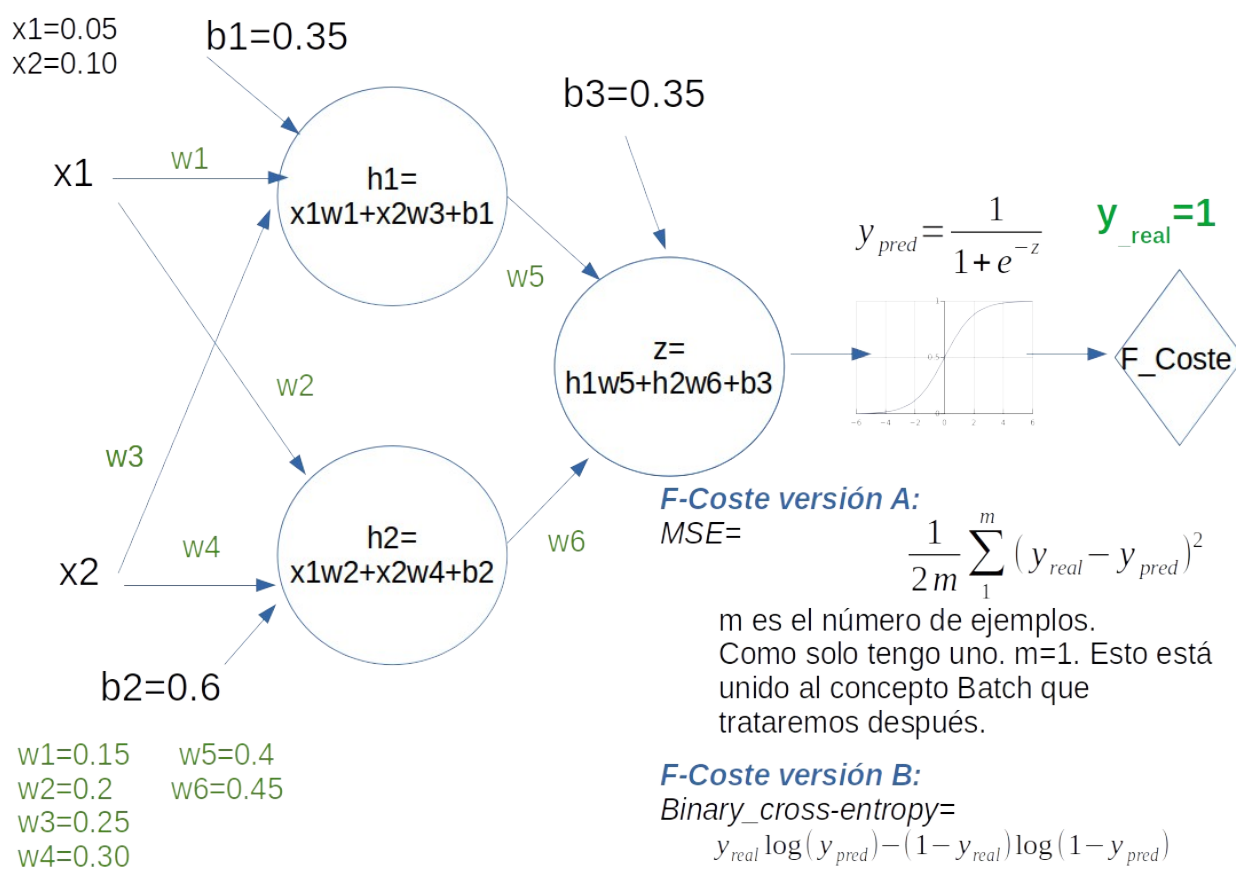


Ejercicio de Clase con Redes Neuronales:

Propagación hacia delante y hacia atrás clasificación binaria:

Calcular la predicción de esta red asumiendo que la función de coste es *MSE* y la clase real es 1. Para ello habrá que realizar lo que se llama una propagación hacia adelante (forward propagation). Es importante apuntar bien los cálculos intermedios ($h_1, h_2, z, y_{pred}, F_Coste$). m es el número de ejemplos que en este caso es 1. Cuando empleamos batches (o agrupaciones de items de entrenamiento) entonces m será el tamaño del batch.

$$MSE = \frac{1}{2m} \sum_1^m (y_{real} - y_{pred})^2$$



Calcular la predicción de esta red asumiendo que la función de coste es *Binary_Cross_entropy*. Para ello habrá que realizar lo que se llama una propagación hacia adelante (forward propagation). Es importante apuntar bien los cálculos intermedios ($h_1, h_2, z, y_{pred}, F_Coste$)

Haz una pasada hacia adelante de esta red asumiendo que la función de coste es

$$\text{Binary_Cross_Entropy} = y_{real} \log(y_{pred}) - (1 - y_{real}) \log(1 - y_{pred})$$

Propagación hacia atrás:

Para llevar a cabo la propagación hacia atrás, es importante entender que la propagación hacia atrás implica emplear el error cometido para las actualizaciones de los parámetros (w) y las derivadas que nos miden “cuanta parte de culpa” tiene una variable con respecto a ese error. Es decir, en cuanto la

variación de esa variable “influye” sobre la función. En nuestro caso queremos saber cuanto ha sido la influencia de w_5 sobre el coste (F-Coste o error) porque esa “cantidad de influencia” es lo que me permitirá actualizar el peso. Así a mayor influencia w_5 sobre F-Coste, mayor debería de ser el cambio, porque ha influido mucho en el error.

