
Book review

Richard O. Duda, Peter E. Hart and David G. Stork, *Pattern Classification*, 2nd edn, Wiley, 2001, 654 pp., £78.95.

This book not only describes pattern classification methods, but also includes new and recently developed techniques. There are nine main chapters and an appendix.

Chapter two covers Bayesian decision theory, a fundamental approach for statistical pattern classification. Several criteria for minimising the error rate are described. The classification is discussed for both two- and multi-categories, especially with normal distributions.

Chapters three and four both contribute to an understanding of supervised learning. Chapter three describes the maximum-likelihood and Bayesian parameter estimation, where the underlying pdf functions are assumed to be known in advance. However some parameters in the pdfs are unknown and will be estimated by the training data. The chapter also discusses topics such as sufficient statistics, problems of dimensionality, component analysis and discriminants, expectation-maximisation and HMM models. Chapter four describes non-parametric techniques that can be used for arbitrary distributions without any assumption of the underlying forms of pdf function. Techniques described in this chapter mainly include the Parzen window approach, k-nearest neighbour and fuzzy classification.

Chapters five and six contribute to approaches that training samples can be used directly to estimate the parameters in a classifier without requiring a knowledge of the underlying pdfs. Chapter five describes various linear discriminant functions, assuming the forms of discriminant functions are linear and known. In such cases, the training samples are used to estimate the parameters in a classifier. Chapter six describes nonlinear multilayer neural networks, where patterns are not linearly separable, hence the decision boundaries cannot be described by a hyperplane. Feed forward operation for classification and backpropagation operation for supervised training are described. The association of backpropagation training to fitting the Bayesian discriminant functions is also described. Other networks, such as radial basis functions, matched filters, convolutional and recurrent networks, are also included.

Chapter seven describes stochastic methods when models become more complex therefore parameter estimation becomes difficult due to multiple maxima (minima). Two general classes of search methods, Boltzmann learning and genetic algorithms, are described.

Chapter eight describes non-metric methods by using a list of attributes for describing the patterns. Several methods for growing decision trees from the training data are described.

Chapter nine describes algorithm-independent machine learning. Since no pattern classification method is inherently superior to any other, several ways are explored to quantify and adjust the match between a learning algorithm and the problem it addresses. The chapter describes how under certain assumptions, one can estimate the accuracy of a classifier and also compare different classifiers.

Chapter ten describes methods for unsupervised learning and clustering. This includes maximum-likelihood estimates using data from a mixture density, k-means (or c-means) and fuzzy k-means clustering, unsupervised Bayesian learning, hierarchical clustering and component analysis.

Overall I think this is an excellent book. It gives clear insights into the methods described, and contains in addition many figures and algorithms which enhance understanding of the subject matter. The book is suitable as a textbook and is a valuable reference book for researchers and engineers.

Irene Gu *Chalmers University of Technology*