Heuristic selection of the optimal neurons in a layer is no longer necessary as regularization algorithms such as Drop-Off, L1 & L2 Regularization ensure that the learned weight distribution implicitly reduces/balances the contribution of any sub-optimally greater number of neurons.

Various Optimization Algorithms are compared based on various features such as model-performance, training-time, final error-rate, and overfitting is judged.