A Simple Dynamic Learning Rate Configuration for the DAC4SGD Track

Tien Dung LE

Overview. We improved the scheduler configuration provided by DAC-4-SGD (2022). The new scheduler is a combination of 2 different sub-schedulers.

`CosineAnnealingLRPolicy`: At the beginning, we use the cosine annealing scheduler (Loshchilov and Hutter, 2017). Starting from an initial learning rate (configurable parameter) reduces the learning rate to 0 following a re-scaled half cosine curve.

`SigmoidAnnealingLRPolicy`: After that, we use the sigmoid annealing scheduler which is similar to the CosineAnnealingLRPolicy but instead of the cosine curve we use the sigmoid curve.

References

DAC-4-SGD (2022). Dac4sgd. Online: `https://github.com/automl/DAC4SGD`.

Loshchilov, I. and Hutter, F. (2017). Sgdr: Stochastic gradient descent with warm restarts. In *Proceedings of the International Conference on Learning Representations (ICLR'17)*. Published online: `iclr.cc`.