Deep Learning Lab Course – Report for assignment 5

The average optimization trajectories (validation error per epoch) for 10 independent runs of both random search and Bayesian optimization are shown in figure 1. One can observe that Bayesian optimization finds configurations with similar performance to the ones found by random search in the first few iterations. After about 13 iterations, the configurations found by Bayesian optimization provide superior performance which steadily improves until it finally converges to a validation error slightly below 0.16 after 40 iterations. Meanwhile, random search occasionally finds configurations which are better than the incumbent (with decreasing probability, as can be seen by the decreasing slope), but it does not converge to the minimum and stays above 0.17 validation error.
The weak performance of Bayesian optimization in the first few iterations is probably due to the procedure not being able to fit a precise approximation of the objective function, as it requires evaluations at more points to make a reasonable estimation.

The estimated cumulative runtime for the configuration evaluation per epoch for both random search and Bayesian optimization is shown in figure 2. For random search, the estimated runtime is ≈ 18905 seconds to evaluate all configurations, whereas for Bayesian optimization it is ≈ 27677 seconds. So the latter method created configurations which on average lead to 46% longer training times. This probably happens because Bayesian optimization chooses a higher number of filters per layer (since the numbers of filters are the parameters which most influence the runtime of the training).

Figure 1: Average optimization trajectories for random search and Bayesian optimization

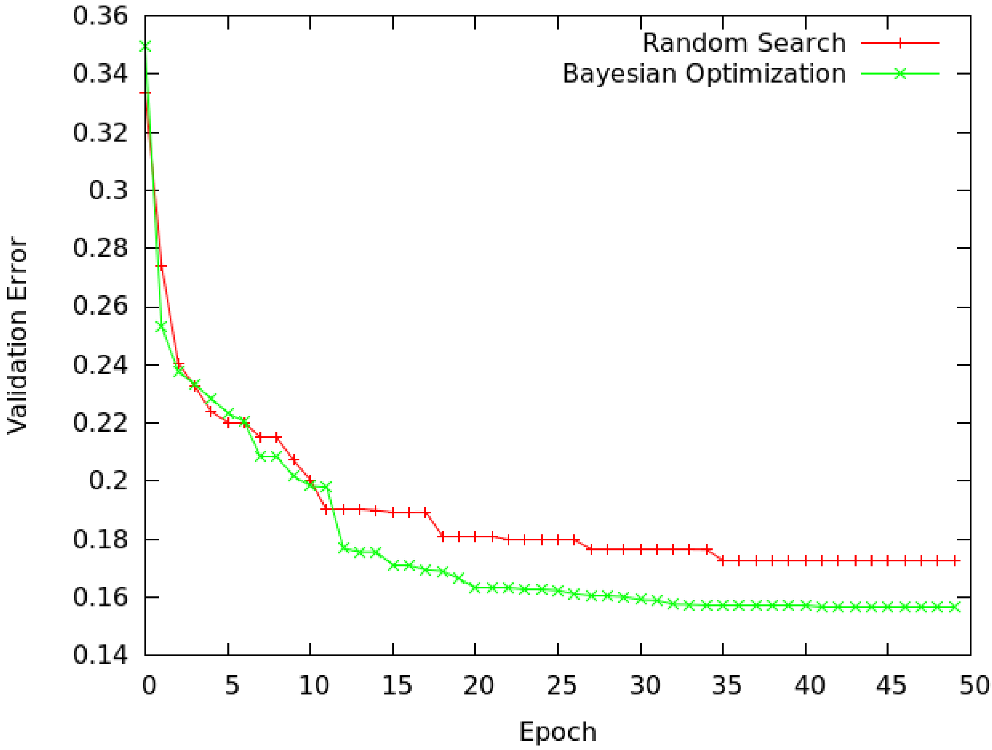

Figure 2: Cumulative runtime per epoch for random search and Bayesian optimization