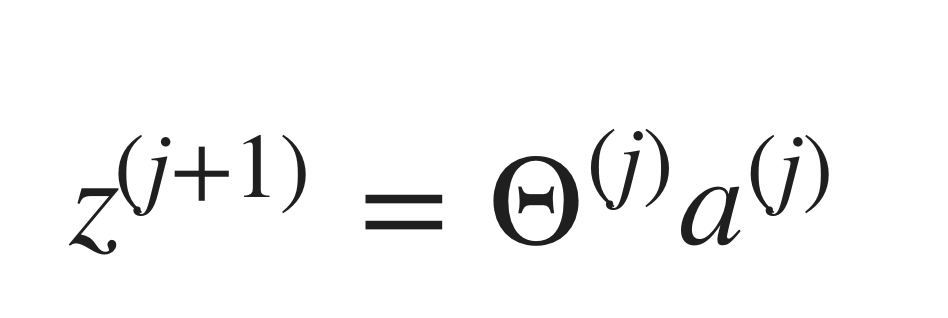
**Neural Network**

We are multiplying our matrix Θ(j−1) with dimensions *sj*​×(*n*+1) (where *sj*​ is the number of our activation nodes) by our vector *a*(*j*−1) with height (n+1). This gives us our vector *z*(*j*) with height *sj*​. Now we can get a vector of our activation nodes for layer j as follows:

A drawing of a person

Description automatically generated

We can then add a bias unit (equal to 1) to layer j after we have computed *a*(*j*). This will be element *a*0(*j*)​ and will be equal to 1. To compute our final hypothesis, let's first compute another z vector:



We get this final z vector by multiplying the next theta matrix after Θ(j−1) with the values of all the activation nodes we just got. This last theta matrix Θ(j) will have only **one row**which is multiplied by one column *a*(*j*) so that our result is a single number. We then get our final result with:

A close up of a logo

Description automatically generated

**Applications**

Where g(z) is the following:

**A close up of a map

Description automatically generated**