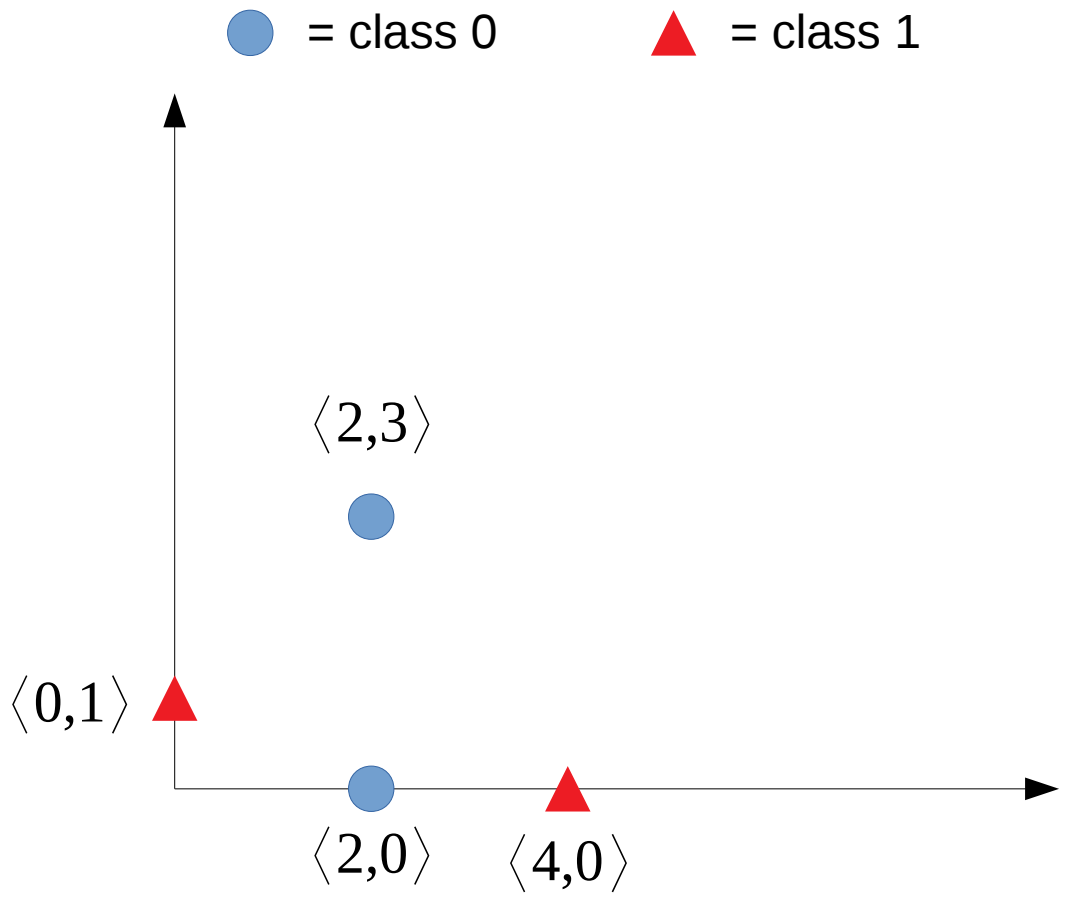
Backprop Demo

MLP Exercise

UNIVERSITY OF LEEDS

