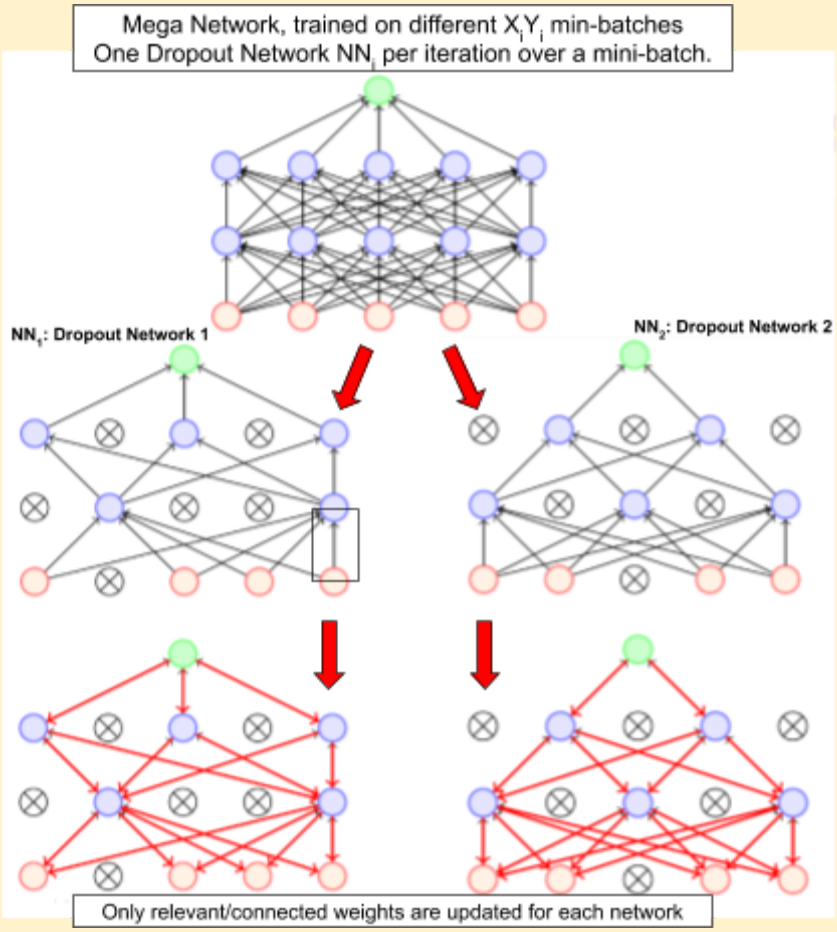
PadhAI: Batch Normalization and Dropout

One Fourth Labs

Training without dropout

What does the training procedure look like?

1. Let's look at the training procedure for dropped-out networks



   a. **Initialise**:
      i. Initialise the parameters
   b. **Iterate over data**:
      i. X_iY_i = Current_mini_batch
      ii. NN_i = dropout Network for that mini-batch (made using coin-toss method for each neuron n)
      iii. The same/relevant weights are shared between the mega-network and the dropped out network.
      iv. $L(\theta) = compute\_loss(NN_i, X_iY_i)$
      v. Backpropagate(NN_i)
      vi. Only connected weights are updated.
   c. **End**:
      i. Repeat till satisfied