

Computer Science and Data Analysis Series

Statistical Learning and Data Science

Edited by

Mireille Gettler Summa, Léon Bottou,
Bernard Goldfarb, Fionn Murtagh,
Catherine Pardoux, and Myriam Touati



CRC Press

Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

A CHAPMAN & HALL BOOK

Contents

Preface	vii
Contributors	xiii
I Statistical and Machine Learning	1
1 Mining on Social Networks	3
<i>Benjamin Chapus, Françoise Fogelman Soulié, Erik Marcadé, and Julien Sauvage</i>	
2 Large-Scale Machine Learning with Stochastic Gradient Descent	17
<i>Léon Bottou</i>	
3 Fast Optimization Algorithms for Solving SVM+	27
<i>Dmitry Pechyony and Vladimir Vapnik</i>	
4 Conformal Predictors in Semisupervised Case	43
<i>Dmitry Adamskiy, Iliia Nouretdinov, and Alexander Gammerman</i>	
5 Some Properties of Infinite VC-Dimension Systems	53
<i>Alexey Chervonenkis</i>	
II Data Science, Foundations, and Applications	61
6 Choriogenesis: the Dynamical Genesis of Space and Its Dimensions, Controlled by Correspondence Analysis	63
<i>Jean-Paul Benzécri</i>	
7 Geometric Data Analysis in a Social Science Research Program: The Case of Bourdieu's Sociology	77
<i>Frédéric Lebaron</i>	
8 Semantics from Narrative: State of the Art and Future Prospects	91
<i>Fionn Murtagh, Adam Ganz, and Joe Reddington</i>	
9 Measuring Classifier Performance: On the Incoherence of the Area under the ROC Curve and What to Do about It	103
<i>David J. Hand</i>	

10 A Clustering Approach to Monitor System Working: An Application to Electric Power Production	113
<i>Alzennyr Da Silva, Yves Lechevallier, and Redouane Seraoui</i>	
11 Introduction to Molecular Phylogeny	125
<i>Mahendra Mariadassou and Avner Bar-Hen</i>	
12 Bayesian Analysis of Structural Equation Models Using Parameter Expansion	135
<i>S�everine Demeyer, Jean-Louis Foulley, Nicolas Fischer, and Gilbert Saporta</i>	
III Complex Data	147
13 Clustering Trajectories of a Three-Way Longitudinal Dataset	149
<i>Mireille Gettler Summa, Bernard Goldfarb, and Maurizio Vichi</i>	
14 Trees with Soft Nodes: A New Approach to the Construction of Prediction Trees from Data	159
<i>Antonio Ciampi</i>	
15 Synthesis of Objects	171
<i>Myriam Touati, Mohamed Djedour, and Edwin Diday</i>	
16 Functional Data Analysis: An Interdisciplinary Statistical Topic	189
<i>Laurent Delsol, Fr�ed�eric Ferraty, and Adela Mart�inez Calvo</i>	
17 Methodological Richness of Functional Data Analysis	197
<i>Wenceslao Gonz�alez Manteiga and Philippe Vieu</i>	
Bibliography	205
Index	225