Meta-Learning - Using Prior Data to Warmstart Optimization

Marc Speckmann, Robin Gieseke, Kristina Thieme

Leibniz University Hannover

1 TL;DR

This AutoML project focuses on the use of prior configuration data to improve optimisation efficiency. Approaches include pre-configuration optimisers and multi-fidelity techniques.
The main objective is to find a multi-class neural network classification system with sufficient performance to solve the DeepWeeds dataset. Overall, the best test accuracy achieved was 75.25%.

2 Motivation & Problem Setting

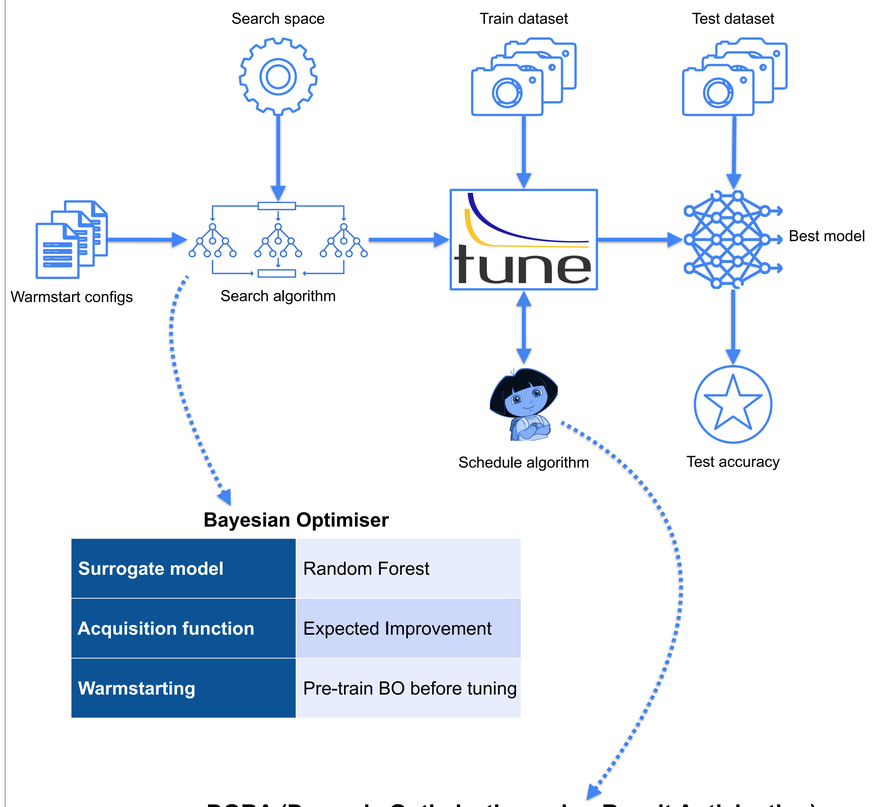
Motivation

- Better Model Performance, Resource Savings & Multi-Fidelity Optimization

Problem Setting

- *Objective*: Find neural network for multi-class classification with test accuracy > 65%
- *Resource Constraints*: Maximum 6h runtime with max 20 epochs per model
- *Prior Configuration Data*
- *Diverse Conditional Configuration Search Space*
- *Dataset:* DeepWeeds (32x32)

3 Approach



Bayesian Optimiser

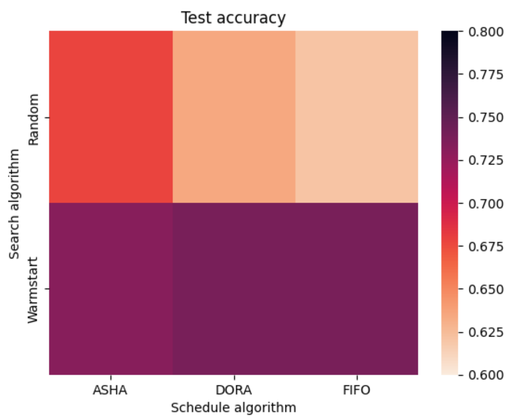
Surrogate model	Random Forest
Acquisition function	Expected Improvement
Warmstarting	Pre-train BO before tuning

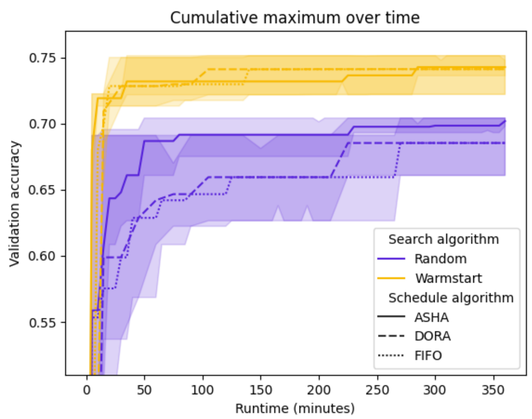
DORA (Dynamic Optimization using Result Anticipation):
The algorithm employs an adaptive mechanism to modulate epoch fidelity by leveraging predictive techniques to anticipate performance outcomes in future training epochs.

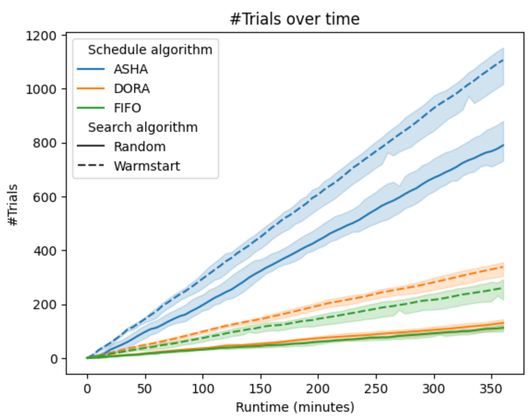
Further details can be found in our public code repository

4 Key Insights

- All experiments were conducted on a NVIDIA A100-SXM4-40GB employing the same 3 randomised seeds.
- The accuracy on the test dataset exhibited a notable enhancement of the warm-started bayesian search algorithm compared to random search.
- The cumulative maximum also underscored the proficiency of the warm-started search.
- The proposed scheduling algorithm (DORA) showed a slightly improvement against FIFO scheduling but not as substantial as the asynchronous HyperBand (ASHA).







5 Future Works

- Warmup the search algorithm by initialising it with configuration data sourced from various datasets without specific constraints.
- Improve the concept of adaptively adjusting fidelity by predicting the future performance of ongoing trials. For example, by utilising the warmstart configurations.