Deep learning with R

François Chollet with J. J. Allaire; *Manning*

The steep evolution of deep learning in the early 2010's has been driven by the enormous amount of data and advances in computing power. Deep learning is an engineering science, built up on empirical findings based on multiple layers of artificial neurons rather than by theory and assumptions. Deep learning has led to major breakthroughs, which couldn't be achieved through previous approaches in machine learning, including natural language processing, image analysis, image/text generation, AI in games, and autonomous driving. Consequently, there is substantial interest in deep learning in the statistics community, and this is a very useful introductory book.

This book mainly introduces `Keras` (a Python library developed by the author of this book, François Chollet) and how to use `Keras` for various deep learning models through an R interface. `Keras` is known to be easy to use and user friendly. The R-version of `Keras` will be especially useful in the statistics community.

In the introductory chapter, the authors provide a broad overview of deep learning and its relationship to machine learning and artificial intelligence. Chapter 2 explains basic mathematical concepts such as tensors and backpropagation. Basic data structure is expressed through tensors. For example, timeseries data are encoded as 3-dimensional tensors (*samples*, *timesteps*, *features*), image data are encoded as 4-dimensional tensors (*samples*, *height*, *width, color depth*), and video data are encoded as 5-dimensional tensors (*samples*, *frames*, *height*, *width, color depth*). Backpropagation is a gradient-based optimization algorithm that computes the gradient of the loss function backward from the last layer to the first layer applying the chain rule. These concepts are essential to understand how neural networks work for the practical examples presented in the following chapters.

Chapter 3 offers some hands-on experience that helps understand various types of network architectures, the right learning configuration, how to train models, and how to know which model gives you the right result. This was the chapter that made me open my computer. After a few hassles to install the most updated "`keras`" package in R (which was not described in the book), I started running the `keras` and sample code provided in the book. I then enjoyed reading the book even more. Chapter 4 provides a conceptual framework for general machine-learning problem solving (not limited to deep learning) focusing on generalizing the algorithm to new data. A blueprint described in Chapter 4.5 is an excellent high-level summary, covering all steps to follow, with helpful tips and key choices needed to be made at each step. A common way of handling missing values is also described in this chapter.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/biom.13224.

Chapters 5 and 6 introduce convolutional neural networks (*ConNets*) and recurrent neural networks (*RNNs*) as deep learning models for computer vision/image and general sequence data, respectively. The sequence data include text data (viewed as sequences of words and characters) and time series data (e.g. weather, stock). The first half of the chapters well describes the entire process of using the *ConvNets* and *RNNs* in the `keras` package, including unique preprocess of data (eg, convolution of image data and tokenization of text data) to validating and visualizing the trained representations. However, as a novice to deep learning, I wish these chapters had begun with the architectural foundations of the *ConvNets* and *RNNs*.

I personally was fascinated by two aspects of deep learning: pre-trained networks and use of *ConvNets* for sequencing data. The pre-learned features on a large dataset can be used across different problems, which makes deep learning very effective with even small datasets. The second half of Chapters 5-6 focuses on how to use a pretrained network. As a faster alternative to RNNs for the sequencing data, 1-dimensional *ConvNets* are presented in Chapter 6.4, given that the natural language process can be viewed as pattern recognition similar to the pattern recognition in pixels.

Combining different types of neural networks is introduced in Chapter 7 as an advanced technique to build models with complex structures such as multi-input models, multi-output models, graphical models, layer sharing, and model sharing. This chapter also adds visualization tools to monitor models during training and to make appropriate adjustments. Current and future directions for deep learning, including generative models that create new work, are discussed in Chapters 8 and 9.

To sum up, this is an excellent introductory textbook for statisticians, data scientists, and graduate students. The book covers most fundamental concepts of deep learning, while focusing on their implementation. As. The chapters are logically connected, the book flows easily. Lots of diagrams, pictures, and annotations embedded right next to sample code improve the clarity and intuition, which makes it straightforward to use. If you need one book to get you started with deep learning, and you would like to take advantage of easy accessible R interface, you will enjoy this book.

Sehee Kim

Department of Biostatistics

School of Public Health

University of Michigan

Email: seheek@umich.edu