Hands on Lab on Support Vector Machines

AGENDA OF THE VIDEO

- What is Support Vector Machines?
- A brief Overview of how does SVMs work?
- Linearly Separable vs non-linearly Separable
- Support Vectors vs Hyperplane
- SVMs as Margin Maximization Problem
- SVM algorithm
- What makes SVMs difficult to scale?
- SVMs for linearly Separable data in sklearn
- SVMs for non-linearly separable data in sklearn
- Task for you...

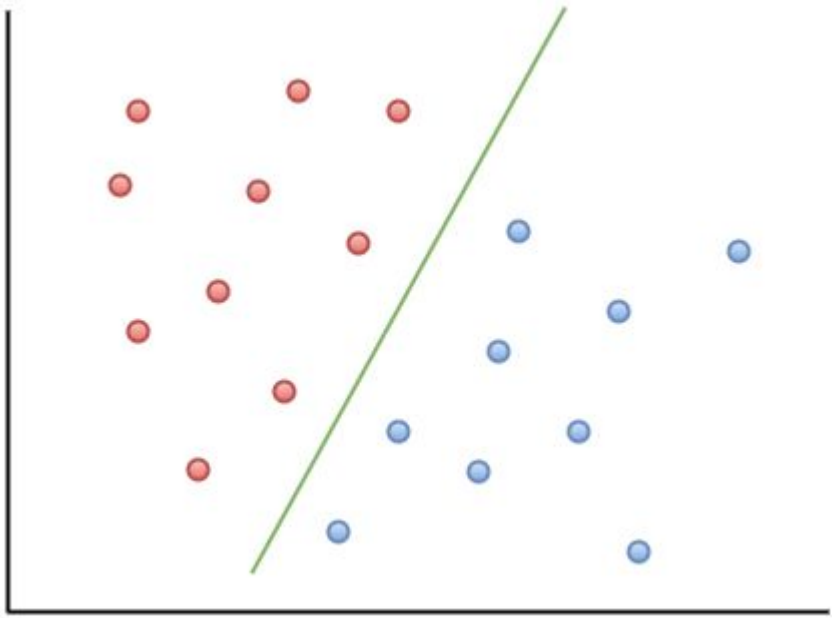
What is Support Vector Machines?

Support Vector Machines is one of the most earliest machine learning algorithms which is developed in early 60s but was never implemented due to lack of computational resources. By default, it is used for binary classification problems but can be easily extended.

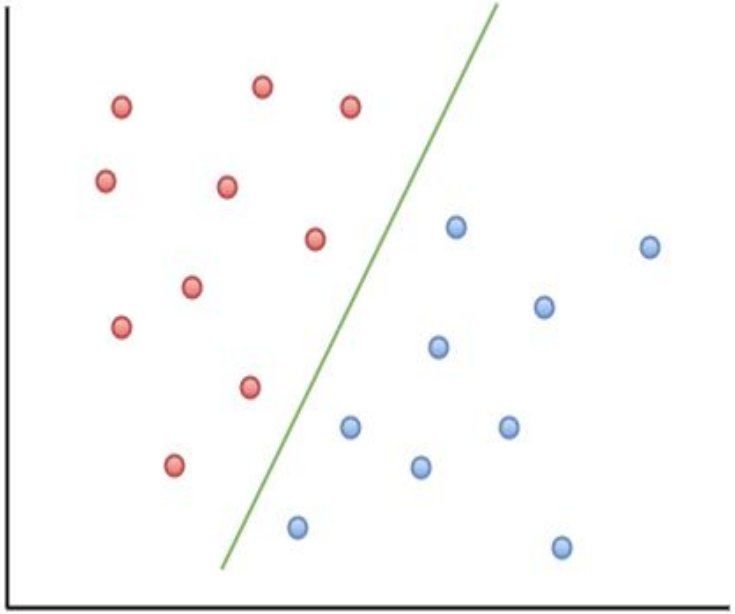
It is made up of three words. Support, Vectors and Machines.

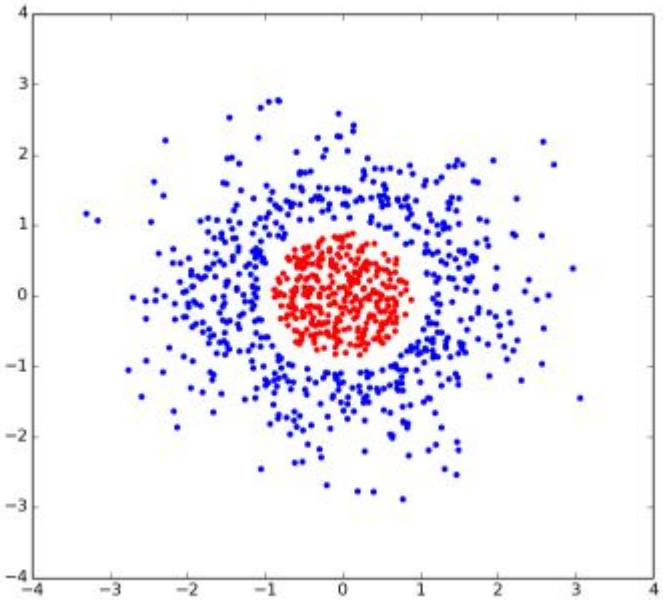
Support Vectors are those data points which supports the classification process, and Machines is just a word used to convey a system.

