Vapnik, V. N.
The nature of statistical learning theory. 2nd ed. (English) ‖Zbl 0934.62009‖
Statistics for Engineering and Information Science. New York, NY: Springer. xix, 314 p. (2000).

[For the review of the first edition from 1995 see Zbl 0833.62008.]

The goal of this book is to describe the nature of statistical learning theory. It is not a survey of the standard learning theory but it is intended to show how abstract reasoning implies new algorithms.

Ch. 1 is to promote the setting of the learning problem. It presents the classical paradigm of solving learning problems, nonparametric methods of density estimation, and the main principles for solving problems with restricted information. A general model of minimizing the risk functional from empirical data is introduced. Ch. 2 describes the conceptual theory of learning processes: Key theorems of learning theory, theory of nonfalsifiability, and concepts that allow the construction of necessary and sufficient conditions for the consistency of the learning processes. Ch. 3 gives the nonasymptotic theory of distribution-independent and dependent bounds on the convergence rate of the learning processes. Ch. 4 deals with the theory for controlling the generalization ability of learning machines which is founded on constructing an inductive principle for minimizing the risc functional for a small sample size of training instances. (Examples of structures for neural networks, local function approximation, MDL and SRM principles, capacity control).

Ch. 5 describes learning algorithms for pattern recognition with generalizations for the regression estimation problem. Along with classical neural networks the minimization of the empirical risc with fixed confidence intervals is implemented. The main topics promote experiments with support vector machines (SVM). Ch. 6 introduces a new type of loss functions, the so-called ε-insensitive loss function. The results obtained for the pattern recognition problem are generalized to the problem of estimating real-valued regressions. Ch. 7 shows a new approach to the problems above combining the ideas of the theory of ill-posed problems, classical nonparametric statistics, and statistical learning theory to obtain a new type of algorithms. Ch. 8 introduces a new principle for minimizing the expected risk called the vicinal risk minimization using the SVM technique.

The book is not easy-reading but written in a concise style. It must be recommended to scientists of statistics, mathematics, physics, and computer science.

Reviewer: Roland Fahrion (Heidelberg)

MSC:

62B10	Statistical aspects of information-theoretic topics
62-02	Research exposition (monographs, survey articles) pertaining to statistics
68T05	Learning and adaptive systems in artificial intelligence

Cited in **2** Reviews
Cited in **542** Documents

Keywords:

nonparametric statistics; vicinal risk minimization; new algorithms; pattern recognition; support vector machines