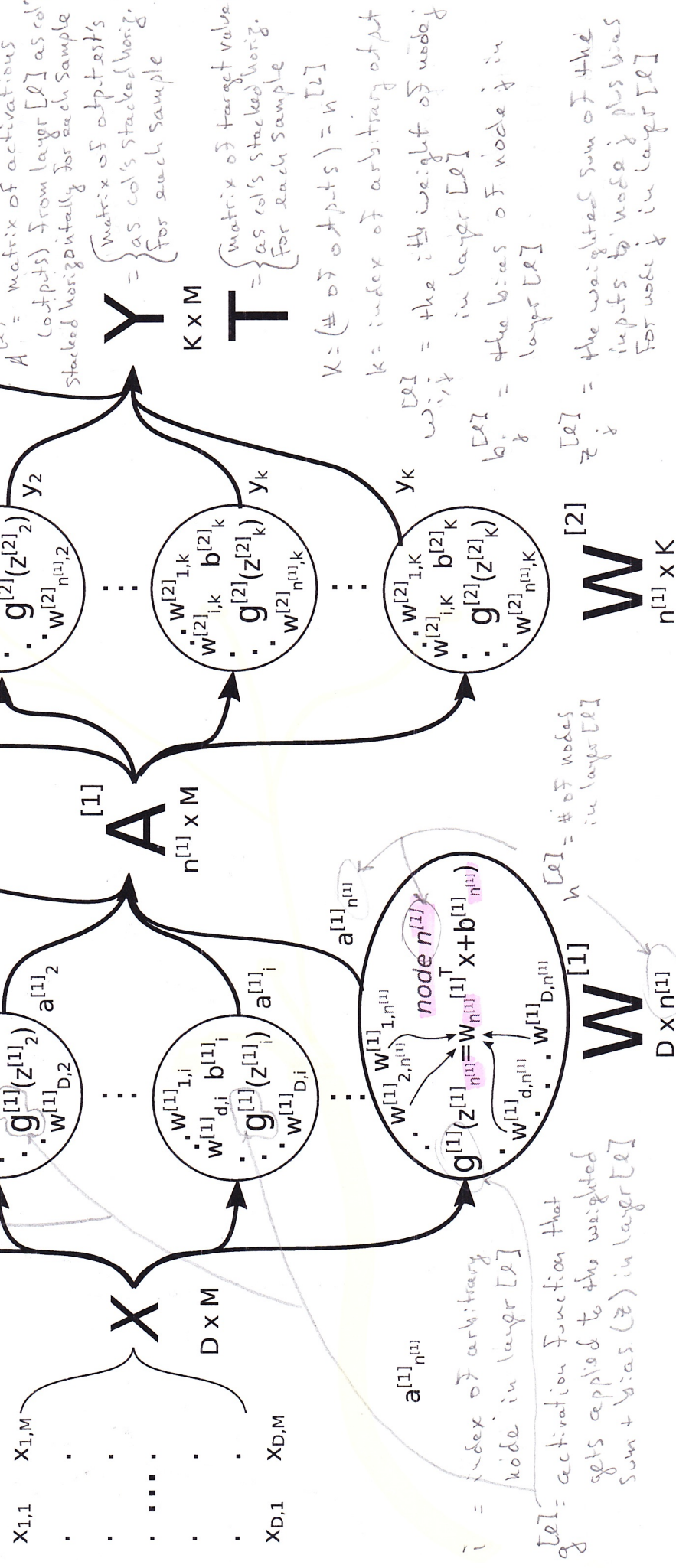


$D = \# \text{ of dimensions / features in the inputs}$
 $d = \text{index of arbitrary dimension / feature}$
 $M = \# \text{ of samples}$
 $m = \text{index of arbitrary sample}$

$L = \# \text{ of layers in the network}$
 $\ell = \text{index of arbitrary layer}$

$y_k = \text{estimate for the } k\text{th output / response}$
 $t_k = \text{target value for the } k\text{th output / response}$
 $a_i^{[2]} = y_1$
 $a_i^{[\ell]} = \text{output activation of node } i \text{ in layer } [\ell]$
 $= g^{[\ell]}(z^{[\ell]})$
 $= g^{[\ell]}(w_j^{[\ell]} a_j^{[\ell-1]} + b_j^{[\ell]})$
 $A^{[\ell]} = \text{matrix of activations (outputs) from layer } [\ell] \text{ as cols stacked horizontally for each sample}$
 $Y = \begin{cases} \text{matrix of output test's as col's stacked horiz.} \\ \text{for each sample} \end{cases}$



$K = (\# \text{ of outputs}) = n^{[2]}$
 $k = \text{index of arbitrary output}$
 $w_{i,j}^{[\ell]} = \text{the } i\text{th weight of node } j \text{ in layer } [\ell]$
 $b_j^{[\ell]} = \text{the bias of node } j \text{ in layer } [\ell]$
 $z_j^{[\ell]} = \text{the weighted sum of the inputs to node } j \text{ plus bias for node } j \text{ in layer } [\ell]$

2-Layer, with Single Hidden Layer Feedforward Neural Network (M samples in columns stacked horizontally)