The one-vs.-one multi-class technique can potentially lead to a large number of constructed classifiers and hence can result in slower processing time. Conversely, the classifiers are more robust to class imbalances when training each classifier.

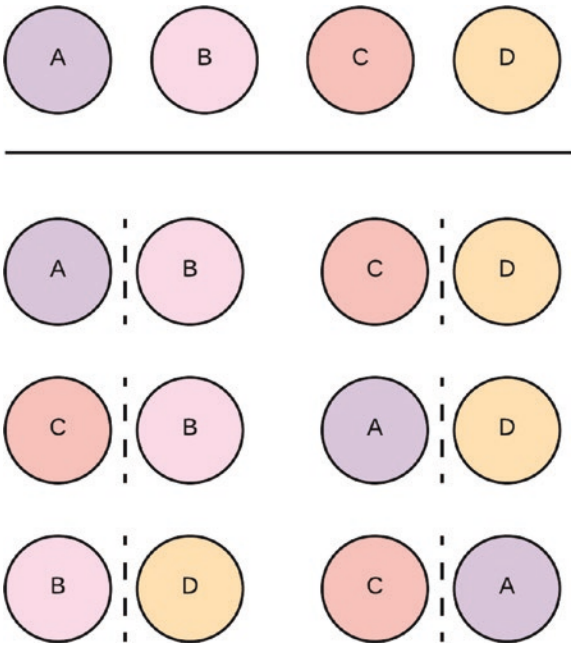

Figure 22-8. *Suppose we have four classes in the dataset labeled A to D, this will result in six different classifiers*

One-vs.-All (OVA)

The one-vs.-all method for fitting an SVM to a multi-classification problem where the number of classes k is greater than 2 consists of fitting each k class against the remaining k – 1 classes. Suppose we have ten classes, each of the classes will be classified against the remaining nine classes. This example is illustrated with four classes in Figure 22-9.

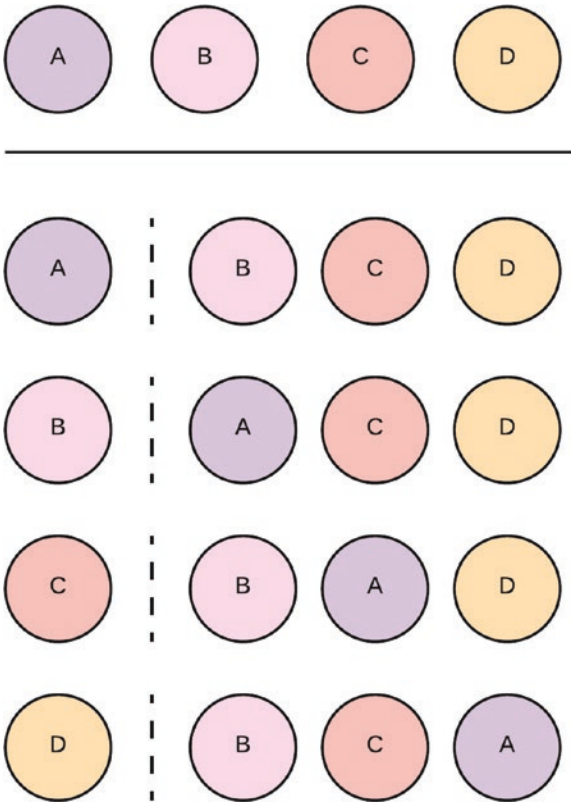

Figure 22-9. *Given four classes in a dataset, we construct four classifiers, with each class fitted against the rest*

The classifiers are evaluated by comparing a test example to each fitted classifier. The classifier for which the margin of the hyperplane is the largest is chosen as the predicted classification target because the classifier margin size is indicative of high confidence of class membership.

The Kernel Trick: Fitting Non-linear Decision Boundaries

Non-linear datasets occur more often than not in real world scenarios.

Technically speaking, the name support vector machine is when a support vector classifier is used with a non-linear kernel to learn non-linear decision boundaries.