

Contents

1 Automated Machine Learning and Bayesian Optimization	1
1.1 Automated Machine Learning	1
1.1.1 Motivation	1
1.1.2 Model Selection	3
1.1.3 Hyperparameter Optimization	4
1.1.4 Combined Algorithm Selection and Hyperparameter Optimization	6
1.1.5 Why Hyperparameter Optimization Is Important?	6
1.2 The Basic Structure of Bayesian Optimization	8
1.2.1 Sequential Model-Based Optimization	8
1.2.2 Surrogate Model	10
1.2.3 Acquisition Function	12
1.3 Automated Machine Learning for Predictive Analytics	14
References	17
2 From Global Optimization to Optimal Learning	19
2.1 A Priori Analysis of Global Optimization Strategies	20
2.2 Lipschitz Global Optimization (LGO)	21
2.3 Random Search	24
2.3.1 General Properties of Uniform Sampling	26
2.3.2 Cluster Analysis	26
2.3.3 Stopping Rules	26
2.4 Bandits, Active Learning and Bayesian Optimization	28
References	32
3 The Surrogate Model	37
3.1 Gaussian Processes	37
3.1.1 Gaussian Processes Regression	37
3.1.2 Kernel: The Data Geometry of Bayesian Optimization	40

- 3.1.3 Embedding Derivative Observations in the Gaussian Process 43
- 3.1.4 Numerical Instability 46
- 3.2 Thompson Sampling 47
- 3.3 Alternative Models 50
 - 3.3.1 Random Forest 51
 - 3.3.2 Neural Networks: Feedforward, Deep and Bayesian 53
- References 55
- 4 The Acquisition Function 57**
 - 4.1 Traditional Acquisition Functions 57
 - 4.1.1 Probability of Improvement 57
 - 4.1.2 Expected Improvement 58
 - 4.1.3 Upper/Lower Confidence Bound 60
 - 4.2 New Acquisition Functions 62
 - 4.2.1 Scaled Expected Improvement 62
 - 4.2.2 Portfolio Allocation 62
 - 4.2.3 Thompson Sampling 63
 - 4.2.4 Entropy-Based Acquisition Functions 65
 - 4.2.5 Knowledge Gradient 66
 - 4.2.6 Look-Ahead 68
 - 4.2.7 K-Optimality 69
 - 4.3 Optimizing the Acquisition Function 70
 - References 71
- 5 Exotic Bayesian Optimization 73**
 - 5.1 Constrained Global Optimization 73
 - 5.2 Support Vector Machine—Constrained Bayesian Optimization 76
 - 5.3 Safe Bayesian Optimization 86
 - 5.4 Parallel Bayesian Optimization 90
 - 5.5 Multi-objective Bayesian Optimization 91
 - 5.6 Multi-source and Multi-fidelity Bayesian Optimization 93
 - References 94
- 6 Software Resources 97**
 - 6.1 Open Source Software 97
 - 6.2 Bayesian Optimization as a Service 101
 - 6.3 Bayesian Optimization-Based Services for Hyperparameters Optimization 102
 - 6.4 Test Functions and Generators 103
 - 6.4.1 Survey and Site/Repository of Test Functions 103
 - 6.4.2 Test Functions Generators 107

6.5 Non-Bayesian Global Optimization Software	107
References	108
7 Selected Applications	111
7.1 Overview of Applications	111
7.2 Smart Water	116
7.2.1 Leakage Localization	116
7.2.2 Pump Scheduling Optimization	118
References	122

1.1 Automated Machine Learning

1.1.1 Motivation

Automated machine learning (AML) is introduced in this chapter with several model classes and hyperparameter tuning and the specific features of the resulting optimization problems.

These issues are important because an algorithm that scores well on one learning task can score poorly on another, as summarized by the “no free lunch” (NFL) theorem (Wolpert and Macready 1997); even the same algorithm, with different hyperparameter values, can show very different performance (Wolpert 2011). NFL theorems have since given rise to a significant research domain that impacts on both the optimization and supervised learning as analyzed in a recent survey paper (Lopes and Wotzinger 2014) in which specific results are given for early stopping and cross-validation with a conduct the training phase.

While this work emphasizes the role of the ML experts in the design of real-life applications, substantial research activity has been focusing on the possibility to exploit the potential of ML by making it easy to be used off the shelf also by non-experts, taking as a reference the “turnkey” idea of the early 1970s (Lopes 2014). This research emphasizes the configuration of the algorithm and the tuning of their hyperparameters.

The more focused approach to tackle the problem is highlighted by the optimization reformulation in 2.2 and 2.3 which were not directly related to the original problem. In this chapter, the reformulation is shown in a way that is more general, which may have implications for the research in this area.

Global optimization is a challenging task that has been extensively studied in the last few decades. In this chapter, the global optimization is reformulated as a supervised learning problem. In 2.2, the global optimization is reformulated as a supervised learning problem. In 2.3, the global optimization is reformulated as a supervised learning problem.