II SUPPORT VECTOR MACHINES

The algorithms for constructing the separating hyperplane considered above will be utilized for developing a battery of programs for pattern recognition.

V. N. Vapnik [560, p. 364]

Now that we have the necessary concepts and tools, we move on to the class of Support Vector (SV) algorithms. SV algorithms are commonly considered the first practicable spin-off of statistical learning theory. We described the basic ideas of Support Vector machines (SVMs) in Chapter 1. It is now time for a much more detailed discussion and description of SVMs, starting with the case of *pattern recognition* (Chapter 7), which was historically the first to be developed.

Following this, we move on to a problem that can actually be considered as being even simpler than pattern recognition. In pattern recognition, we try to distinguish between patterns of at least two classes; in single-class classification (Chapter 8), however, there is only one class. In the latter case, which belongs to the realm of unsupervised learning, we try to learn a model of the data which describes, in a weak sense, what the training data looks like. This model can then be used to assess the "typicality" or novelty of previously unseen patterns, a task which is rather useful in a number of application domains.

Chapter 9 introduces SV algorithms for regression estimation. These retain most of the properties of the other SV algorithms, with the exception that in the regression case, the choice of the loss function, as described in Chapter 3, becomes a more interesting issue.

After this, we give details on how to implement the various types of SV algorithms (Chapter 10), and we describe some methods for incorporating prior knowledge about invariances of a given problem into SVMs (Chapter 11).

We conclude this part of the book by revisiting statistical learning theory, this time with a much stronger emphasis on elements that are specific to SVMs and kernel methods (Chapter 12).