Tapani Raiko, Aalto University

Combining Supervised and Unsupervised Learning (and the Ladder Network)

Most of the best results in deep learning have recently been purely supervised, that is, training is based on mimicking the given outputs in the training set. However, it is expected that unsupervised learning will become far more important in the long term (LeCun et al., Nature 2015), and it was unsupervised layerwise pretraining that first allowed building deep networks (Hinton and Salakhutdinov, Science 2006). We recently introduced a new approach called the Ladder network (Rasmus et al., arXiv 2015), whose unsupervised objective is well compatible with the supervised objective, improving the state of the art in both semisupervised and supervised benchmarks. I will describe a lengthy background covering semisupervised learning, denoising autoencoders, connection between probabilistic modelling and denoising, and layerwise pretraining, before describing the Ladder network.