x_1

x_2

x_3

$+1$

$+1$

$a_1^{(2)}$

$a_2^{(2)}$

$a_3^{(2)}$

$h_\alpha(x)$

Convention: $a_i^{(\ell)}$ is the output of the ith neuron in the ℓth layer.

output layer (K entries - in this case just one).

bias units

input layer
(p entries plus the bias term)

hidden layer
(M entries - in this case 3 - plus bias unit)