Application of Quantum Annealing to Training of Deep Neural Networks

Steve Adachi

Lockheed Martin Space Systems Company, 1111 Lockheed Martin Way, Sunnyvale CA 94089

In Deep Learning, a well-known approach for training a Deep Neural Network starts by training a generative Deep Belief Network model, typically using Contrastive Divergence (CD), then fine-tuning the weights using backpropagation or other discriminative techniques. However, the generative training can be time-consuming due to the slow mixing of Gibbs sampling. We investigated an alternative approach that estimates model expectations of Restricted Boltzmann Machines using samples from a D-Wave quantum annealing machine. We tested this method on a coarse-grained version of the MNIST data set. In our tests we found that the quantum sampling-based training approach achieves comparable or better accuracy with significantly fewer iterations of generative training than conventional CD-based training. Further investigation is needed to determine whether similar improvements can be achieved for other data sets, and to what extent these improvements can be attributed to quantum effects.