Preface

Rule learning is not only one of the oldest but also one of the most intensively investigated, most frequently used, and best developed fields of machine learning. In more than 30 years of intensive research, many rule learning systems have been developed for propositional and relational learning, and have been successfully used in numerous applications. Rule learning is particularly useful in intelligent data analysis and knowledge discovery tasks, where the compactness of the representation of the discovered knowledge, its interpretability, and the actionability of the learned rules are of utmost importance for successful data analysis.

The aim of this book is to give a comprehensive overview of modern rule learning techniques in a unifying framework which can serve as a basis for future research and development. The book provides an introduction to rule learning in the context of other machine learning and data mining approaches, describes all the essential steps of the rule induction process, and provides an overview of practical systems and their applications. It also introduces a feature-based framework for rule learning algorithms which enables the integration of propositional and relational rule learning concepts.

The book starts by introducing the necessary machine learning and data mining terminology. Chapter 1 provides the motivation for the research, development, and application of rule learning systems. Descriptions of the inductive rule learning task, basic propositional rule learning algorithms, and the basic rule quality evaluation measures are found in Chap. 2. The chapter ends with an overview of the best established rule learning systems. In Chap. 3, coverage space is defined, which enables qualitative and quantitative comparisons of predictive rule learning systems. Coverage space is extensively used in the rest of the book for the evaluation and illustration of various elements constituting the rule learning process.

The central part of the book consists of Chaps. 4–10. It starts by the presentation of features as unifying building blocks for propositional and relational rule learning. Additionally, Chap. 4 presents methods for detecting and eliminating irrelevant features as well as for handling missing and imprecise attribute values which are important constituents of the process of feature construction. Chapter 5 is devoted to relational rule learning: it describes relational feature construction as an upgrade of

propositional feature construction, which enables the methodology described in the following chapters to be applied regardless of the propositional or relational nature of the features. Chapter 6 describes algorithms learning single rules by searching the hypothesis space for rules that cover many positive and few negative examples. These objectives are usually captured in search heuristics, which are extensively discussed in Chap. 7. Chapter 8 shows how individual rules can be combined into complete theories such as rule sets or decision lists. Chapter 9 addresses pruning of rules and rule sets with the goal of learning simple concepts in order to increase interpretability and to prevent overfitting. This part of the book ends with Chap. 10, which contains an extensive presentation of approaches used for classification into multiclass and structured output spaces, unsupervised learning, and regression.

The presentation of modern descriptive rule learning methods and their applications is in the final part of the book. Chapter 11 includes the presentation of contrast set mining, emerging pattern mining, and subgroup discovery approaches. It shows that these supervised descriptive rule learning approaches are compatible with standard rule learning definitions. This chapter ends by presenting subgroup visualization methods which are a useful ingredient of descriptive data analysis. A detailed description of four successful applications of rule learning is given in Chap. 12. These include a social science application in a demographic domain, a biological functional genomics analysis application concerning the analysis of gene expression, and two medical applications. The applications are used to illustrate the data analysis process, including data collection, data preparation, results evaluation, and results interpretation, and to illustrate the performance of selected rule learning algorithms described in the previous chapters of the book on real-world data. The book includes also a rich list of almost 300 references that will help the interested reader in further work.

The book is prepared as a monograph summarizing relevant achievements in the field of rule learning. The target audience is researchers looking for references for further study and teachers in need of detailed material for their machine learning courses. The book is structured as a self-contained introduction to the field of rule learning (Chaps. 1–3) followed by in-depth discussions of relevant subtopics. After reading the introductory part, a reader is prepared to directly skip to any of the following chapters. Jumping to Chap. 12 describing applications of rule learning may be a good motivation for studying the details of the methodology.

The book is written by authors who have been working in the field of rule learning for many years and who themselves developed several of the algorithms and approaches presented in the book. Although rule learning is assumed to be a well-established field with clearly defined concepts, it turned out that finding a unifying approach to present and integrate these concepts was a surprisingly difficult task. This is one of the reasons why the preparation of this book took more than 5 years of joint work.

A good deal of discussion went into the notation to use. The main challenge was to define a consistent notational convention to be used throughout the book because there is no generally accepted notation in the literature. The used notation is gently introduced throughout the book, and is summarized in Table I in a section on

notational conventions immediately following this preface (pp. xi–xiii). We strongly believe that the proposed notation is intuitive. Its use enabled us to present different rule learning approaches in a unifying notation and terminology, hence advancing the theory and understanding of the area of rule learning.

Many researchers have substantially contributed to the content of this book, as most of the chapters are based on our original papers which were written in coauthorship. Most notably, the contributions of Peter Flach and Sašo Džeroski, with whom we have successfully collaborated in rule learning research for nearly a decade, are acknowledged. Other collaborators who have in several aspects contributed to our work include Francesco Billari, Bruno Crémilleux, Alexia Fürnkranz-Prskawetz, Marko Grobelnik, Eyke Hüllermeier, Frederik Janssen, Arno Knobbe, Antonija Krstačić, Goran Krstačić, Eneldo Loza Mencía, Petra Kralj Novak, Sang-Hyeun Park, Martin Scholz, Ashwin Srinivasan, Jan-Nikolas Sulzmann, Jakub Tolar, Geoffrey Webb, Lorenz Weizsäcker, Lars Wohlrab, and Filip Železný; we are grateful for their collaboration and contributions.

We are grateful to our institutions, Technical University of Darmstadt (Germany), Rudjer Bošković Institute (Zagreb, Croatia), and Jožef Stefan Institute (Ljubljana, Slovenia), for providing us with stimulating working environments, and to the national funding agencies who have supported and cofunded this work. We are also grateful to Springer, the publisher of this work, and especially to Ronan Nugent who has shown sufficient patience waiting for this book to be ready for publication, which took much longer than initially expected.

Our final thanks go to Geoffrey Webb for his inspiring Foreword to this book, as well as Richard Wheeler and Peter Ross, who proofread the book for language, consistency, and mathematical correctness. Richard's contribution was especially invaluable due to his passionate reading of the entire book, proving that the book contains sufficiently interesting material also for a very knowledgeable reader.

Darmstadt, Germany   Johannes Fürnkranz
Zagreb, Croatia   Dragan Gamberger
Ljubljana, Slovenia   Nada Lavrač
January 2012