Infinitely Wide Neural Networks for Small Data Tasks.

Theoretical research has shown that an infinitely wide neural network trained under L2 loss by gradient descent with an infinitesimally small learning rate is equivalent to kernel regression with respect to a so called Neural Tangent Kernel (NTK)**.** Since Kernel Methods are highly efficient on small data tasks, we attempt to test the Neural Tangent Kernel on the UCI dataset collection, since they have a large collection of small-data data sets and compare their performance to other tried and tested models such as Random Forests, Support Vector Machines (SVM), and (Deep) Neural Networks.

We use the work of Arora et al. (2019) to implement NTK. In addition, we also use the following research (Fernández et al. 2014; Olson 2018) to develop the benchmarks to evaluate the performance of NTK for classification.

We have developed theoretical understanding of NTK and its relationship to infinitely wide neural networks. Further we collected UCI datasets and have examined the performance of Random Forest, Decision Trees, and SVM. For the remaining duration of the internship, we plan to build neural network and NTKs and compare their performance with other models mentioned above.

References

Arora, S., Du, S. S., Li, Z., Salakhutdinov, R., Wang, R., & Yu, D. (2019). Harnessing the power of infinitely wide deep nets on small-data tasks. arXiv preprint arXiv:1910.01663.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. The journal of machine learning research, 15(1), 3133-3181.

Olson, M., Wyner, A., & Berk, R. (2018). Modern neural networks generalize on small data sets. Advances in neural information processing systems (NIPS), 31.