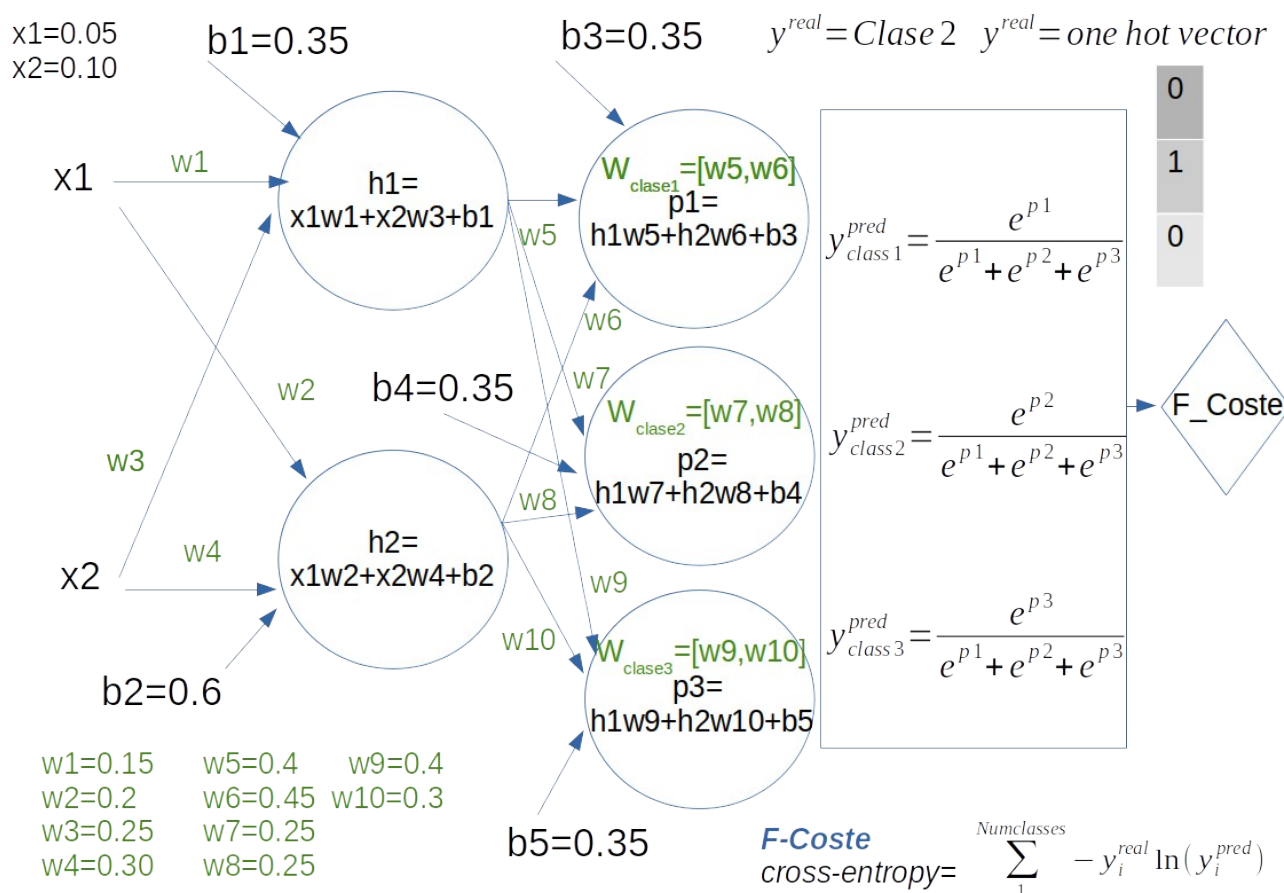
La “cantidad de actualización” de w_5 será:

$$w_5 = w_5 - \eta \nabla_{w_5}(F \text{ Coste})$$

¿Cómo calculo esto?

$$\nabla_{w_5}(F \text{ Coste}) = \frac{\partial(F \text{ Coste})}{\partial(y_{pred})} \frac{\partial(y_{pred})}{\partial(z)} \frac{\partial(z)}{\partial(w_5)}$$

Propagación hacia delante y hacia atrás multiclas:



Recordad que la salida real Y es un one-hot vector que tendrá un 1 en una de las i -s (la que corresponda) y un 0 en las otras. En el ejemplo que nos ocupa tendrá un 1 cuando la i valga 2 y 0 en el resto de las i -s. Así el sumatorio y por lo tanto el valor del error quedará:

$$-0 \ln(y_{pred1}) - 1 \ln(y_{pred2}) - 0 \ln(y_{pred3}) \mid y_{predi} \text{ es el softmax correspondiente a cada clase } i$$