# Perceptron and Neural Nets:

## Perceptron:

1. Since its given that any document $d_i$ has no more than 100 words, $||\mathbf{x}_i|| \leq \sqrt{100}$, the maximum being when the word ($b_i \in x_i$) occurs all the 100 times in document $d_i$. Given that $||\mathbf{x}_i||^2 < R^2$
   $\implies$ The upper bound on R is 10.

   The number of mistakes ($k$) that the perceptron makes before converging to a correct hypothesis $v_k$ is given as:
   $k < R^2/\gamma^2$
   Its also given that, there exists some vector u such that $\forall, y_i u.x_i > 1 \implies \gamma = 1$ (since $y_i u.x_i > \gamma$)
   The upper bound on the number of mistakes is just $R^2 = 100$.

2. No, a single percetron can't compute the XOR function, since a perceptron is linear, and you wouls need two perceptrons to compute the XOR function.

## Neural Networks:

Given $i_1 = 1.0$ and $i_2 = 0.0$

1. Output by node $h_1$=
$$\frac{1}{1 + exp(-w_{i_1,h_1} * i_1)} = \frac{1}{1 + exp(-0.5)} = 0.6225$$

2. Output by node $h_2$=
$$\frac{1}{1 + exp(-w_{i_2,h_2} * i_2)} = \frac{1}{1 + exp(0)} = 0.5$$

3. Output by node $o_1$=
$$\frac{1}{1 + exp(-(w_{h_1,o_1} * h_1 + w_{h_2,o_1} * h_2))} = \frac{1}{1 + exp(0.499)} = 0.3778$$

4. Output by node $o_2$=
$$\frac{1}{1 + exp(-(w_{h_1,o_2} * h_1 + w_{h_2,o_2} * h_2))} = \frac{1}{1 + exp(-0.5735)} = 0.6396$$

   Given $t_1 = 1.0$ and $t_2 = 1.0$

5. The value of $\delta_{o_1}$=
$$o_1 * (1 - o_1) * (t_1 - o_1) = 0.1463$$

6. The value of $\delta_{o_2}$=
$$o_2 * (1 - o_2) * (t_2 - o_2) = 0.0831$$

7. The value of $\delta_{h_1}$=
$$o_1 * (1 - o_1) * \sum_{k \in outputs} w_{h1,k}\delta_k = 0.3778 * (1 - 0.3778) * (-0.4 * 0.1463 + 0.6 * 0.0831) = -0.002$$

   Leaning rate of $\eta = 0.1$

8. New value of $w_{h_1,o_1}$=

$$w_{h_1,o_1} + \Delta w_{h_1,o_1} = w_{h_1,o_1} + \eta\delta_{o_1} * h_1 = -0.4 + 0.1 * 0.1463 * 0.6225 = -0.3909$$

9. New value of $w_{h_1,o_2}$=

$$w_{h_1,o_2} + \Delta w_{h_1,o_2} = w_{h_1,o_2} + \eta\delta_{o_2} * h_1 = 0.6 + 0.1 * 0.0831 * 0.6225 = 0.6052$$

10. New value of $w_{i_1,h_1}$=

$$w_{i_1,h_1} + \Delta w_{i_1,h_1} = w_{i_1,h_1} + \eta\delta_{h_1} * i_1 = 0.5 + 0.1 * -0.002 * 0.5 = 0.4999$$