

1. Record Nr.	UNINA9910548277503321
Autore	Brazdil Pavel
Titolo	Metalearning : Applications to Automated Machine Learning and Data Mining
Pubbl/distr/stampa	Cham, : Springer Nature, 2022 Cham, : Springer International Publishing AG, , 2022 ©2022
ISBN	3-030-67024-4
Edizione	[2nd ed.]
Descrizione fisica	1 online resource (349 pages)
Collana	Cognitive Technologies
Classificazione	COM004000COM021030
Altri autori (Persone)	van RijnJan N SoaresCarlos VanschorenJoaquin
Disciplina	006.31
Soggetti	Artificial intelligence Data mining Machine learning Aprendentatge automàtic Mineria de dades Llibres electrònics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Contents -- Part I Basic Concepts and Architecture -- 1 Introduction -- 1.1 Organization of the Book -- 1.2 Basic Concepts and Architecture (Part I) -- 1.3 Advanced Techniques and Methods (Part II) -- 1.4 Repositories of Experimental Results (Part III) -- References -- 2 Metalearning Approaches for Algorithm Selection I (Exploiting Rankings) -- 2.1 Introduction -- 2.2 Different Forms of Recommendation -- 2.3 Ranking Models for Algorithm Selection -- 2.4 Using a Combined Measure of Accuracy and Runtime -- 2.5 Extensions and Other Approaches -- References -- 3 Evaluating Recommendations of Metalearning/AutoML Systems -- 3.1 Introduction -- 3.2 Methodology for Evaluating Base-Level Algorithms -- 3.3 Normalization of Performance for Base-Level Algorithms -- 3.4 Methodology for Evaluating Metalearning and AutoML Systems -- 3.5

Evaluating Recommendations by Correlation -- 3.6 Evaluating the Effects of Recommendations -- 3.7 Some Useful Measures -- References -- 4 Dataset Characteristics (Metafeatures) -- 4.1 Introduction -- 4.2 Data Characterization Used in Classification Tasks -- 4.3 Data Characterization Used in Regression Tasks -- 4.4 Data Characterization Used in Time Series Tasks -- 4.5 Data Characterization Used in Clustering Tasks -- 4.6 Deriving New Features from the Basic Set -- 4.7 Selection of Metafeatures -- 4.8 Algorithm-Specific Characterization and Representation Issues -- 4.9 Establishing Similarity Between Datasets -- References -- 5 Metalearning Approaches for Algorithm Selection II -- 5.1 Introduction -- 5.2 Using Regression Models in Metalearning Systems -- 5.3 Using Classification at Meta-level for the Prediction of Applicability -- 5.4 Methods Based on Pairwise Comparisons -- 5.5 Pairwise Approach for a Set of Algorithms -- 5.6 Iterative Approach of Conducting Pairwise Tests -- 5.7 Using ART Trees and Forests.

5.8 Active Testing -- 5.9 Non-propositional Approaches -- References -- 6 Metalearning for Hyperparameter Optimization -- 6.1 Introduction -- 6.2 Basic Hyperparameter Optimization Methods -- 6.3 Bayesian Optimization -- 6.4 Metalearning for Hyperparameter Optimization -- 6.5 Concluding Remarks -- References -- 7 Automating Workflow/Pipeline Design -- 7.1 Introduction -- 7.2 Constraining the Search in Automatic Workflow Design -- 7.3 Strategies Used in Workflow Design -- 7.4 Exploiting Rankings of Successful Plans (Workflows) -- References -- Part II Advanced Techniques and Methods -- 8 Setting Up Configuration Spaces and Experiments -- 8.1 Introduction -- 8.2 Types of Configuration Spaces -- 8.3 Adequacy of Configuration Spaces for Given Tasks -- 8.4 Hyperparameter Importance and Marginal Contribution -- 8.5 Reducing Configuration Spaces -- 8.6 Configuration Spaces in Symbolic Learning -- 8.7 Which Datasets Are Needed? -- 8.8 Complete versus Incomplete Metadata -- 8.9 Exploiting Strategies from Multi-armed Bandits to Schedule Experiments -- 8.10 Discussion -- References -- 9 Combining Base-Learners into Ensembles -- 9.1 Introduction -- 9.2 Bagging and Boosting -- 9.3 Stacking and Cascade Generalization -- 9.4 Cascading and Delegating -- 9.5 Arbitrating -- 9.6 Meta-decision Trees -- 9.7 Discussion -- References -- 10 Metalearning in Ensemble Methods -- 10.1 Introduction -- 10.2 Basic Characteristics of Ensemble Systems -- 10.3 Selection-Based Approaches for Ensemble Generation -- 10.4 Ensemble Learning (per Dataset) -- 10.5 Dynamic Selection of Models (per Instance) -- 10.6 Generation of Hierarchical Ensembles -- 10.7 Conclusions and Future Research -- References -- 11 Algorithm Recommendation for Data Streams -- 11.1 Introduction -- 11.2 Metafeature-Based Approaches -- 11.3 Data Stream Ensembles -- 11.4 Recurring Meta-level Models.

