**CONSTRAINED AND PARTITIONED TRAINING OF NEURAL NETWORKS**

TIFFANY VLAAR

One half of the talk will discuss the use of constrained stochastic differential equations to train deep neural networks. Common techniques used to improve the generalisation performance of deep neural networks (such as e.g. L2 regularisation and batch normalisation) are tantamount to imposing a parameter constraint, but despite their widespread use are often not well understood. In the talk I will describe an approach for efficiently incorporating hard constraints into a stochastic gradient Langevin dynamics framework. Our constraints offer direct control of the parameter space, which allows us to study their effect on generalisation. In the other half of the talk, I will focus on the role played by individual layers and substructures of neural networks: layer-wise sensitivity to the choice of initialisation and optimiser hyperparameter settings varies and training neural network layers differently may lead to enhanced generalisation and/or reduced computational cost. Embedding the loss gradient in a second order Langevin dynamics framework and using low temperatures in combination with partitioned integrators can lead to enhanced generalization performance of neural networks on certain classification tasks. Further, I will show that a multirate approach can be used to train deep neural networks for transfer learning applications in half the time, without reducing the generalisation performance of the model.