

Chapter 1

Introduction



Eva Bartz

It's about resources, time, money, and effort. It's about how science serves our society.

Ever wondered how ice cream manufacturers picked the favorite flavor of the year? Might there be a "fruit conspiracy" (Sect. 5.2)?

The COVID-19 pandemic showed the importance of forecasts concerning health-care workers, protective equipment, vaccines, and so forth. How do we model the data we have to generate sound results and robust conclusions?

How many places for childcare will we need in the future? How many teachers? Local politicians will need numbers to prepare good politics caring for our future, raising our children with the hope for a good education.

The needs and the possibilities for good forecasts and predictions are numerous in our society. Coming from the business end of things I was deeply impressed, how scientific research in general and hyperparameter tuning in particular changes and contributes to our society again and again.

Bartz & Bartz GmbH initiated the methods described in this book to achieve better results in hyperparameter tuning faster, with less effort and costs. Because, let's face it, computational time entails a number of costs. First and foremost it entails the time of the researcher, furthermore a lot of energy. All this equals money. So if we manage to achieve better results in hyperparameter tuning in less time, everybody profits. On a larger scale the methods described may contribute a small part to address some of the challenges we face as a society.

Having initiated the methods in an expertise funded by the Federal Statistical Office of Germany (destatis), we realized that a number of people and businesses might benefit from our knowledge. To be able to enlarge our entrepreneurial effort into a book, scientists from the Institute for Data Science, Engineering, and Analytics

E. Bartz (✉)

Bartz & Bartz GmbH, Gummersbach, Germany

e-mail: eva.bartz@bartzundbartz.de

© The Author(s) 2023

E. Bartz et al. (eds.), *Hyperparameter Tuning for Machine and Deep Learning with R*,
https://doi.org/10.1007/978-981-19-5170-1_1

of the Technische Hochschule Köln (THK) took over. We added the academic point of view to the business consulting, Bartz & Bartz GmbH provided.

Our clients from destatis contributed Chap. 7, discussing the “Hyperparameter Tuning in German Official Statistics”.

We link academic and entrepreneurial requirements, hoping to create a very broad theoretical overview with high practical value for our readers. Thus this book can be used as a handbook as well as a text book. It provides hands-on examples that illustrate how hyperparameter tuning can be applied in practice and gives deep insights into the working mechanisms of Machine Learning (ML) and Deep Learning (DL) methods. Programming code is provided so that users can reproduce the results.

ML and DL methods are becoming more and more important and are used in many industrial production processes, e.g., Cyber-physical Production Systems (CPPS). Several hyperparameters of the methods used have to be set appropriately. Previous projects carried out produced inconsistent results in this regard. For example, with Support Vector Machines (SVMs) it could be observed that the tuning of the hyperparameters is critical to success with the same data material, with random forests the results do not differ too much from one another despite different selected hyperparameter values. While some methods have only one or a few hyperparameters, others provide a large number. In the latter case, optimization using a (more or less) fine grid (grid search) quickly becomes very time-consuming and can therefore no longer be implemented. In addition, the question of how the optimality of a selection can be measured in a statistically valid way (test problem: training/validation/test data and resampling methods) arises for both many and a few hyperparameters. In real-world projects, DL experts have gained profound knowledge over time as to what reasonable hyperparameters are, i.e., Hyper Parameter Tuning (HPT) skills are developed. These skills are based on human expert and domain knowledge and not on valid formal rules.

Figure 1.1 illustrates how data scientists select models, specify metrics, preprocess data, etc. Kedziora et al. (2020) present a similar description. Chollet and Allaire (2018) describe the situation as follows:

If you want to get to the very limit of what can be achieved on a given task, you can't be content with arbitrary [hyperparameter] choices made by a fallible human. Your initial decisions are almost always suboptimal, even if you have good intuition. You can refine your choices by tweaking them by hand and retraining the model repeatedly—that's what machine-learning engineers and researchers spend most of their time doing.

But it shouldn't be your job as a human to fiddle with hyperparameters all day—that is better left to a machine.