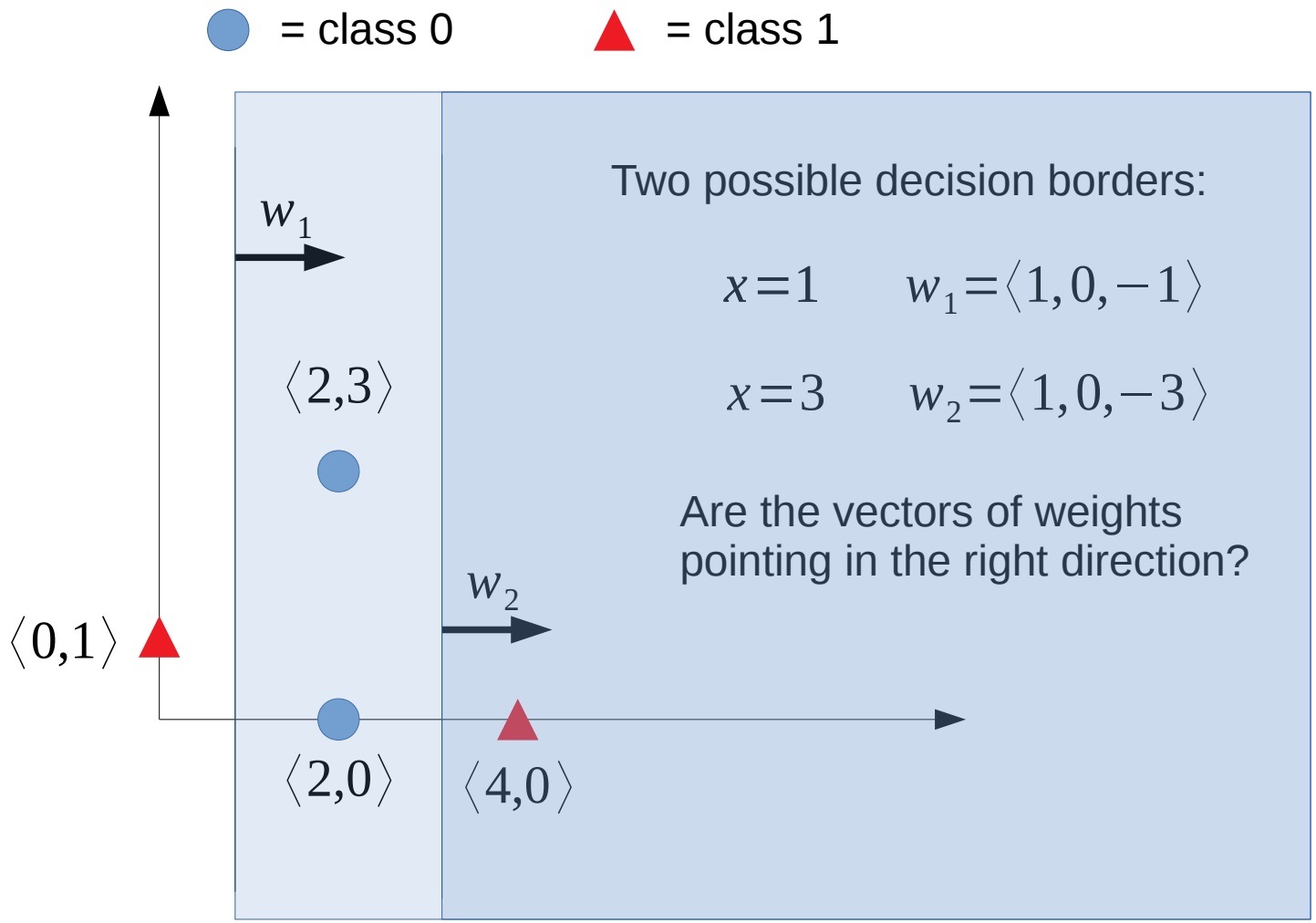
Construct an MLP able to classify the following dataset:



