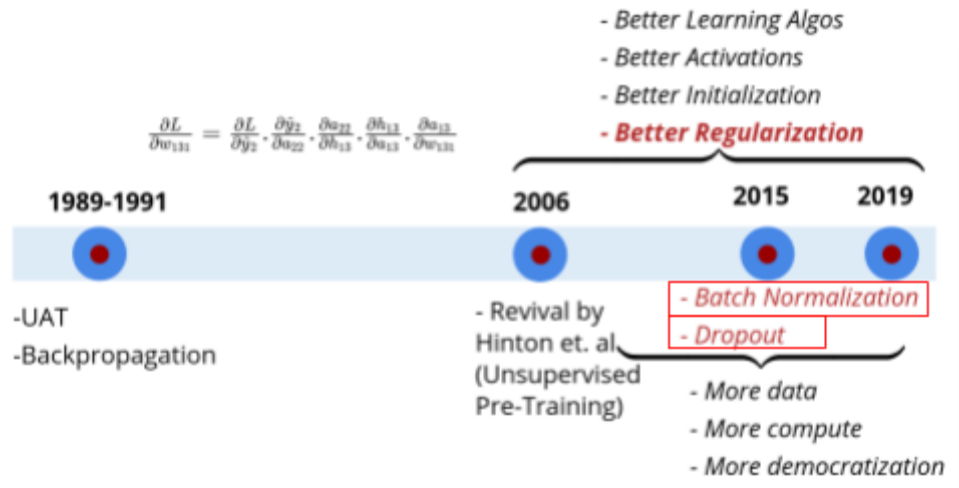
PadhAI: Batch Normalization and Dropout

One Fourth Labs

Summary and what next?

Let's summarize what we've completed so far

1. Let's have another brief look at our Deep Learning timeline



2. Here's a brief overview of the topics covered so far
3. Feedforward Neural networks
   a. Universal approximation theorem (UAT)
   b. Backpropagation
   c. Learning Algorithms
   d. Activation functions
   e. Regularization methods
4. Convolutional Neural networks
   a. Convolutional operation
   b. CNN Architectures (AlexNet, VGGNet, ResNet etc)
   c. Batch Normalization:
      i. An interesting point to note is that batch normalization can either be performed pre-activation or post-activation (after non-linearity like ReLU is applied)
   d. Dropout:
      i. For hidden layers, we normally use probability p = 0.5
      ii. For input layer, we use p = 0.8
      iii. Using 50% probability in the input layer would cause too much noise to the network.
5. The next set of topics to cover are as follows:
   a. Recurrent Neural Networks (RNNs)
   b. Encoder-Decoder Models