# ­­­­Chapter 1 - Introduction

## Setting the stage

### 1.1 - Another book on machine learning?

This book is about machine learning, data science, and artificial intelligence, primarily focusing on machine learning. I am sure there are many books on these topics and many more to come. But this book has a different, broader audience. Specifically, I wanted to write something for those of us who find ourselves on highly interdisciplinary technology teams but who are not data scientists or machine learning engineers

|  |  |  |
| --- | --- | --- |
|  | A picture containing shape  Description automatically generated |  |
|  | **Figure 1.1** - A machine learning algorithm made the above logo. I gave it a name and a style, and this logo is one of the many options it created. |  |

### 1.2 - Concepts, abstractions, intuitions over derivation and implementation

The diverse roles, disciplines, skills, and backgrounds within such a team will make cross-disciplinary communication hard, as each team member has their way of describing complex ideas and concepts. In short, it creates the technological equivalent of the tower of Babel.

|  |  |  |
| --- | --- | --- |
|  | Icon  Description automatically generated |  |
|  | **Figure 1.2** - I hope you'll have a few Eureka moments along the way. Sudden insights and the sense of gained intuition can be highly enriching. |  |

This book and its code demos aim to contain just enough depth and detail that it remains accessible to those who hail from a different discipline. Enough detail that complex topics could still be grasped by those unfamiliar with the field. Enough detail for those with whom I share specific skills, like engineers, so it is still helpful to learn something new quickly. And enough for the experts to agree that it is not a complete mockery of the field. It is not an exhaustive overview of everything data science, machine learning, and AI. It is not a deep dive into mathematical proofs. It will not help you find how to parallelize your deep belief neural network over multiple cloud GPU instances to optimize computational resources at a system level (a mouthful, I know). Instead, it is about getting an intuition for the ideas, methods, concepts, and applications. I hope that at the end of it, you will see through the terminology, nomenclature, and equation-heavy lingo and the connections between the ideas behind these disparate machine learning and data science algorithms. And an intuition for what might work, even though you need some help implementing it. And finally, a sense of where it could go wrong. In this way, you can contribute to any discussion on using a particular algorithm to solve a specific problem with equal enthusiasm (knowing what can be done) and skepticism (addressing potential pitfalls).

|  |  |  |
| --- | --- | --- |
|  | Table  Description automatically generated |  |
|  | Figure 1.3 - What letter do you see? This book aims to have you focus on the letter H, and I will leave to more specialized textbooks to deal with all the e's. A quick note about the origin of this image: it has been used in various neuropsychological tests, including autism. Autism is more than a problem in understanding and communicating emotions and affect. You may have heard the phrase 'not seeing the forest for the trees', typically expressed as a criticism of failing to see the 'big picture' and instead focusing on the details. At the very core, people with autism seem to have a more focused, detailed perception of the world than their neurotypical counterparts. They see the e's but fail to see the bigger picture, which is the letter H, composed of all the smaller e's. It is a somewhat controversial claim, and the results aren't entirely conclusive. Still, it fits well with the rare savant with unparalleled memory for small details, whether it manifests as photographic memory or an encyclopedic knowledge of the fact, as well as the observation that to autistic people, the world can be chaotic and overwhelming. I would also be overwhelmed if every single detail entered my consciousness. |  |