MLP Exercise

UNIVERSITY OF LEEDS

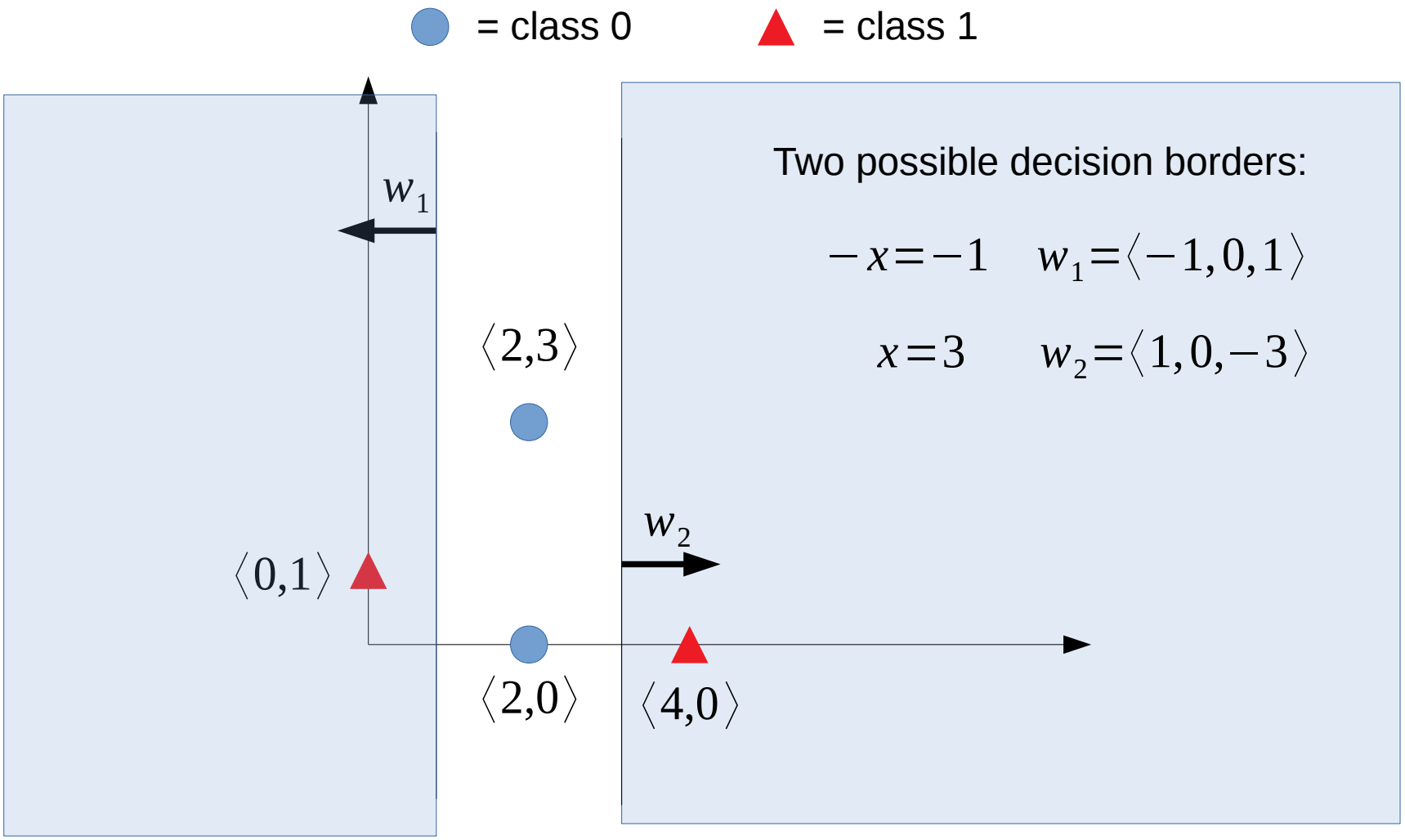
Construct an MLP able to classify the following dataset:



MLP Exercise

UNIVERSITY OF LEEDS

Construct an MLP able to classify the following dataset:

