Preface

The continuous growth of successful applications in machine learning and data mining has led to an apparent view of real progress towards the understanding of the nature and mechanisms of learning machines. From the automated classification of millions of luminous objects stored in astronomical images, to the complex analysis of long sequences of genes in biomedical data sets, machine learning has positioned itself as an indispensable tool for data analysis, pattern recognition, and scientific discovery. This apparent progress in the search for accurate predictive models relies on the design of learning algorithms exhibiting novel functionalities. The history of machine learning shows a research community devoted to the study and improvement of a highly diverse set of learning algorithms such as nearest neighbors, Bayesian classifiers, decision trees, neural networks, and support vector machines (to name just a few). While the design of new learning algorithms is certainly important in advancing our ability to finding accurate data models, so is the understanding of the relation between data set characteristics and the particular mechanisms embedded in the learning algorithm. Rather than testing multiple algorithms to assess which one would perform satisfactorily on a certain data set, the end user needs guidelines pointing to the best learning strategy for the particular problem at hand. Researchers and practitioners in machine learning have a clear need to answer the following question: what works well where? There is a strong need to characterize both data distributions and learning mechanisms to construct a theory of learning behavior. Moreover, we advocate the development of a new generation of learning algorithms that are capable of profound adaptations in their behavior to the input data. This may include changes to the model language itself.

Despite different interpretations of the term metalearning, in this book we pursue the goal of finding principled methods that can make learning algorithms adaptive to the characteristics of the data. This can be achieved in many ways as long as there is some form of feedback relating learning performance with data distributions. Thus one can think of the problem of algorithm selection (or ranking), or algorithm combination, as frameworks that exploit past performance to guide the selection of a final model. The ultimate end is to design learning algorithms that adapt to the problem at hand, rather than invoking the same fixed mechanisms independent of the nature of the data under analysis. We explain in this book that a unifying theme in metalearning is that of exploiting experience or metaknowledge to achieve flexible learning systems.

These ideas have brought us together to write a book that summarizes the current state of the art in the field of metalearning. The motivation for such a book can be traced back to the METAL project [166], in which the first three authors were active participants. Our first (electronic) meetings regarding this book took place in the second half of 2005 and have continued over the following three years until the second half of 2008. The project proved challenging in many ways, most particularly in unifying our view concerning the scope of the term metalearning. After long discussions we finally agreed on the definition provided in Chapter 1. Equally challenging was to decide on a list of topics that stand as clearly representative of the field. We hope the reader will find our selection appropriate and sufficiently broad to offer adequate coverage. Finally, it is our hope that this book will serve not only to place many of the ideas now dispersed all over the field into a coherent volume, but also to encourage researchers in machine learning to consider the importance of this fascinating area of study.

Book Organization

The current state of diverse ideas in the field of metalearning is not yet mature enough for a textbook based on a solid conceptual and theoretical framework of learning performance. Given this, we have decided to cover the main topics where there seems to be a clear consensus regarding their relevance and legitimate membership in the field. In the following, we briefly describe the contents of the book acknowledging the contribution of each of the authors. In Chapter 1, all of us worked on introducing the main ideas and concepts that we believe are essential to understanding the field of metalearning. Chapters 2–4 have a more practical flavor, illustrating the important problem of selecting and ranking learning algorithms, with a description of several currently operational applications. In Chapter 2, Soares and Brazdil describe a simple meta-learning system that, given a dataset, provides the user with guidance concerning which learning algorithm to use. The issues that are involved in developing such a system are discussed in more detail by Soares in Chapter 3, including a survey of existing approaches to address them. In Chapter 4, Giraud-Carrier describes a number of systems that incorporate some form of automatic user guidance in the data mining process.

Chapters 5–7, on the other hand, have a more conceptual flavor, covering the combination of classifiers, learning from data streams, and knowledge transfer. Chapter 5, authored by Giraud-Carrier, describes the main concepts behind model combination, including classical techniques such as bagging and boosting, as well as more advanced techniques such as delegating, arbitrating and meta-decision trees. We invited Gama and Castillo to contribute to Chapter 6; the chapter discusses the dynamics of the learning process and general strategies for reasoning about the evolution of the learning process itself. The main characteristics and new constraints on the design of learning algorithms imposed by large volumes of data that evolve over time are described, including embedding change-detection mechanisms in the learning algorithm and the trade-off between the cost of update and the gain in performance. Chapter 7, authored by Vilalta, covers the important topic of knowledge transfer across tasks; the chapter covers topics such as multitask learning, transfer in kernel methods, transfer in parametric Bayesian methods, theoretical models of learning to learn, and new challenges in transfer learning with examples in robotics. Lastly, Chapter 8, authored by Brazdil, discusses the important role of metalearning in the construction of complex systems through the composition of induced subsystems. It is shown how domain-specific metaknowledge can be used to facilitate this task.

Acknowledgements

We wish to express our gratitude to all those who helped in bringing this project to fruition. We are grateful to the University of Porto and Faculty of Economics and also to the Portuguese funding organization FCT for supporting the R&D laboratory LIAAD (Laboratory of Artificial Intelligence and Decision Support) where a significant part of the work associated with this book was carried out. We also acknowledge support from the Portuguese funding organization FCT for the project ALES II – Adaptive LEarning Systems. This work was also partially supported by the US National Science Foundation under grant IIS-0448542.

The motivation for this book came from the involvement of the first three authors in an earlier grant from the European Union (ESPRIT project METAL [166]). We would like to acknowledge financial support from this grant and the contribution of our colleagues: Hilan Bensusan, Helmut Berrer, Sašo Džeroski, Peter Flach, Johannes Fürnkranz, Melanie Hilario, Jörg Keller, Tom Khabaza, Alexandros Kalousis, Petr Kuba, Rui Leite, Guido Lindner, Reza Nakhaeizadeh, Iain Paterson, Yonghong Peng, Rui Pereira, Johann Petrak, Bernhard Pfahringer, Luboš Popelínský, Ljupčo Todorovski, Dietrich Wettschereck, Gerhard Widmer, Adam Woznica and Bernard Zenko.

We are greatly indebted to several other colleagues for their many comments and suggestions that helped improve earlier versions of this book: Bart Bakker, Theodoros Evgeniou, Tom Heskes, Rich Maclin, Andreas Maurer, Tony Martinez, Massimiliano Pontil, Rajat Raina, André Rossi, Peter Stone, Richard Sutton, Juergen Schmidhuber, Matthew Taylor, Lisa Torrey and Roberto Valerio.

Pavel Brazdil would also like to gratefully acknowledge that Robert Kowalski drew his attention to the topic of metareasoning in the 70's and

to express his gratitude to the late Donald Michie for having accepted the proposal to include metalearning in the preparatory stages of the StatLog project [169] in 1989.

Finally, we are most grateful to our editor at Springer, Ronan Nugent, for his patience, gentle prodding and encouragement throughout this project.

Porto, Portugal; Provo, Houston, USA
July 2008

Pavel Brazdil
Christophe Giraud-Carrier
Carlos Soares
Ricardo Vilalta

http://www.springer.com/978-3-540-73262-4

Metalearning
Applications to Data Mining
Brazdil, P.; Giraud Carrier, C.; Soares, C.; Vilalta, R.
2009, XI, 176 p. 53 illus., Hardcover
ISBN: 978-3-540-73262-4