**1.3 – Outline**

To do all this, we must go over much material rapidly. I divided this book into five sections. Like most textbooks, we will start the following few chapters by defining some key terms and discussing AI's history. This current text perhaps deviates because we will also discuss, quite early on, the philosophical implications of AI and how we should even begin to compare our AI with the biological systems, us humans included. Why are some things trivial to an AI (like playing chess or GO), and why are some other mundane tasks, like finding an item in a supermarket, far beyond our AI-powered systems? Finally, as food for thought for the reader, we will touch upon the idea of consciousness, sentience, and agency. At what point do the criteria for being human apply to machines as well? We have powerful AI to translate what we say between languages, even Mongolian into Swahili. But to what extent does the machine understand the translated message? I dedicate the next part of the book to the basics of geometry and algebra within the context of machine learning. And no, it will not be an entire crash course on the topics, just enough insights to move forward. We can view machine learning from the perspective of software engineering (the correct algorithm for the right job) or statistics (do the observations provide sufficient evidence for the presence or absence of an unseen variable?). But ultimately, machine learning is based on the mathematics of algebra and geometry. These are at the heart of the computations in practically all machine learning algorithms and offer the best intuition and understanding of how machine learning algorithms transform input data into meaningful and predictive outputs. Finally, getting to machine learning proper, we will use two foundational machine learning algorithms, classification, and regression, to talk in-depth about how these algorithms learn and how we measure the success of what has been learned. We follow this discussion with chapters on various machine learning algorithms, such as decision trees and clustering methods. The remaining two chapters in this sequence are on dimensionality reduction and manifold learning. These methods have specific use cases but are also helpful representations of the geometry underlying machine learning and AI. Finally, we dedicate some chapters in this section to domain-specific techniques, like graph theory and natural language processing. After we discuss these conventional machine learning algorithms, we turn to neural networks. We will use several chapters to build intuition and draw parallels with the methods discussed in the preceding chapters. At first, we will be building logic gates from scratch. At the very end, you should have the tools to create a modern-day-sized neural network using specialized software (e.g., TensorFlow) and hardware (GPUs). Notice a chapter on biological vision within this series on artificial neural networks. I included this chapter as a companion chapter on convolutional neural networks. These networks closely mimic computational processes one can find in a biological system, like the human visual system. Computer vision still serves as a benchmark for the capabilities of neural networks. Understanding how an artificial system mimics a biological one to produce its high accuracy should make a lot of concepts in deep belief neural networks more transparent. Finally, we finish our winding path with machine learning methods that are lesser known but by no means less fascinating. They offer new ways of training our AI, new perspectives, and paradigms that yield new ideas and testable theories. And some of these methods, from quantum computing to genetic algorithms, can drastically change our world. So, to due diligence, we finish with a necessary discussion on some of the legal, ethical, and moral implications of building and deploying AI and glance into the future of AI, if only very, very briefly.

### 1.4 - Examples and Demos

Starting around Chapter 10, when algorithms are initially discussed in some detail, readers can find example Python notebooks that break down some of these algorithms as demos. I have done my best to use novel data sets, with some exceptions where established data sets are better suited to building an understanding and intuition.

|  |  |  |
| --- | --- | --- |
|  | Logo, company name  Description automatically generated |  |
|  | Figure 1.4 - The demos will use traditional toy datasets, but we will use unique data sets whenever possible. All demos are interactive Jupyter notebooks. Sometimes the demo will code an entire algorithm from scratch, using only core Python functionality. We will use widely used toolboxes like Scikit-learn, Keras, and TensorFlow for more complicated machine learning algorithms. Coding algorithms from scratch, depending on your understanding of their inner workings, coding algorithms using toolboxes allow us to dispense with a large amount of code that isn't necessarily relevant to making the point. |  |

### 1.5 - Errata

I am not perfect. There will be bugs and errors, and oversights. If you find one of those, please email me at gbrouwer5151@gmail.com.

### 1.6 – References

Python - https://www.python.org/

Numpy - https://numpy.org/

SciPy - https://scipy.org/

Keras - https://keras.io/

PyTorch - https://pytorch.org/

TensorFlow - https://www.tensorflow.org

Scikit-learn - https://scikit-learn.org/stable/

Nvidia CUDA - https://developer.nvidia.com/cuda-toolkit

AWS SageMaker - https://aws.amazon.com/pm/sagemaker

Google Teachable machine - https://teachablemachine.withgoogle.com/

Google MediaPipe - https://google.github.io/mediapipe/