Please compare this to Fig. 2.2, which shows how the automated tuning process works. But is it reasonable to transfer the power or decision-making entirely to a machine? I don't think so and you probably don't either. But how do we accomplish that?

This book deals with the hyperparameter tuning of ML and DL algorithms and keeps the human in the loop. In particular, it provides

- a survey of important model parameters;

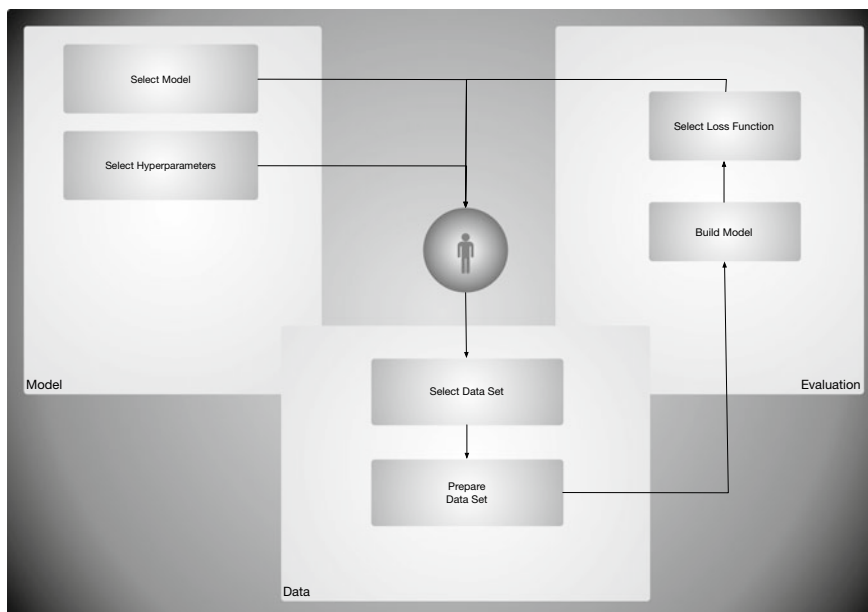


Fig. 1.1 Elements of the HPT process. For a given ML or DL model and its hyperparameters, the data scientist selects a data set and prepares the corresponding data. The model is built and a loss function is computed to evaluate the results

- three parameter tuning studies;
- one extensive global parameter tuning study;
- statistical analysis of the performances of the ML and DL methods based on severity; and
- a new way, based on consensus ranking, to analyze results from multiple algorithms.

More than 30 hyperparameters from six relevant ML methods and DL methods are analyzed. We extend the well-established SPOT framework that improves the optimization of ML methods and DL while increasing the transparency and keeping the human in the loop. The case studies presented in this book can be run on a regular desktop or notebook computer. No high-performance computing facilities are required. Interactive tools for visualization with the popular R package `ploty` are provided. We hope that you achieve better results with significantly less time, costs, effort, and resources using the methods described in this book. We wish you a successful implementation.

This book is structured as follows: Chap. 2 introduces the methodology. Chapter 3 presents models (algorithms or methods) and hyperparameters. HPT methods are introduced in Chap. 4. Chapter 5 discusses result aggregation and severity. Chapter 6 describes the relevance of HPT in industry. Chapter 7 presents HPT in official statistics. Four case studies are presented next. These HPT studies are using the Census-

Income (KDD) Data Set (CID), which will be described in Sect. 8.2.1. The first case study, which describes HPT for random forests, is presented in Chap. 8. This case study might serve as a starting point for the interested reader. The second case study analyzes Extreme Gradient Boosting (XGBoost) and is presented in Chap. 9. The third case study analyzes hyperparameter tuning for DL in Chap. 10. To expand on the example in Chap. 10, which considered tuning a Deep Neural Network (DNN), Chap. 11 also deals with neural networks, but focuses on a different type of learning task: Reinforcement Learning (RL). A global study, which analyzes *tunability*, is presented in Chap. 12.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

