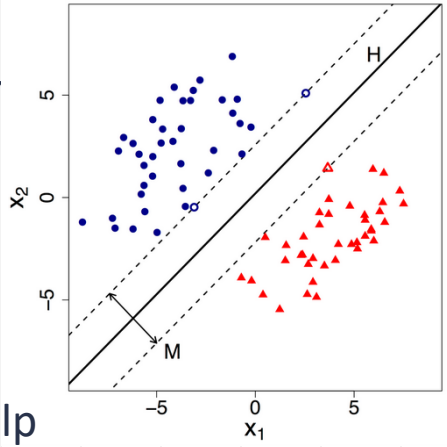
Difference between hard and soft SVM

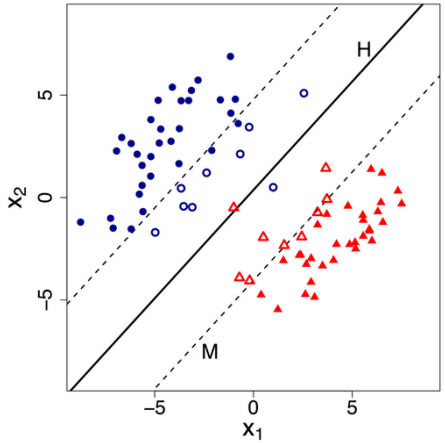
Hard margin

Does not allow any misclassification to happen.

Soft margin

Allows some misclassification to happen. Is implemented with the help of the **Regularization parameter (C).**





In this case HP1 (figure 1) is a Hard SVM and the HP2 (figure 2) is a Soft SVM.

Parameters of SVM

Regularization parameter (C)

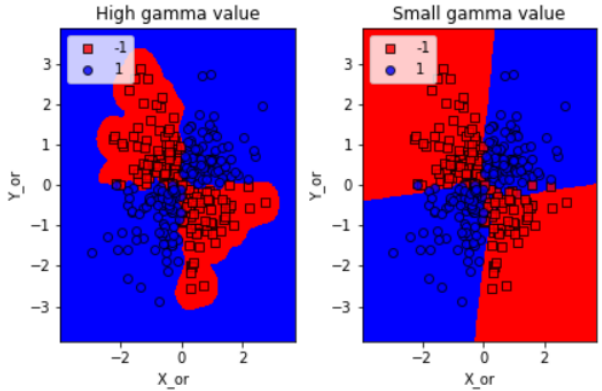
It tells us how much misclassification we want to avoid.

Large Gamma

Fewer data points will influence the decision boundary

Small Gamma

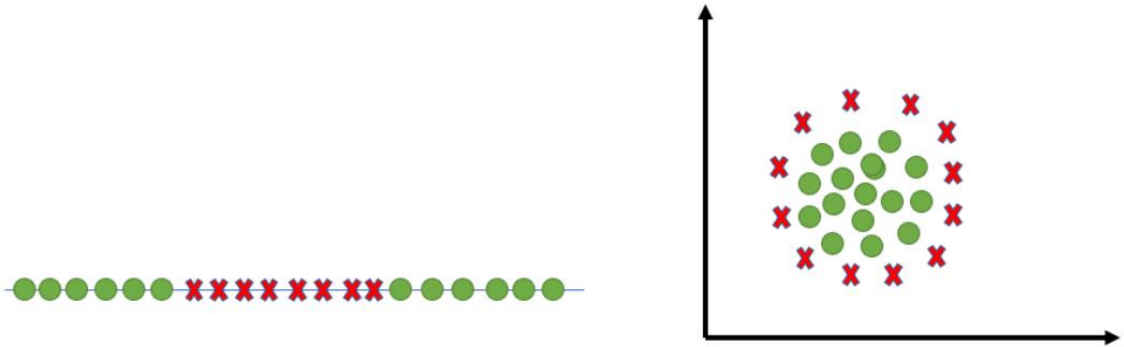
More data points will influence the decision boundary



Kernel Trick

1st step

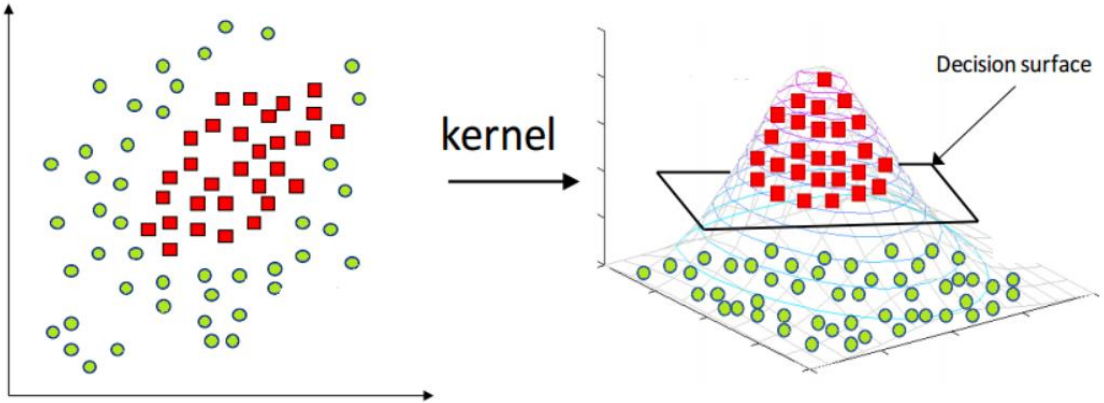
SVM deals with nonlinear data by transforming it into a higher dimension where it is linearly separable. We use “Kernel Trick” which makes it easier to classifies the points.



Kernel Trick

2nd step

Try converting this lower dimension space to a higher dimension space using some quadratic functions which will allow us to find a decision boundary that clearly divides the data points



Advantages of SVM

1. SVM works better when the data is Linear
2. It is more effective in high dimensions
3. With the help of the kernel trick, we can solve any complex problem
4. SVM is not sensitive to outliers
5. Can help us with Image classification

Disadvantages of SVM

1. Choosing a good kernel is not easy
2. It doesn’t show good results on a big dataset
3. The SVM hyperparameters are Cost -C and gamma. It is not that easy to fine-tune these hyper-parameters. It is hard to visualize their impact