Contribution ID: **21**   Type: **not specified**

A ground-up construction of deep learning

Wednesday 11 November 2015 10:45 (45 minutes)

I propose to give a ground up construction of deep learning as it is in it's modern state. Starting from it's beginnings in the 90's, I plan on showing the relevant (for physics) differences in optimization, construction, activation functions, initialization, and other tricks that have been accrued over the last 20 years. In addition, I plan on showing why deeper, wider basic feedforward architectures can be used. Coupling this with MaxOut layers, modern GPUs, and including both l1 and l2 forms of regularization, we have the current "state of the art" in basic feedforward networks. I plan on discussing pre-training using deep autoencoders and RBMs, and explaining why this has fallen out of favor when you have lots of labeled data. While discussing each of these points, I propose to explain why these particular characteristics are valuable for HEP. Finally, the last topic on basic feedforward networks – interpretation. I plan on discussing latent representations of important variables (i.e., mass, pT) that are contained in a dense or distributed fashion inside the hidden layers, as well as nifty ways of extracting variable importance.

I also propose a short discussion on dark knowledge – i.e., training very deep, very wide neural nets then using the outputs of these as targets for a smaller, shallower neural networks – this has been shown to be incredibly useful for focusing the network to learn important information. Why is this relevant for physics? well, we could think of trigger level or hardware level applications, where we need FPGA level (for example) implementations of nets that cannot be very deep.

Then I propose to discuss (relatively briefly) the uses cases of Convolution Networks (one current area of research for me) and recurrent neural networks in physics, as well as giving a broad overview of what they are and what domains they typically belong to – i.e., jet image work with convolutional nets, or jet tagging that can read in info from each track in the case of RNNs.

Presenter: DE OLIVEIRA, Luke Percival (SLAC National Accelerator Laboratory (US))

Session Classification: Wednesday Morning Session