TIFFANY VLAAR, McGill University / Mila
Constrained and Multirate Training of Neural Networks

I will describe algorithms for regularizing and training deep neural networks. Soft constraints, which add a penalty term to the loss, are typically used as a form ofexplicit regularization for neural network training. In this talk I describe a method for efficiently incorporating constraints into a stochast ic gradient Langevin framework for the training of deep neural networks. In contrast to soft constraints, our constraints offer direct control of the parameter space, which allows us to study their effect on generalization. In the second part of the talk, I illustrate the presence of latent multiple time scales in deep learning applications. Different features present in the data can be learned by training a neural network on different time scales simultaneously. By choosing appropriate partitionings of the network parameters into fast and slow parts I show that our multirate techniques can be used to train deep neural networks for transfer learning applications in vision and natural language processing in half the time, without reducing the generalization performance of the model.