Brief Overview of SVMs

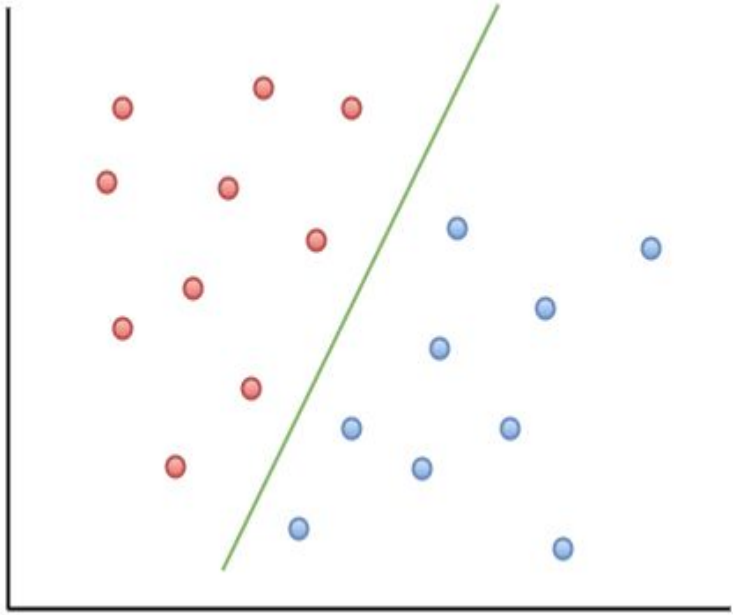


Linearly and nonlinearly Separable





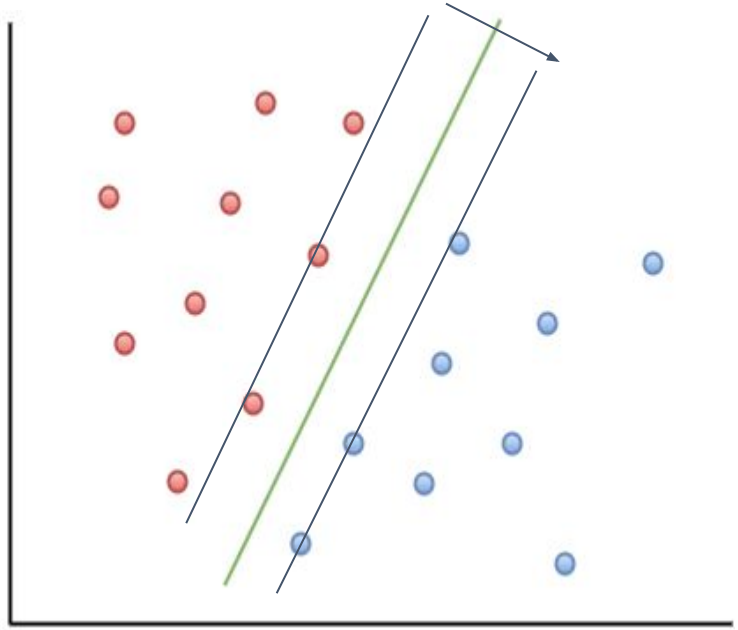
Support Vectors and Hyperplanes



Hyperplane is a fancy name for decision boundary.

Support vectors are the points that are closest to the decision boundary.

SVMs as a Margin Maximization Problem



To get the hyperplane in the exact between of these support vectors, we have maximize the so called margin between support vectors are the hyperplane.

SVM Algorithm

Equation of Hyperplane: **x.w + b = 0**
Equation of Support Vectors: **x.w + b = 1 and x.w + b = -1**

X is our input vector. What is the Decision Rule? sign(x.w + b)

But how do we get w and b? So, we need to estimate it? It is an Optimization Problem..

Maximize: the margin with respect to some constraints.

It turns out the optimization problem is a quadratic problem and can get a little difficult to understand. So, we use sklearn for its implementation.

What makes SVMs difficult to scale?

For really huge dataset, the training time is really huge. For such datasets, we need supercomputers with massive computational powers.

But once the SVMs are trained, we no longer need to store the data in memory, we just need to remember the weights and bias vectors, these are all what we need to make a new prediction.

SVM for linearly separable data in sklearn

X1	X2	Class
15	50	Pos
20	25	Pos
17	75	Pos
15	25	Pos
19	51	Pos
10	100	Neg
8	75	Neg
5	25	Neg
4	50	Neg
6	60	Neg
5	90	Neg

SVM for non linearly separable data in sklearn

X1	X2	Class
11	25	Pos
9	26	Pos
4	49	Pos
4	60	Pos
3	75	Pos
11	80	Pos
12	85	Pos
16	85	Pos
17	76	Pos
19	60	Pos
18	49	Pos
17	26	Pos
16	27	Pos

X1	X2	Class
11	52	Neg
12	60	Neg
13	60	Neg
13	52	Neg
15	40	Neg
15	52	Neg
15	75	Neg
16	76	Neg
11	70	Neg
14	51	Neg

Task For You...

This task encourages to participate in a data science competition for some classification problem and apply SVM on that dataset.

Now, your work is not finished just by applying the SVC class in sklearn. Rather, you need to configure the best hyperparameters of SVC by trying various different combinations of configurations. This way you will understand the power of hit and trial in Model building and in improving the model accuracy.

Thank You.