Introduction

Before talking about this book and about neural networks allow me a brief historical digression about Lev Landau, the famous Nobelist Russian physicist and his “Theoretical Minimum”. The reader will realize soon how this is related to this book and to neural networks.

In 1932 Lev Landau founded the *Kharkiv Theoretical Physics* School. All the students that wanted to be admitted had to pass an exam that was known as the “Theoretical Minimum”. The exam covered all aspects of theoretical physics from classical mechanics to field theory. To pass the exam one had to ask Landau, and if he accepted one would have to pass eight exams on different topics. The exam was so difficult that from 1933 to 1961 only 43 people manage to pass it. Landau would invite the candidate to his apartment and there, by using only paper and pencils, the candidate would have to solve two or three complicated problems. Boris Ioffe recalls[[1]](#footnote-1) how during exams Landau would check over the candidate’s shoulder and if he was silent, then this was a good sign. If he would say “hmmm”, then this was a very bad sign. Two “hmmm” and the exam was over. The list of students that passed Landau’s theoretical minimum was on Landau’s office door and can be seen in Figure 1.

Text, letter

Description automatically generated

Figure 1: the list of students that passed Landau’s Theoretical Minimum from 1943 to 1961. Boris Ioffe can be found at place 13.

Passing the Theoretical Minimum exam would guarantee that a student mastered the necessary fundamentals to do research work in theoretical physics. Without a deep understanding of the basics, it was clear that no impactful and profound discoveries could be made. Having this profound knowledge would allow Landau’s students (note that not all the passed his theoretical minimum went to become his graduate students) to take an advanced research topics and study it in detail. Landau would take a recent research result and choose a student who was supposed to explain and report on the results in a seminar. The speaker not only had to review the paper but was supposed to understand how the results were obtained, including all the necessary mathematical steps and proofs. Boris Ioffe recalls that statements like “the author claims…” were not accepted. Failing to keep up to Landau’s standards in this seminar meant that the student would not be allowed to review any recent results anymore.

When students today want to learn how machine learning, and in particular neural networks work they often cannot find a clear study curriculum. There is no accepted “theoretical minimum” for neural networks and machine learning. Courses give weight to different topics, but no real common curriculum can be found in any book or course. This book goal is to try to address this issue. This book is made of three parts: neural *network theoretical minimum*, *applications with Keras* and *advanced topics*. In the *neural network theoretical minimum* part, no code is used or discussed. If the student masters this part, she or he will have a well-rounded knowledge that will allow her or him to proceed in her or his career in research in machine learning or neural networks. Exercises are provided in each chapter to allow the student to try on her or his own to solve more challenging problems. This part has grown from the first edition of this book (quite different than this edition, as it had a much more applied nature) and from many university courses that I have given over the last few years on deep learning. This part has somewhat higher pre-requisites than the *applications with Keras* one. It requires some intermediate to advanced knowledge of algebra, calculus, and statistics. For more detail you can check the Pre-requisites section below. In the *applications with Keras* part I have taken a more hands-on approach. The easiest concepts are explained again (although mostly in a much easier form), and code examples (by using Keras) are shown to demonstrate how to translate theoretical concepts in working Python code. Note that in the chapters only relevant snippets of code are discussed. The online book <https://adl.toelt.ai> contains many complete examples with commented code and is freely available and should be used as a companion to this book. The online book structure reflects the one of this book, thus finding the right examples should be easy. The last part, *advanced topics*, covers a few interesting areas, as autoencoders, and explains them in more details than usually found. Complete code and examples can again be found on the online book.

I strongly believe that without a proper *neural networks theoretical minimum* no researchers can do any proper science with neural networks. I hope that this book will help and that it will give students and practitioners that wants to work with neural networks a solid basis. I am sure that many will disagree on the choices of topics I made, but I strongly feel that the topics I chose *must* be mastered. There is surely additional material that is important but including too much would have made this book too long and unreadable. I think that the length of the *neural networks theoretical minimum* part is perfect for a one semester introductory course on neural network theory. Such a course could be augmented with some lectures on a few advanced topics given in the second half of the semester and with hands-on exercises taken from the middle part of this book.

# Audience for the book

This book has been written mainly for undergraduate and graduate students in scientific fields: physics, chemistry, biology, computer science or mathematics. But it can be used very successfully by practitioners in a wide variety of areas (from insurance, to banking to pharma just to cite a few). The book focus is on the methods and not on use-cases, so no specific use cases will be found in the book. But after mastering the content of the book, it should be easy for researchers and practitioners to implement the methods described in their own job or projects.

# Structure of the book

As I already mentioned, this book is divided in three parts: *neural network theoretical minimum*, *applications with Keras* and *advanced topics*. Each part can be read independently from each other and have different goals. The *neural network theoretical minimum* has the goal of giving the student a solid understanding of how neural network works and how to use them to do scientifically sound research. The part *applications with Keras* covers the easiest concept again with a large number of Keras examples. The goal in this second part is to help the student in understanding how to translate theoretical concepts in working Python code to use in their own projects and experiments. The last part, *advanced topics*, try to cover a few interesting advanced applications of neural networks (as variational autoencoders) with enough depth for the students to really understand them.

# How to use this book

The student should try to work through the three parts in order. If this is done, the students will start the part *applications with Keras* with a solid theoretical foundation and understanding why the code is written as it is will be much easier to understand. If the student does not have a solid Python knowledge, the time in which she or he study the theory can also be used to expand her or his Python knowledge so that when starting the second part, coding will be much easier. If this book is used for a semester long course, the first and second part could be used in parallel to do theory and hands-on exercises. If you are a teacher and you are unsure how feel free to get in touch with me. Will be happy to help you.

# Pre-requisites

The three parts that make this book have different pre-requisites. For the *neural network theoretical minimum* a good understanding (at an undergraduate level) of algebra, calculus and statistics is required. For this part no coding is required. For the *applications with Keras*, the only prerequisite is an intermediate experience in Python. Know-how with TensorFlow and keras is of course helpful, although not entirely necessary as the code used is for the most part not so complex. For the last part, *advanced topics*, the student needs an advanced mathematical background and a very good Python and Keras experience. Advanced concepts are used and therefore without enough experience trying to understand the topics could be frustrating for the beginners. Note that this part can be omitted in a first read of the book without compromising the understanding of the other parts.

# Final words

I hope that this book will give students a clear curriculum to study neural networks in the most structured and easy way. The topics are not easy and required effort and time. Thus, the students should not be discouraged. Unfortunately to do machine learning is very different than doing copy-paste from some blog in internet. Programming is only a part of it, and without knowing how the algorithms work writing code will be useless, and in the worst case will give wrong results.

Dübendorf, 27th May, 2021

1. B.I. Ioffe, *Concluding Talk at the Workshop «QCD At the Threshold of the Fourth Decade/IOFFEFEST”*, (2002), <https://arxiv.org/abs/hep-ph/0204295v1> [↑](#footnote-ref-1)