

Classification

Recall that classification means we have some blue points and some red points, and we want to find a boundary that separates them.

But now we want the boundary to be as far away from the points as possible. To do this we add two boundaries (shown in the video as equidistant parallel lines to the main line), and try to maximize the distance between the two.

There are two ways to measure the model:

- By how many points it mis-classifies
- By how wide the margin is

We can 'punish' the misclassified points, which means make them part of our classification error. We also do not want any points in the margin, so we would 'punish' those points as well and include them in the classification error.

Lastly, we want the margin to be as wide as possible and will see how to calculate a margin error a little later. Our new error for this algorithm is going to be a classification error + a margin error. Minimizing this error is what's going to give us the algorithm for Support Vector Machines.

Quiz Question

Which of the following is true about SVM?

- a. We want the boundary to be as close as possible to the points.
- b. We want the boundary to be as far away from the points as possible.
- c. The boundary should intersect at least 1 of the points

Quiz Question

Which of the following is not true about SVM?

- a. The error for the SVM algorithm is classification error + margin error
- b. The error for the SVM algorithm is classification error - margin error
- c. minimizing the (classification+ margin) error is the goal