CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday – May 11, 2018

Deep Neural Network as Gaussian Processes

Dr. Guofei Pang

It has long been known that a single-layer fully-connected neural network with an i.i.d. prior over its parameters is equivalent to a Gaussian process (GP), in the limit of infinite network width. This correspondence enables exact Bayesian inference for infinite width neural networks on regression tasks by means of evaluating the corresponding GP. Recently, kernel functions which mimic multi-layer random neural networks have been developed, but only outside of a Bayesian framework. As such, previous work has not identified that these kernels can be used as covariance functions for GPs and allows fully Bayesian prediction with a deep neural network. In this work, we derive the exact equivalence between infinitely wide deep networks and GPs. We further develop a computationally efficient pipeline to compute the covariance function for these GPs. We then use the resulting GPs to perform Bayesian inference for wide deep neural networks on MNIST and CIFAR-10. We observe that trained neural network accuracy approaches that of the corresponding GP with increasing layer width, and that the GP uncertainty is strongly correlated with trained network prediction error. We further find that test performance increases as finite-width trained networks are made wider and more similar to a GP, and thus that GP predictions typically outperform those of finite-width networks. Finally we connect the performance of these GPs to the recent theory of signal propagation in random neural networks.