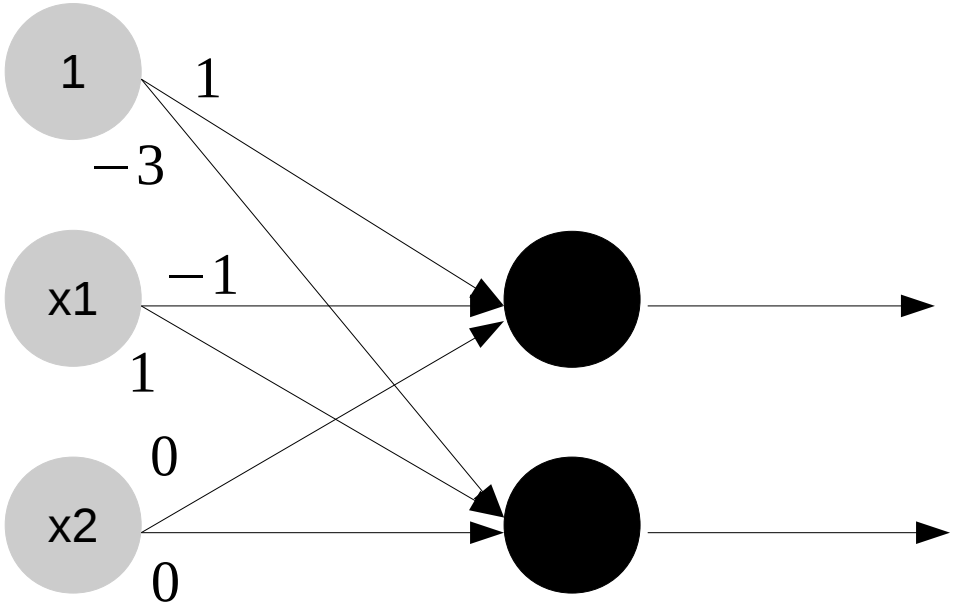


Hidden Layer

UNIVERSITY OF LEEDS

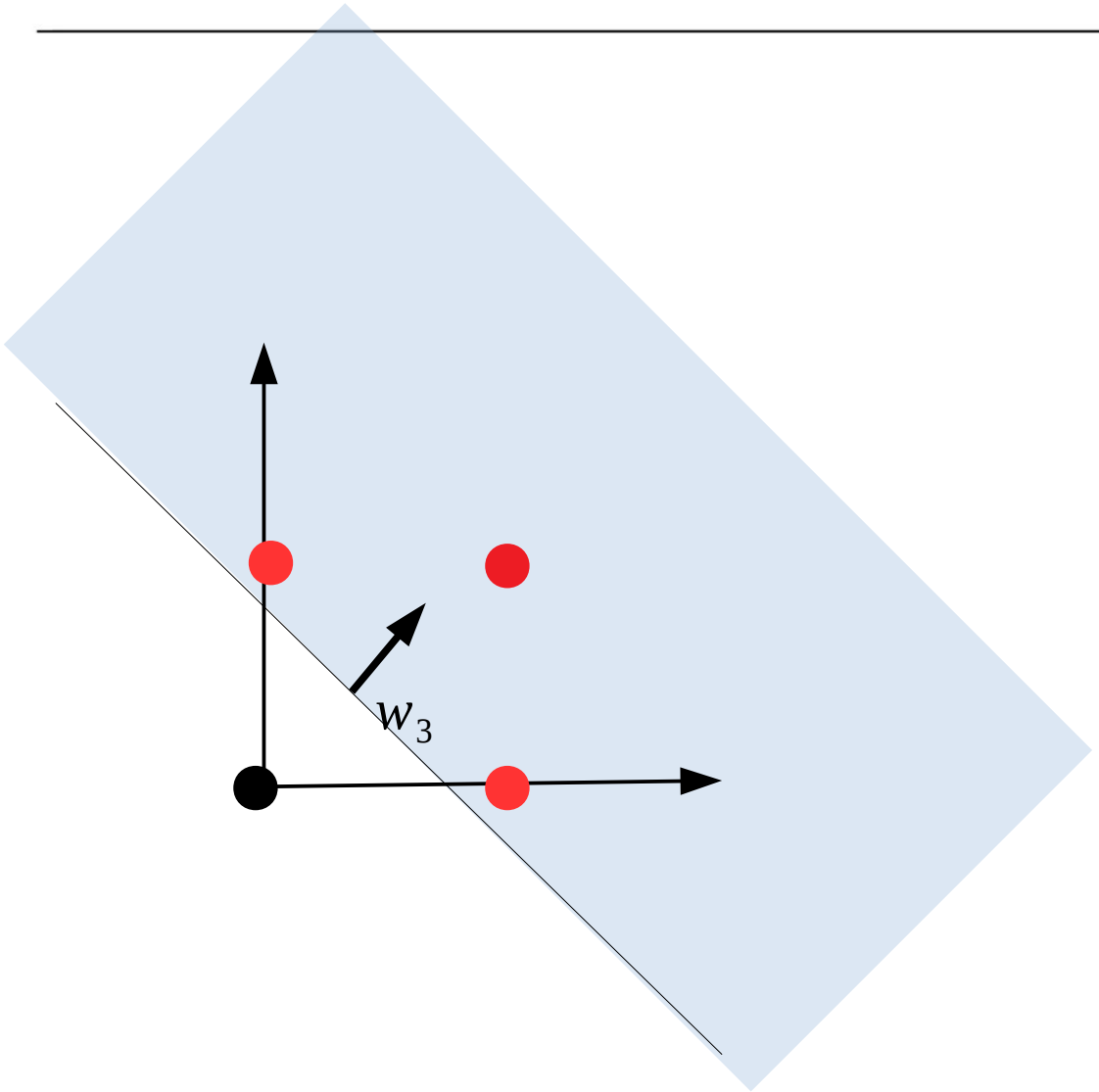
$$w_1 = \langle -1, 0, 1 \rangle$$

$$w_2 = \langle 1, 0, -3 \rangle$$



Output Layer: OR

UNIVERSITY OF LEEDS



