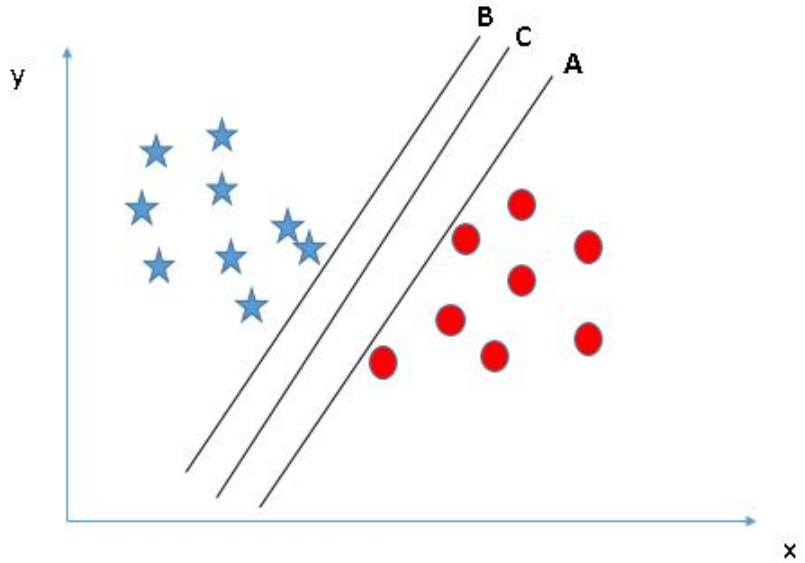
Support Vector Machines
(And Why We Like Them)

Support Vector Algorithm Creators

Vladimir Vapnik

Alexey Chervonenkis

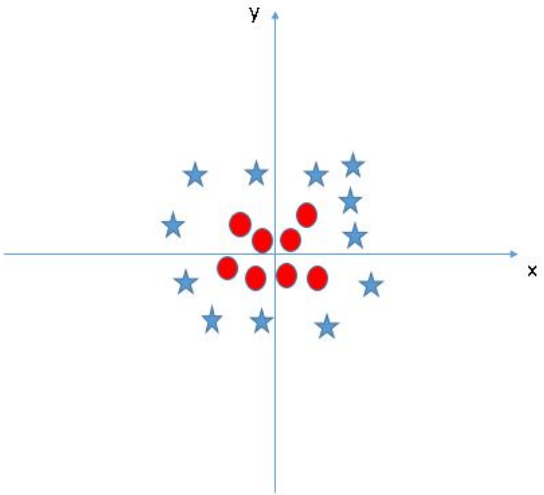
Identifying the Right Hyperplane



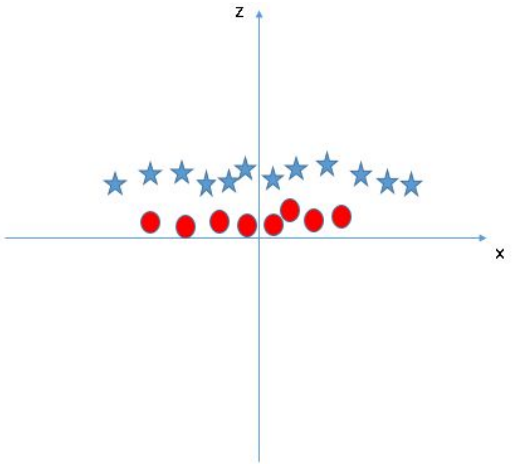
Here, you can see that the margin for hyper-plane C is high as compared to both A and B. Hence, we name the right hyper-plane as C. Another lightning reason for selecting the hyper-plane with higher margin is robustness. If we select a hyper-plane having low margin then there is high chance of miss-classification.

Non-Linear Hyperplanes

What to do?



SVM can solve this problem. Easily! It solves this problem by introducing additional feature. Here, we will add a new feature **$z=x^2+y^2$.** Now, let's plot the data points on axis x and z:



Kernel trick: functions which take low dimensional input space and transform it to a higher dimensional space

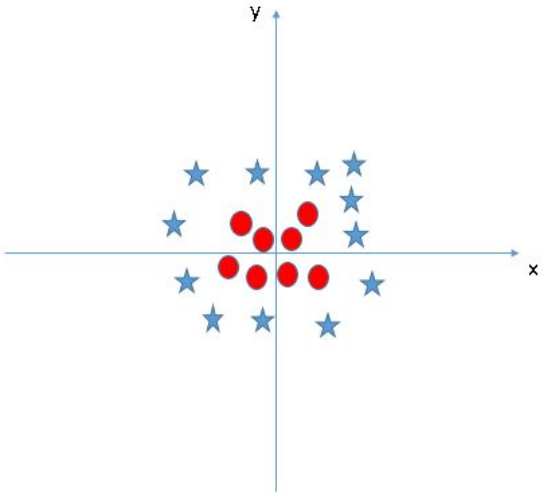
Visual Demonstration

[Youtube Link]

Remember:

$$z=x^2+y^2$$

If X and Y both = 2, what does Z equal?



Advantages Vs. Disadvantages

Pros?

- It works really well with clear margin of separation
- It is effective in high dimensional spaces.
- It is effective in cases where number of dimensions is greater than the number of samples.
- It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.

Cons?

- It doesn’t perform well, when we have large data set because the required training time is higher
- It also doesn’t perform very well, when the data set has more noise i.e. target classes are overlapping
- SVM doesn’t directly provide probability estimates, these are calculated using an expensive five-fold cross-validation. It is related SVC method of Python scikit-learn library.

Common Uses

-What are support vector machines often used for?

- Text Classification
- Facial Recognition
- Bioinformatics