**Practical Bayesian Optimization of Machine Learning Algorithms**

•The use of machine learning algorithms frequently involves careful tuning of learning parameters and model hyperparameters. Unfortunately, this tuning is of-ten a “black art” requiring expert experience, rules of thumb, or sometimes brute-force search. There is therefore great appeal for automatic approaches that canoptimize the performance of any given learning algorithm to the problem at hand.In this work, we consider this problem through the framework of Bayesian opti-mization, in which a learning algorithm’s generalization performance is modeledas a sample from a Gaussian process (GP). We show that certain choices for thenature of the GP, such as the type of kernel and the treatment of its hyperparame-ters, can play a crucial role in obtaining a good optimizer that can achieve expert-level performance. We describe new algorithms that take into account the variablecost (duration) of learning algorithm experiments and that can leverage the pres-ence of multiple cores for parallel experimentation. We show that these proposedalgorithms improve on previous automatic procedures and can reach orsurpasshuman expert-level optimization for many algorithms including latent Dirichletallocation, structured SVMs and convolutional neural networks. Machine learning algorithms are rarely parameter-free: parameters controlling the rate of learningor the capacity of the underlying model must often be specified. These parameters are often con-sidered nuisances, making it appealing to develop machine learning algorithms with fewer of them.Another, more flexible take on this issue are to view the optimization of such parameters as a proce-dure to be automated. Specifically, we could view such tuning as the optimization of an unknownblack-box function and invoke algorithms developed for such problems.

•