Foreword

Machine learning is the latest in a long line of attempts to capture human knowledge and reasoning into a form that is suitable for constructing machines and engineering automated systems. As machine learning becomes more ubiquitous and its software packages become easier to use it is natural and desirable that the low-level technical details are abstracted away and hidden from the practitioner. However, this brings with it the danger that a practitioner becomes unaware of the design decisions and, hence, the limits of machine learning algorithms. The enthusiastic practitioner who is interested to learn more about the magic behind successful machine learning algorithms currently faces a daunting set of pre-requisite knowledge:

- Programming languages and data analysis tools
- Large-scale computation and the associated frameworks
- Mathematics and statistics and how machine learning builds on it

At universities, introductory courses on machine learning tend to spend early parts of the course covering some of these pre-requisites. For historical reasons, courses in machine learning tend to be taught in the computer science department, where students are often trained in the first two areas of knowledge, but not so much in mathematics and statistics. Current machine learning textbooks try to squeeze in one or two chapters of background mathematics, either at the beginning of the book or as appendices. This book brings the mathematical foundations of basic machine learning concepts to the fore and collects the information in a single place.

Why Another Book on Machine Learning?

Machine learning builds upon the language of mathematics to express concepts that seem intuitively obvious but which are surprisingly difficult to formalize. Once properly formalized we can then use the tools of mathematics to derive the consequences of our design choices. This allows us to gain insights into the task we are solving and also the nature of intelligence. One common complaint of students of mathematics around the globe is that the topics covered seem to have little relevance to practical problems. We believe that machine learning is an obvious and direct motivation for people to learn mathematics.

"Math is linked in the popular mind with phobia and anxiety. You'd think we're discussing spiders." (Strogatz, 2014)

This book is intended to be a guidebook to the vast mathematical literature that forms the foundations of modern machine learning. We motivate the need for mathematical concepts by directly pointing out their usefulness in the context of fundamental machine learning problems. In the interest of keeping the book short, many details and more advanced concepts have been left out. Equipped with the basic concepts presented here, and how they fit into the larger context of machine learning, the reader can find numerous resources for further study, which we provide at the end of the respective chapters. For readers with a mathematical background, this book provides a brief but precisely stated glimpse of machine learning. In contrast to other books that focus on methods and models of machine learning (Mackay, 2003; Bishop, 2006; Alpaydin, 2010; Rogers and Girolami, 2016; Murphy, 2012; Barber, 2012; Shalev-Shwartz and Ben-David, 2014) or programmatic aspects of machine learning (Müller and Guido, 2016; Raschka and Mirjalili, 2017; Chollet and Allaire, 2018) we provide only four representative examples of machine learning algorithms. Instead we focus on the mathematical concepts behind the models themselves, with the intent of illuminating their abstract beauty. We hope that all readers will be able to gain a deeper understanding of the basic questions in machine learning and connect practical questions arising from the use of machine learning with fundamental choices in the mathematical model.

Who is the Target Audience?

As applications of machine learning become widespread in society we believe that everybody should have some understanding of its underlying principles. This book is written in an academic mathematical style, which enables us to be precise about the concepts behind machine learning. We encourage readers unfamiliar with this seemingly terse style to persevere and to keep the goals of each topic in mind. We sprinkle comments and remarks throughout the text, in the hope that it provides useful guidance with respect to the big picture. The book assumes the reader to have mathematical knowledge commonly covered in high-school mathematics and physics. For example, the reader should have seen derivatives and integrals before, and geometric vectors in two or three dimensions. Starting from there we generalize these concepts. Therefore, the target audience of the book includes undergraduate university students, evening learners who and people who participate in online machine learning courses.

In analogy to music, there are three types of interaction, which people have with machine learning:

Astute listener

The democratization of machine learning by the provision of open-source software, online tutorials, and cloud-based tools allows users to not worry about the nitty gritty details of pipelines. Users can focus on extracting

insights from data using off-the shelf tools. This enables non-tech savvy domain experts to benefit from machine learning. This is similar to listening to music; the user is able to choose and discern between different types of machine learning, and benefits from it. More experienced users are like music critics, asking important questions about the application of machine learning in society such as ethics, fairness, and privacy of the individual. We hope that this book provides a framework for thinking about the certification and risk management of machine learning systems, and allow them to use their domain expertise to build better machine learning systems.

Experienced artist

Skilled practitioners of machine learning are able to plug and play different tools and libraries into an analysis pipeline. The stereotypical practitioner would be a data scientist or engineer who understands machine learning interfaces and their use cases, and is able to perform wonderful feats of prediction from data. This is similar to virtuosos playing music, where highly skilled practitioners can bring existing instruments to life, and bring enjoyment to their audience. Using the mathematics presented here as a primer, practitioners would be able to understand the benefits and limits of their favorite method, and to extend and generalize existing machine learning algorithms. We hope that this book provides the impetus for more rigorous and principled development of machine learning methods.

Fledgling composer

As machine learning is applied to new domains, developers of machine learning need to develop new methods and extend existing algorithms. They are often researchers who need to understand the mathematical basis of machine learning and uncover relationships between different tasks. This is similar to composers of music who, within the rules and structure of musical theory, create new and amazing pieces. We hope this book provides a high-level overview of other technical books for people who want to become composers of machine learning. There is a great need in society for new researchers who are able to propose and explore novel approaches for attacking the many challenges of learning from data.