Quiz #7 – Solutions

1. Match the terms to the appropriate descriptions:

a) Maximal margin classifier
b) Support vector classifier
c) Support vector machine

- Linear classifier for data that cannot be perfectly separated with a hyperplane (B)

- Simple classifier for data that can be perfectly separated with a hyperplane (A)

- Non-linear classifier for data that cannot be perfectly separated with a hyperplane (C)

2. Select the true statements from the following statements about maximal margin classifiers in the context of a two-class problem. (Select all that apply.)

a) Maximal margin classifiers only work for training data that can be perfectly separated using a hyperplane.
b) The maximal margin hyperplane is the hyperplane that has the closest maximum distance to the training observations.
c) We can be more confident about class assignments for observations close to the maximal margin hyperplane than for those far from it.
d) All training observations in one class are on one side of the maximal margin hyperplane and all training observations in the other class are on the other side.

Solution: A, D

3. Select the true statements from the following statements about support vector classifiers in the context of a two-class problem. (Select all that apply.)

a) Support vector classifiers allow some training observations to be misclassified in order to improve classification for the remaining observations.
b) The tuning parameter for support vector classifiers controls the number of training observations that are allowed to be misclassified.
c) When only a few training observations are allowed to be misclassified the resulting classifier has high bias and low variance; by contrast, when many training observations are allowed to be misclassified the resulting classifier has low bias and high variance.
d) All the training observations determine the fitted support vector classifier.

Solution: A, B

4. Select the true statements from the following statements about support vector machines in the context of a two-class problem. (Select all that apply.)

a) A support vector machine is constrained to have a linear decision boundary between the classes.
b) The classifier for a support vector machine can be characterized as a linear combination of kernels, where the kernels are functions of the observation in question and each support vector.
c) A support vector machine that uses linear kernels is simply a support vector classifier.
d) A support machine that uses polynomial or radial kernels leads to a more flexible, nonlinear decision boundary between the classes.

Solution: B, C, D

5. What is one advantage of using kernels with support vector machines rather than simply applying a support vector classifier to an enlarged feature space (containing squared predictor terms, interactions, etc.)?

a) Using kernels with support vector machines allows non-linear decision boundaries whereas applying a support vector classifier to an enlarged feature space constrains the decision boundary to be linear.
b) Using kernels with support vector machines always results in better predictive performance on the test set than applying a support vector classifier to an enlarged feature space.
c) Using kernels with support vector machines is more computationally efficient than applying a support vector classifier to an enlarged feature space.

Solution: C

6. True or false? It is possible to extend SVMs for more than two classes?

Solution: True