Infinitely ~~Small Data Tasks~~Wide Neural Networks ~~For~~ 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~~ and other models~~.

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: Not Bold

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.

~~We take high inspiration from the paper "Harnessing the Power of Infinitely Wide Deep Nets on Small-data Tasks".~~

~~Along with the work of "Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?" which gave us various models to try from and chose the best ones.~~

~~Initially research was done on the Neural Tangent Kernel to try and understand their working and their relation with infinitely wide neural nets.~~

~~After this, the UCI dataset (around ~120 datasets) was collected on which we then trained and collected the accuracies of a few machine learning models such as Random Forests, Decision Trees, and Support Vector Machines.~~

~~Further, we still need to build models corresponding to NNs and also the so called NTKs and compare them to see how the NTK models stands amongst other models.~~

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.