

Chapter 5

1. What is the fundamental idea behind Support Vector Machines?
 - a. The main idea behind SVMs is to have the largest boundary between the two classes and the training set in the SVM.
2. What is a support vector?
 - a. A support vector is just an instance in the boundary between the classes.
3. Why is it important to scale the inputs when using SVMs?
 - a. If the inputs are too small SVMs may ignore or toss out smaller features, which you may not want.
4. Can an SVM classifier output a confidence score when it classifies an instance? What about a probability?
 - a. Yes, an SVM classifier can output values that can be used as a confidence score. For example, the distance between the test set and the decision boundary.
 - b. No, a probability cannot be output.
5. Should you use the primal or the dual form of the SVM problem to train a model on a training set with millions of instances and hundreds of features?
 - a. In this scenario, the primal form would perform better because the dual forms computational complexity is roughly between m^2 and m^3 while the primal form is m . Where m is the training instances.
6. Say you've trained an SVM classifier with an RBF kernel, but it seems to underfit the training set. Should you increase or decrease γ (gamma)? What about C ?
 - a. In this scenario you would likely need to increase both gamma and C .
7. How should you set the QP parameters (H , F , A , and b) to solve the soft margin linear SVM classifier problem using an off-the-shelf QP solver?
 - a. I am not sure how to solve this problem.