Master's Thesis

Acceleration of Hyperparameter Learning via ℓ_p-bilevel Optimization Problems

Guidance

Professor Nobuo YAMASHITA

Tian ZHONG

Department of Applied Mathematics and Physics

Graduate School of Informatics

Kyoto University

February 2023

Abstract

We focus on tackling the problem of hyperparameter tuning in the ℓ_p-regularization optimization model for machine learning tasks. One of the natural ways is to formulate the problem into a bilevel optimization problem where decision valuables include hyperparameters, its objective function reflects the predictive performance of the model, and the lower-level optimization problem is the original machine learning model with the hyperparameters. The existing method adopts the smoothing approach to transform the ℓ_p-regularization term into a differentiable function. Then using the Karush-Kuhn-Tucker conditions, the differentiable bilevel optimization problem becomes a solvable single-level optimization problem. Such a method has good interpretability, but it must be time-consuming for large-scale learning tasks with a large sample set. To overcome this issue, we propose reducing the sample size based on the dual sparse model used in the support vector regression. By combining a grid search of sample size with bilevel smoothing optimization, we propose a method that can tune the hyperparameters of a model while excluding the influence of irrelevant samples. Numerical experiments demonstrate that our proposed method significantly reduces computation time.