

Acknowledgments

Writing the acknowledgments is always the nicest part! A lot of people have helped us, and we relish this opportunity to thank them. This book has arisen out of the machine learning research project in the Computer Science Department at the University of Waikato, New Zealand. We received generous encouragement and assistance from the academic staff members early on in that project: John Cleary, Sally Jo Cunningham, Matt Humphrey, Lyn Hunt, Bob McQueen, Lloyd Smith, and Tony Smith. Special thanks go to Geoff Holmes, the project leader and source of inspiration, and Bernhard Pfahringer, both of whom also had significant input into many different aspects of the Weka software. All who have worked on the machine learning project here have contributed to our thinking: We would particularly like to mention early students Steve Garner, Stuart Inglis, and Craig Nevill-Manning for helping us to get the project off the ground in the beginning, when success was less certain and things were more difficult.

The Weka system that illustrates the ideas in this book forms a crucial component of it. It was conceived by the authors and designed and implemented principally by Eibe Frank, Mark Hall, Peter Reutemann, and Len Trigg, but many people in the machine learning laboratory at Waikato made significant early contributions. Since the first edition of this book, the Weka team has expanded considerably: So many people have contributed that it is impossible to acknowledge everyone properly. We are grateful to Remco Bouckaert for his Bayes net package and many other contributions, Lin Dong for her implementations of multi-instance learning methods, Dale Fletcher for many database-related aspects, James Foulds for his work on multi-instance filtering, Anna Huang for information bottleneck clustering, Martin Gütlein for his work on feature selection, Kathryn Hempstalk for her one-class classifier, Ashraf Kibriya and Richard Kirkby for contributions far too numerous to list, Niels Landwehr for logistic model trees, Chi-Chung Lau for creating all the icons for the Knowledge Flow interface, Abdelaziz Mahoui for the implementation of K^* , Stefan Mutter for association-rule mining, Malcolm Ware for numerous miscellaneous contributions, Haijian Shi for his implementations of tree learners, Marc Sumner for his work on speeding up logistic model trees, Tony Voyle for least-median-of-squares regression, Yong Wang for Pace regression and the original implementation of $M5'$, and Xin Xu for his multi-instance learning package, *JRip*, logistic regression, and many other contributions. Our sincere thanks go to all these people for their dedicated work, and also to the many contributors to Weka from outside our group at Waikato.

Tucked away as we are in a remote (but very pretty) corner of the southern hemisphere, we greatly appreciate the visitors to our department who play a crucial role in acting as sounding boards and helping us to develop our thinking. We would like to mention in particular Rob Holte, Carl Gutwin, and Russell Beale, each of whom visited us for several months; David Aha, who although he only came for a few days did so at an early and fragile stage of the project and performed a great

service by his enthusiasm and encouragement; and Kai Ming Ting, who worked with us for two years on many of the topics described in Chapter 8 and helped to bring us into the mainstream of machine learning. More recent visitors include Arie Ben-David, Carla Brodley, and Stefan Kramer. We would particularly like to thank Albert Bifet, who gave us detailed feedback on a draft version of the third edition, most of which we have incorporated.

Students at Waikato have played a significant role in the development of the project. Many of them are in the above list of Weka contributors, but they have also contributed in other ways. In the early days, Jamie Littin worked on ripple-down rules and relational learning. Brent Martin explored instance-based learning and nested instance-based representations, Murray Fife slaved over relational learning, and Nadeeka Madapathage investigated the use of functional languages for expressing machine learning algorithms. More recently, Kathryn Hempstalk worked on one-class learning and her research informs part of Section 7.5; likewise, Richard Kirkby's research on data streams informs Section 9.3. Some of the exercises in Chapter 17 were devised by Gabi Schmidberger, Richard Kirkby, and Geoff Holmes. Other graduate students have influenced us in numerous ways, particularly Gordon Paynter, YingYing Wen, and Zane Bray, who have worked with us on text mining, and Quan Sun and Xiaofeng Yu. Colleagues Steve Jones and Malika Mahoui have also made far-reaching contributions to these and other machine learning projects. We have also learned much from our many visiting students from Freiburg, including Nils Weidmann.

Ian Witten would like to acknowledge the formative role of his former students at Calgary, particularly Brent Krawchuk, Dave Maulsby, Thong Phan, and Tanja Mitrovic, all of whom helped him develop his early ideas in machine learning, as did faculty members Bruce MacDonald, Brian Gaines, and David Hill at Calgary, and John Andreae at the University of Canterbury.

Eibe Frank is indebted to his former supervisor at the University of Karlsruhe, Klaus-Peter Huber, who infected him with the fascination of machines that learn. On his travels, Eibe has benefited from interactions with Peter Turney, Joel Martin, and Berry de Bruijn in Canada; Luc de Raedt, Christoph Helma, Kristian Kersting, Stefan Kramer, Ulrich Rückert, and Ashwin Srinivasan in Germany.

Mark Hall thanks his former supervisor Lloyd Smith, now at Missouri State University, who exhibited the patience of Job when his thesis drifted from its original topic into the realms of machine learning. The many and varied people who have been part of, or have visited, the machine learning group at the University of Waikato over the years deserve a special thanks for their valuable insights and stimulating discussions.

Rick Adams and David Bevans of Morgan Kaufmann have worked hard to shape this book, and Marilyn Rash, our project manager, has made the process go very smoothly. We would like to thank the librarians of the Repository of Machine Learning Databases at the University of California, Irvine, whose carefully collected datasets have been invaluable in our research.

Our research has been funded by the New Zealand Foundation for Research, Science, and Technology and the Royal Society of New Zealand Marsden Fund. The Department of Computer Science at the University of Waikato has generously supported us in all sorts of ways, and we owe a particular debt of gratitude to Mark Apperley for his enlightened leadership and warm encouragement. Part of the first edition was written while both authors were visiting the University of Calgary, Canada, and the support of the Computer Science department there is gratefully acknowledged, as well as the positive and helpful attitude of the long-suffering students in the machine learning course, on whom we experimented. Part of the second edition was written at the University of Lethbridge in Southern Alberta on a visit supported by Canada's Informatics Circle of Research Excellence.

Last, and most of all, we are grateful to our families and partners. Pam, Anna, and Nikki were all too well aware of the implications of having an author in the house ("Not again!"), but let Ian go ahead and write the book anyway. Julie was always supportive, even when Eibe had to burn the midnight oil in the machine learning lab, and Immo and Ollig provided exciting diversions. Bernadette too was very supportive, somehow managing to keep the combined noise output of Charlotte, Luke, Zach, and Kyle to a level that allowed Mark to concentrate. Among us, we hail from Canada, England, Germany, Ireland, New Zealand, and Samoa: New Zealand has brought us together and provided an ideal, even idyllic, place to do this work.