11.5 Challenges for Future Research -- References -- 12 Transfer of Knowledge Across Tasks -- 12.1 Introduction -- 12.2 Background, Terminology, and Notation -- 12.3 Learning Architectures in Transfer Learning -- 12.4 A Theoretical Framework -- References -- 13 Metalearning for Deep Neural Networks -- 13.1 Introduction -- 13.2 Background and Notation -- 13.3 Metric-Based Metalearning -- 13.4 Model-Based Metalearning -- 13.5 Optimization-Based Metalearning -- 13.6 Discussion and Outlook -- References -- 14 Automating Data Science -- 14.1 Introduction -- 14.2 Defining the Current Problem/Task -- 14.3 Identifying the Task Domain and Knowledge -- 14.4 Obtaining the Data -- 14.5 Automating Data Preprocessing and Transformation -- 14.6 Automating Model and Report Generation -- References -- 15 Automating the Design of Complex Systems -- 15.1

Introduction -- 15.2 Exploiting a Richer Set of Operators -- 15.3
Changing the Granularity by Introducing New Concepts -- 15.4 Reusing New Concepts in Further Learning -- 15.5 Iterative Learning -- 15.6
Learning to Solve Interdependent Tasks -- References -- Part III
Organizing and Exploiting Metadata -- 16 Metadata Repositories -- 16.1 Introduction -- 16.2 Organizing the World Machine Learning Information -- 16.3 OpenML -- References -- 17 Learning from Metadata in Repositories -- 17.1 Introduction -- 17.2 Performance Analysis of Algorithms per Dataset -- 17.3 Performance Analysis of Algorithms across Datasets -- 17.4 Effect of Specific Data/Workflow Characteristics on Performance -- 17.5 Summary -- References -- 18 Concluding Remarks -- 18.1 Introduction -- 18.2 Form of Metaknowledge Used in Different Approaches -- 18.3 Future Challenges -- References -- Index.

Sommario/riassunto

This open access book as one of the fastest-growing areas of research in machine learning, metalearning studies principled methods to obtain efficient models and solutions by adapting machine learning and data mining processes. This adaptation usually exploits information from past experience on other tasks and the adaptive processes can involve machine learning approaches. As a related area to metalearning and a hot topic currently, automated machine learning (AutoML) is concerned with automating the machine learning processes. Metalearning and AutoML can help AI learn to control the application of different learning methods and acquire new solutions faster without unnecessary interventions from the user. This book offers a comprehensive and thorough introduction to almost all aspects of metalearning and AutoML, covering the basic concepts and architecture, evaluation, datasets, hyperparameter optimization, ensembles and workflows, and also how this knowledge can be used to select, combine, compose, adapt and configure both algorithms and models to yield faster and better solutions to data mining and data science problems. It can thus help developers to develop systems that can improve themselves through experience. This book is a substantial update of the first edition published in 2009. It includes 18 chapters, more than twice as much as the previous version. This enabled the authors to cover the most relevant topics in more depth and incorporate the overview of recent research in the respective area. The book will be of interest to researchers and graduate students in the areas of machine learning, data mining, data science and artificial intelligence. ; Metalearning is the study of principled methods that exploit metaknowledge to obtain efficient models and solutions by adapting machine learning and data mining processes. While the variety of machine learning and data mining techniques now available can, in principle, provide good model solutions, a methodology is still needed to guide the search for the most appropriate model in an efficient way. Metalearning provides one such methodology that allows systems to become more effective through experience. This book discusses several approaches to obtaining knowledge concerning the performance of machine learning and data mining algorithms. It shows how this knowledge can be reused to select, combine, compose and adapt both algorithms and models to yield faster, more effective solutions to data mining problems. It can thus help developers improve their algorithms and also develop learning systems that can improve themselves. The book will be of interest to researchers and graduate students in the areas of machine learning, data mining and artificial intelligence.