Caleb Princewill Nwokocha February 04, 2023

Convergence of Comprehensive Network

Understanding convergence of comprehensive node and comprehensive layer is required for understanding convergence of comprehensive network. Changes in the node error determines whether the node converges to its objective during training. The gradient of $(e_i)^\phi$ with respect to example i can be found through partial differentiation; where $(e_i)^\phi$ is the node error raised to ϕ. If $\frac{\partial (e_i)^\phi}{\partial i}$ is negative, then $(e_i)^\phi$ at example i is lesser than $(e_{i-1})^\phi$ at example $i-1$. Inversely, if $\frac{\partial (e_i)^\phi}{\partial i}$ is positive, then $(e_i)^\phi$ at example i is greater than $(e_{i-1})^\phi$ at example $i-1$.

Recall that $\bar{e}_i = \left(\frac{(i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi}{i}\right)^{\frac{1}{\phi}}$. To find $(e_i)^\phi$, consider $\bar{e}_i = \frac{\left[(i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi\right]^{\frac{1}{\phi}}}{i^{\frac{1}{\phi}}}$.

Multiplying both sides of $\bar{e}_i = \frac{\left[(i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi\right]^{\frac{1}{\phi}}}{i^{\frac{1}{\phi}}}$ by $i^{\frac{1}{\phi}}$ will give $\bar{e}_i i^{\frac{1}{\phi}} = \left[(i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi\right]^{\frac{1}{\phi}}$, and then raising both sides of $\bar{e}_i i^{\frac{1}{\phi}} = \left[(i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi\right]^{\frac{1}{\phi}}$ to ϕ will give $\left(\bar{e}_i i^{\frac{1}{\phi}}\right)^\phi = (i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi$. The equation $\left(\bar{e}_i i^{\frac{1}{\phi}}\right)^\phi = (i-1)(\bar{e}_{i-1})^\phi + (e_i)^\phi$ can be rearranged to $\left(\bar{e}_i i^{\frac{1}{\phi}}\right)^\phi - (i-1)(\bar{e}_{i-1})^\phi = (e_i)^\phi$ and $\left(\bar{e}_i i^{\frac{1}{\phi}}\right)^\phi - i(\bar{e}_{i-1})^\phi + (\bar{e}_{i-1})^\phi = (e_i)^\phi$. By power rule, $\left(\frac{\bar{e}_i}{\phi}\left(i^{\frac{1-\phi}{\phi}}\right)\right)^\phi - (\bar{e}_{i-1})^\phi = \frac{\partial (e_i)^\phi}{\partial i}$.

A comprehensive node is converging if $\frac{\partial (e_i)^\phi}{\partial i}$ is negative. Furthermore, since a comprehensive layer consists of n comprehensive nodes, it is true that the layer is converging if $\left(\frac{\partial (e_i)^\phi}{\partial i}\right)_1, \cdots, \left(\frac{\partial (e_i)^\phi}{\partial i}\right)_n$ are all negative. The same is true for comprehensive network: If all $\left(\frac{\partial (e_i)^\phi}{\partial i}\right)_1, \cdots, \left(\frac{\partial (e_i)^\phi}{\partial i}\right)_n$ of the network output layer are negative, then the network is converging.