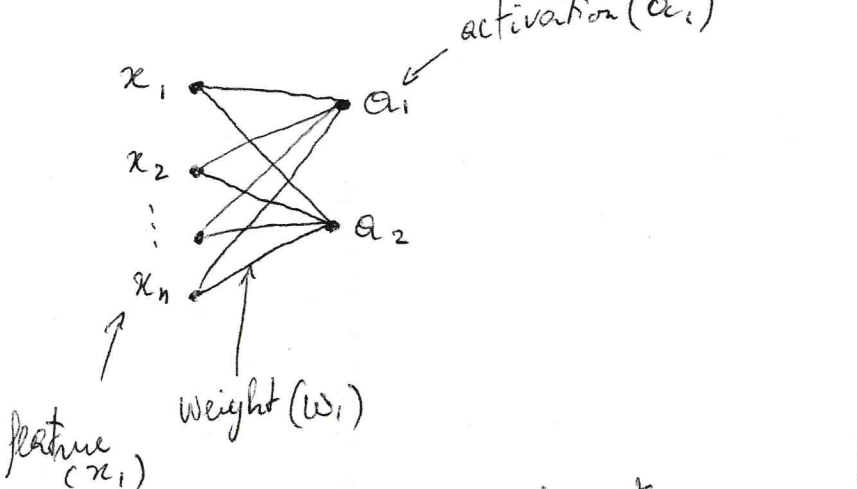
Forward propagation

$f(w_{ij}, x_i) = W \cdot X + b$ (1)

Assuming:



To calculate the activations

$a_1 = w_{1,1} \times x_1 + w_{1,2} \times x_2 + \cdots + w_{1,n} \times x_n + b$

The above applies for all activations

By hand, and that way this is tedious!

Using Matrix notation:

$$\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \vdots & & & \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

weight of size 2×4

data $n \times 1$

bias 2×1

activation 2×1

<u>In the CPU</u>: Individual operation are done at a time

$: w_{1,1} \times x_1 + w_{1,2} \times x_2 \cdots$

The CPU is pretty fast but if the are many rows and columns it, will take time.

<u>In the GPU</u> Batch operations:

Instead of a single value computed, the GPU takes in a whole row multiplied by a column.