Georgios Paliouras Vangelis Karkaletsis
Constantine D. Spyropoulos (Eds.)

Machine Learning and Its Applications

Advanced Lectures

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saabrücken, Germany

Volume Editors

Georgios Paliouras
Vangelis Karkaletsis
Constantine D. Spyropoulos
National Centre for Scientific Research "Demokritos"
Institute of Informatics and Telecommunications
P.O. Box 60228, Ag. Paraskevi, 15310 Athens, Greece
E-mail: {paliourg,vangelis,costass}@iit.demokritos.gr

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Machine learning and its applications : advanced lectures / Georgios Paliouras ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2001
(Lecture notes in computer science ; 2049 : Lecture notes in artificial intelligence)
ISBN 3-540-42490-3

CR Subject Classification (1998): I.2, 4.3.3, H.2.8, H.5.2, J.1, F.4.1

ISBN 3-540-42490-3 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Stefan Sossna
Printed on acid-free paper SPIN: 10781488 06/3142 5 4 3 2 1 0

Preface

In the last few years machine learning has made its way into the areas of administration, commerce, and industry, in an impressive way. Data mining is perhaps the most widely known demonstration of this phenomenon, complemented by less publicized applications of machine learning, such as adaptive systems in various industrial settings, financial prediction, medical diagnosis, and the construction of user profiles for WWW-browsers. This transfer of machine learning from the research labs to the "real world" has caused increased interest in learning techniques, dictating further effort in informing people from other disciplines about the state of the art in machine learning and its uses.

The objective of this book is to provide the reader with sufficient information about the current capabilities of machine learning methods, as well as ideas about how one could make use of these methods to solve real-world problems. The book is based primarily on the material that was presented in the Advanced Course in Artificial Intelligence (ACAI '99), which took place in Chania, Greece and was attended by research students, professionals, and researchers. However, the book goes beyond the material covered in the course, in that it contains several position papers on open research issues of machine learning.

The book is structured in a way that reflects its objective of educating the reader on how machine learning works, what the open issues are, and how it can be used. It is divided into two parts: methods and applications.

The first part consists of 10 chapters covering to a large extent the field of machine learning, from symbolic concept learning and conceptual clustering to case-based reasoning, neural networks, and genetic algorithms. The research issues addressed include the relationship of machine learning to knowledge discovery in databases, the handling of noisy data, and the modification of the learning problem through function decomposition. This part of the book concludes with two chapters examining the basic principles of learning methods. The first of the two chapters examines the approaches to selecting the appropriate method for a particular problem or modifying the problem representation to suit a learning method. In contrast, the last chapter of the section reviews the approaches to integrating different machine learning methods, in order to handle difficult learning tasks.

The second part of the book exposes the reader to innovative applications of machine learning. This part consists of 11 chapters, presenting a range of applications, from data mining in finance, marketing, and economics to learning in human language technology and user modeling. Most of these chapters are based on the work that was presented during the workshops of ACAI '99. Each such chapter describes the types of problem that have been approached with the use of machine learning in a particular domain and gives an overview of the work in this area, as presented at the relevant workshop.

We hope that the combination of theoretical and empirical knowledge in this book will be of use to the reader who is interested in entering this exciting research field and using mature machine learning techniques to solve real-world problems. The editors of the book would like to thank the distinguished authors for their willingness and cooperation in making this special volume a reality.

May 2001

Georgios Paliouras
Vangelis Karkaletsis
Constantine D. Spyropoulos

Table of Contents

Methods

Comparing Machine Learning and Knowledge Discovery in DataBases: An Application to Knowledge Discovery in Texts 1
Yves Kodratoff

Learning Patterns in Noisy Data: The AQ Approach 22
Ryszard S. Michalski and Kenneth A. Kaufman

Unsupervised Learning of Probabilistic Concept Hierarchies 39
Wayne Iba and Pat Langley

Function Decomposition in Machine Learning 71
Blaž Zupan, Ivan Bratko, Marko Bohanec, and Janez Demšar

How to Upgrade Propositional Learners to First Order Logic: A Case Study 102
Wim Van Laer and Luc De Raedt

Case-Based Reasoning 127
Ramon Lopez de Mantaras

Genetic Algorithms in Machine Learning 146
Jonathan Shapiro

Pattern Recognition and Neural Networks 169
Sergios Theodoridis and Konstantinos Koutroumbas

Model Class Selection and Construction: Beyond the Procrustean Approach to Machine Learning Applications 196
Maarten van Someren

Integrated Architectures for Machine Learning 218
Lorenza Saitta

Applications

The Computational Support of Scientific Discovery 230
Pat Langley

Support Vector Machines: Theory and Applications 249
Theodoros Evgeniou and Massimiliano Pontil

Pre- and Post-processing in Machine Learning and Data Mining 258
Ivan Bruha

Machine Learning in Human Language Technology 267
Nikos D. Fakotakis and Kyriakos N. Sgarbas

Machine Learning for Intelligent Information Access 274
Grigoris Karakoulas and Giovanni Semeraro

Machine Learning and Intelligent Agents 281
Themis Panayiotopoulos and Nick Z. Zacharis

Machine Learning in User Modeling 286
Christos Papatheodorou

Data Mining in Economics, Finance, and Marketing 295
Hans C. Jessen and Georgios Paliouras

Machine Learning in Medical Applications 300
George D. Magoulas and Andriana Prentza

Machine Learning Applications to Power Systems 308
Nikolaos Hatziargyriou

Intelligent Techniques for Spatio-Temporal Data Analysis in Environmental Applications 318
Nikolaos Vassilas, Elias Kalapanidas, Nikolaos Avouris, and Stavros Perantonis

Author Index **325**