All data: What are the sizes & things:

$W^{(1)}$ is $k \times n$ $\quad$ $\vec{b}^{(1)}$ is $k \times 1$

$W^{(2)}$ is $m \times k$ $\quad$ $\vec{b}^{(2)}$ is $m \times 1$

Given $\underline{X}$ is $n \times p$ $\quad$ (Net is $n-k-m$)

$P^{(1)} = W^{(1)} \underline{X} + \vec{b}^{(1)}$ $\quad$ ($P^{(1)}$ is $k \times p$)

$S^{(1)} = \sigma(P^{(1)})$ $\quad$ ($S^{(1)}$ is $k \times p$)

$dS^{(1)} = \sigma'(P^{(1)})$ $\quad$ ($dS^{(1)}$ is $k \times p$)

$P^{(2)} = W^{(2)} S^{(2)} + \vec{b}^{(2)}$ $\quad$ ($P^{(2)}$ is $m \times p$)

$S^{(2)} = \sigma(P^{(2)})$ $\quad$ ($S^{(2)}$ is $m \times p$)

$dS^{(2)} =$ ones $\quad$ (Matrix of 1's)

$\Delta^{(2)} = T - Y$ or $\frac{2}{p} * (T - Y)$ $\quad$ $m \times p$

$\Delta^{(1)} = (W^{(2)})^T \Delta^{(2)} .* dS^{(1)}$ $\quad$ $k \times p$

($k \times m$, $m \times p$, $k \times p$)

$\Delta W^{(1)} = \Delta^{(1)} \underline{X}^T$ $\quad$ dims: $k \times m = (k \times p)(p \times n)$

$\Delta W^{(2)} = \Delta^{(2)} (S^{(1)})^T$

($m \times p$, $p \times k$)

$\Delta b^{(1)} =$ sum $\Delta^{(1)}$ to $k \times 1$ $\quad$ ($\Delta^{(1)}$ is $k \times p$)

$\Delta b^{(2)} =$ sum $\Delta^{(2)}$ to $m \times 1$ $\quad$ ($\Delta^{(2)}$ is $m \times p$)