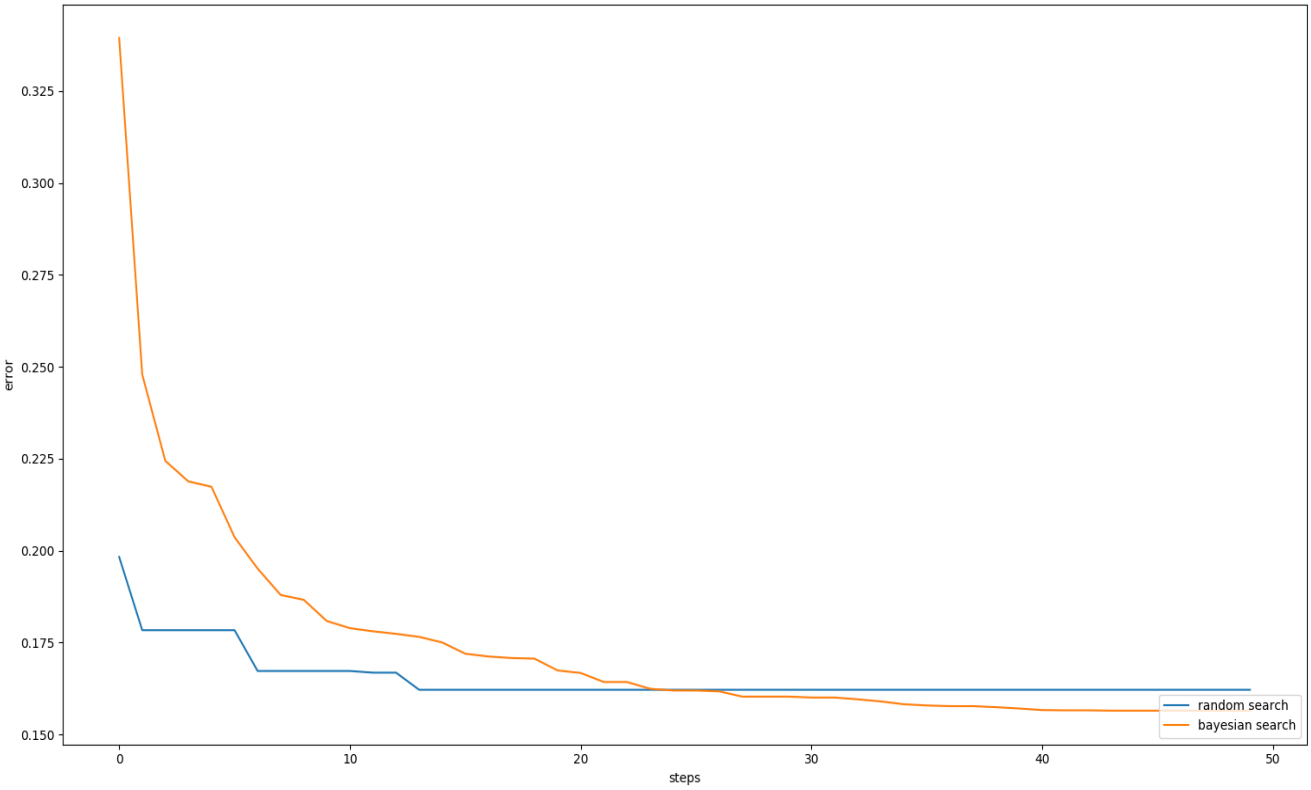
Deep Learning Lab
Exercise 5
Hyper-parameter Optimization

by
Muhammad Hamiz Ahmed

This exercise was aimed to optimize hyper parameters of an already trained neural network, using Bayesian optimization and Random Search. The hyper-parameters of the convolutional neural network and their lower and upper bounds are: learning rate = [10^-6 , 10^0], batch size = {32, 512} and number of filters for each layer between {24 , 210}.

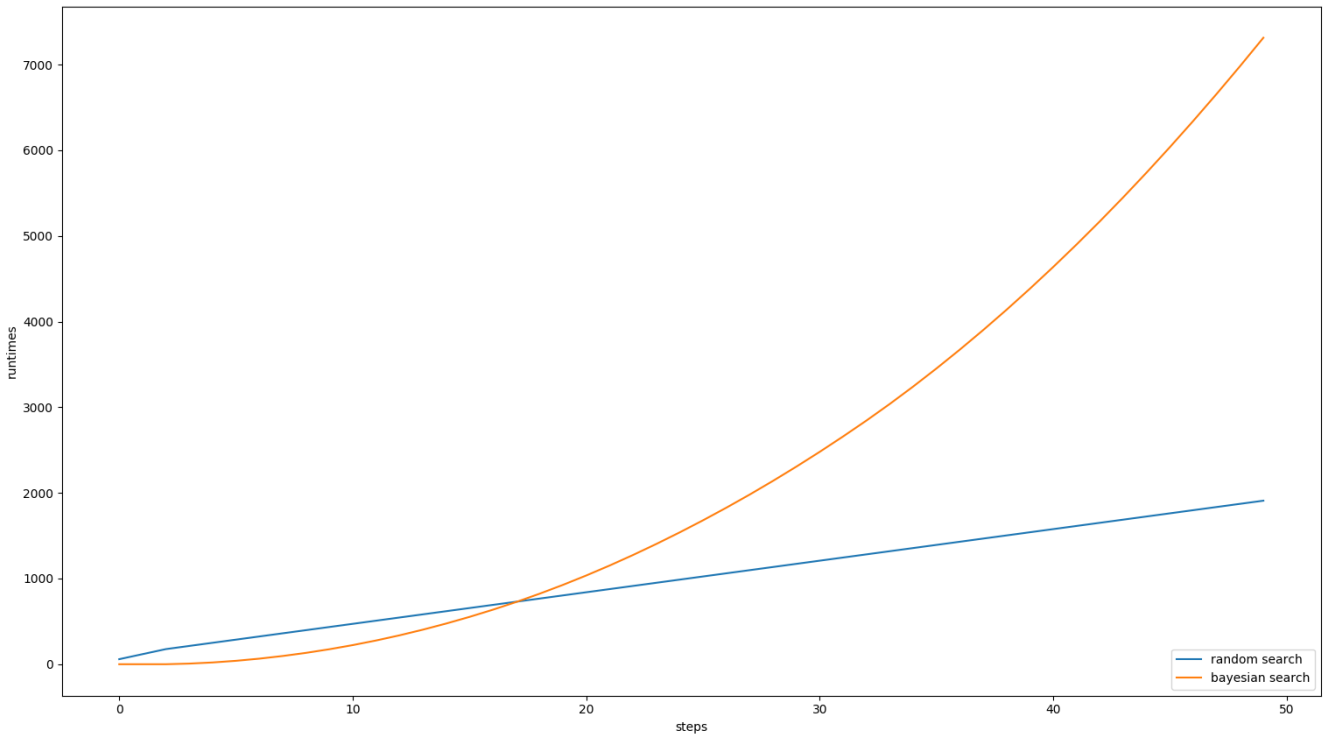
The following figure shows the mean performance of the incumbents in Bayesian and Random Search.



It can be seen. The error gets reduced with the number of iterations. However, initially, Random search starts with a lower error value compared to the Bayesian Search. This is because the first randomly chosen configuration would have given a low error value. It takes around 40 steps for Bayesian to achieve lower error value than Random Search. This

is because Bayesian takes new decision based on its previous performance. Therefore, it improves with time.

The cumulative runtime plot between the two search strategies is as follows:



It can be seen that Bayesian takes a lot more time than the Random Search. Also, the runtime would have been much more if we had used the true objective function.