
Preface

This book is a general introduction to machine learning that can serve as a textbook for students and researchers in the field. It covers fundamental modern topics in machine learning while providing the theoretical basis and conceptual tools needed for the discussion and justification of algorithms. It also describes several key aspects of the application of these algorithms.

We have aimed to present the most novel theoretical tools and concepts while giving concise proofs, even for relatively advanced results. In general, whenever possible, we have chosen to favor succinctness. Nevertheless, we discuss some crucial complex topics arising in machine learning and highlight several open research questions. Certain topics often merged with others or treated with insufficient attention are discussed separately here and with more emphasis: for example, a different chapter is reserved for multi-class classification, ranking, and regression.

Although we cover a very wide variety of important topics in machine learning, we have chosen to omit a few important ones, including graphical models and neural networks, both for the sake of brevity and because of the current lack of solid theoretical guarantees for some methods.

The book is intended for students and researchers in machine learning, statistics and other related areas. It can be used as a textbook for both graduate and advanced undergraduate classes in machine learning or as a reference text for a research seminar. The first three chapters of the book lay the theoretical foundation for the subsequent material. Other chapters are mostly self-contained, with the exception of chapter 5 which introduces some concepts that are extensively used in later ones. Each chapter concludes with a series of exercises, with full solutions presented separately.

The reader is assumed to be familiar with basic concepts in linear algebra, probability, and analysis of algorithms. However, to further help him, we present in the appendix a concise linear algebra and a probability review, and a short introduction to convex optimization. We have also collected in the appendix a number of useful tools for concentration bounds used in this book.

To our knowledge, there is no single textbook covering all of the material presented here. The need for a unified presentation has been pointed out to us

every year by our machine learning students. There are several good books for various specialized areas, but these books do not include a discussion of other fundamental topics in a general manner. For example, books about kernel methods do not include a discussion of other fundamental topics such as boosting, ranking, reinforcement learning, learning automata or online learning. There also exist more general machine learning books, but the theoretical foundation of our book and our emphasis on proofs make our presentation quite distinct.

Most of the material presented here takes its origins in a machine learning graduate course (*Foundations of Machine Learning*) taught by the first author at the Courant Institute of Mathematical Sciences in New York University over the last seven years. This book has considerably benefited from the comments and suggestions from students in these classes, along with those of many friends, colleagues and researchers to whom we are deeply indebted.

We are particularly grateful to Corinna Cortes and Yishay Mansour who have both made a number of key suggestions for the design and organization of the material presented with detailed comments that we have fully taken into account and that have greatly improved the presentation. We are also grateful to Yishay Mansour for using a preliminary version of the book for teaching and for reporting his feedback to us.

We also thank for discussions, suggested improvement, and contributions of many kinds the following colleagues and friends from academic and corporate research laboratories: Cyril Allauzen, Stephen Boyd, Spencer Greenberg, Lisa Hellerstein, Sanjiv Kumar, Ryan McDonald, Andres Muñoz Medina, Tyler Neylon, Peter Norvig, Fernando Pereira, Maria Pershina, Ashish Rastogi, Michael Riley, Umar Syed, Csaba Szepesvári, Eugene Weinstein, and Jason Weston.

Finally, we thank the MIT Press publication team for their help and support in the development of this text.