Support Vector Machine

Tuesday, April 16, 2019 2:11 PM

Admin

Class 8
June 7 2:10

Office Hours
Thursdays 2-3

From Last Week:

Cross Validation
Always expect purity and completeness to act inversely

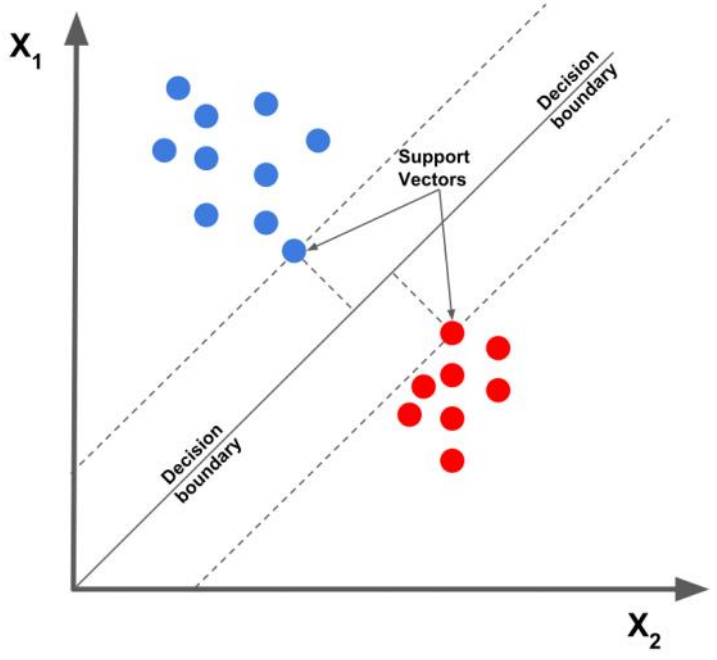
Classification

Support Vector Machine (SVM)
Generalized Technique
Build the model (decision function), then use for classification
More computationally efficient.

$$x_{\{1,\ldots,n\}} = [\leftarrow D \rightarrow] \quad y = \begin{matrix}\uparrow \\ N \\ \downarrow\end{matrix}$$

SVM - finds the hyperplane that maximally separates the classes.

Don't need training data once you have the decision function.



Decision boundary also called the decision function or the hyperplane

Types of SVM: Linear, Poly, RBF

How do support vectors cope when the separation isn't this clean?

Tuning parameter: C, how many misclassifications are allowed.

Max(m)

Subject to

$$\frac{1}{||\beta||} y_i(\beta_0 + \beta^T x_i) \geq m \, \forall_i$$

$\beta_o : intercept\ of\ hyperplane\ (line\ in\ 2D)$

$\beta : slope$

$x_i : a\ point\ y_i : a\ label$

$m : a\ margin$

An analog to this is $y = mx + b$ if you let $\frac{1}{||\beta||} = m$ this simplifies greatly.

$$y_i(\beta_0 + \beta^T x_i) \geq 1$$

$$y_i(\beta_0 + \beta^T x_i) \geq 1 - \xi_i \forall_i$$
$$\xi_i \geq 0$$
$\sum_i \xi_i \leq C$ The max amount of fudge allowed. C is an integer

Cross validation is how you derive an optimal value of C.

svc_linear.fit(xdata, ylables)

Levels is how far from the margin of your decision function you are.

Increasing the number of classes you have will give you more decision functions.
You'll need to reshape svc_linear.decision_function

For 2 classes you get 1 decision function.
For 3 classes you get 3 decision functions.

For real data that does not separate out cleanly you'll want to move away from linear.

The default is the radial basis functions or RBF, these are also referred to as kernels.

$\gamma : size\ of\ the\ radius\ of\ the\ basis\ function$

$c : contamination\ allowed$

People will multiply completeness and purity together to find the optimal solution.
(Gamma in this case.)