Efficient Hyperparameter Optimization of Deep Learning Algorithms using Deterministic RBF Surrogates

Deep learning algorithms are powerful but are very sensitive to the many hyperparameters they have: number of layers and nodes, learning rate, weights initialization... Optimizing the validation error with respect to the hyperparameters involves the minimization of highly multimodal and expensive function in high dimensions. We propose an algorithm that matches the performance of the state-of-the-art hyperparameter optimization algorithms while using up to 6 times fewer evaluations.

We circumvent the expensive evaluation with a deterministic RBF surrogate defined as:

\[ S_n(x) = \sum_{i=1}^{n} \lambda_i \phi(||x - x_i||) + p(x) \]

We tackle the high-dimensionality by reducing the probability \( \varphi \) of searching along dimension:

\[ \varphi_n = \varphi_0 \left[ 1 - \frac{\ln(n - n_0 + 1)}{\ln(N_{\text{max}} - n_0)} \right] \]

We escape local minima by computing a compound score for each candidate point. The score is a dynamic weighted average of a distance metric based on the distance from the best found solution and of a surrogate value metric:

\[ V_{\text{ev}}(t) = \begin{cases} \frac{\Delta_{\text{max}} - \Delta(t)}{\Delta_{\text{max}} - \Delta_{\text{min}}} & \text{if } \Delta_{\text{max}} \neq \Delta_{\text{min}}, \\ 1, & \text{otherwise.} \end{cases} \]

\[ V_{\text{ev}}(t) = \begin{cases} \frac{s(t) - s_{\text{min}}}{s_{\text{max}} - s_{\text{min}}}, & \text{if } s_{\text{max}} \neq s_{\text{min}}, \\ 1, & \text{otherwise.} \end{cases} \]

Final candidate score:

\[ W(t) = w V_{\text{ev}}(t) + (1 - w) V_{\text{dm}}(t) \]

Optimization of 19 CNN hyperparameters

Optimization of 15 CNN hyperparameters

Optimization of 6 MLP hyperparameters

Optimization of 8 CNN hyperparameters

Table: Evaluations required by HORD to match the best found error by state-of-the-art hyperparameter optimization algorithms

Comparison

GitHub: bit.ly/hord-aaai

Supplement and more at: ilija139.github.io

a) Graduate School for Integrative Sciences and Engineering
b) Industrial and Systems Engineering
c) Electrical and Computer Engineering