$$y = -x + 0.5$$

$$x + y - 0.5 = 0$$

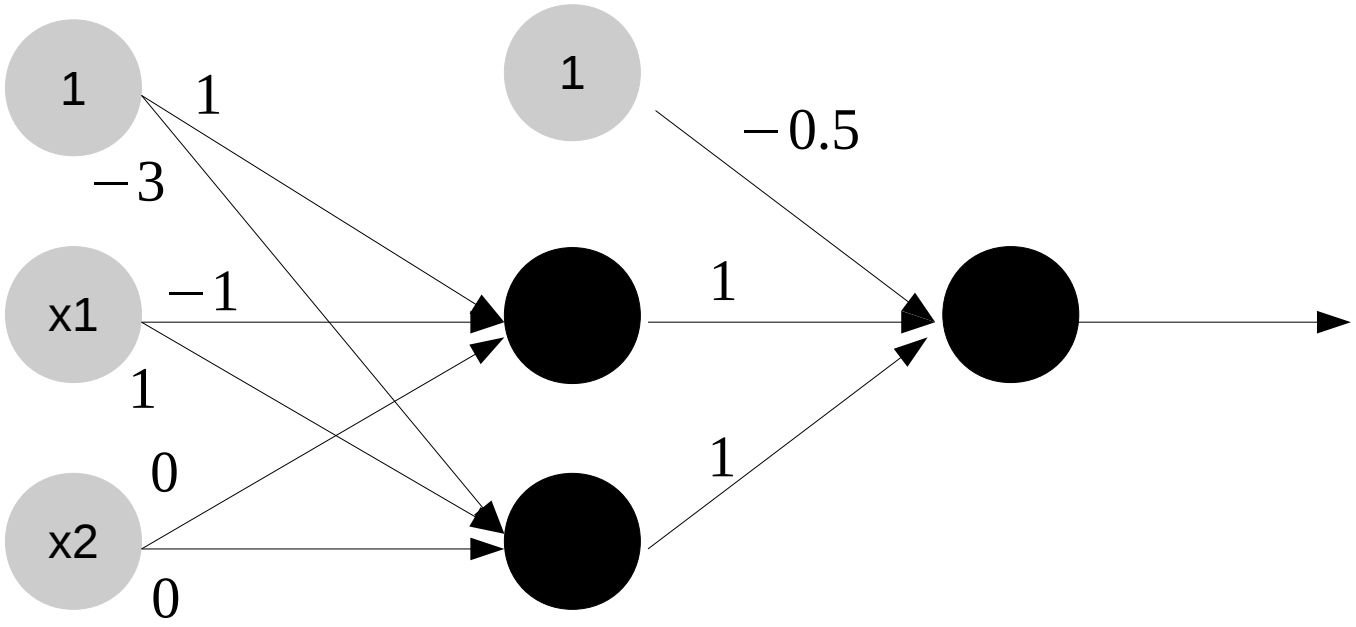
$$w_3 = \langle 1, 1, -0.5 \rangle$$

Final MLP

UNIVERSITY OF LEEDS

$$\boldsymbol{V} = \begin{bmatrix} 1 & -3 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}$$

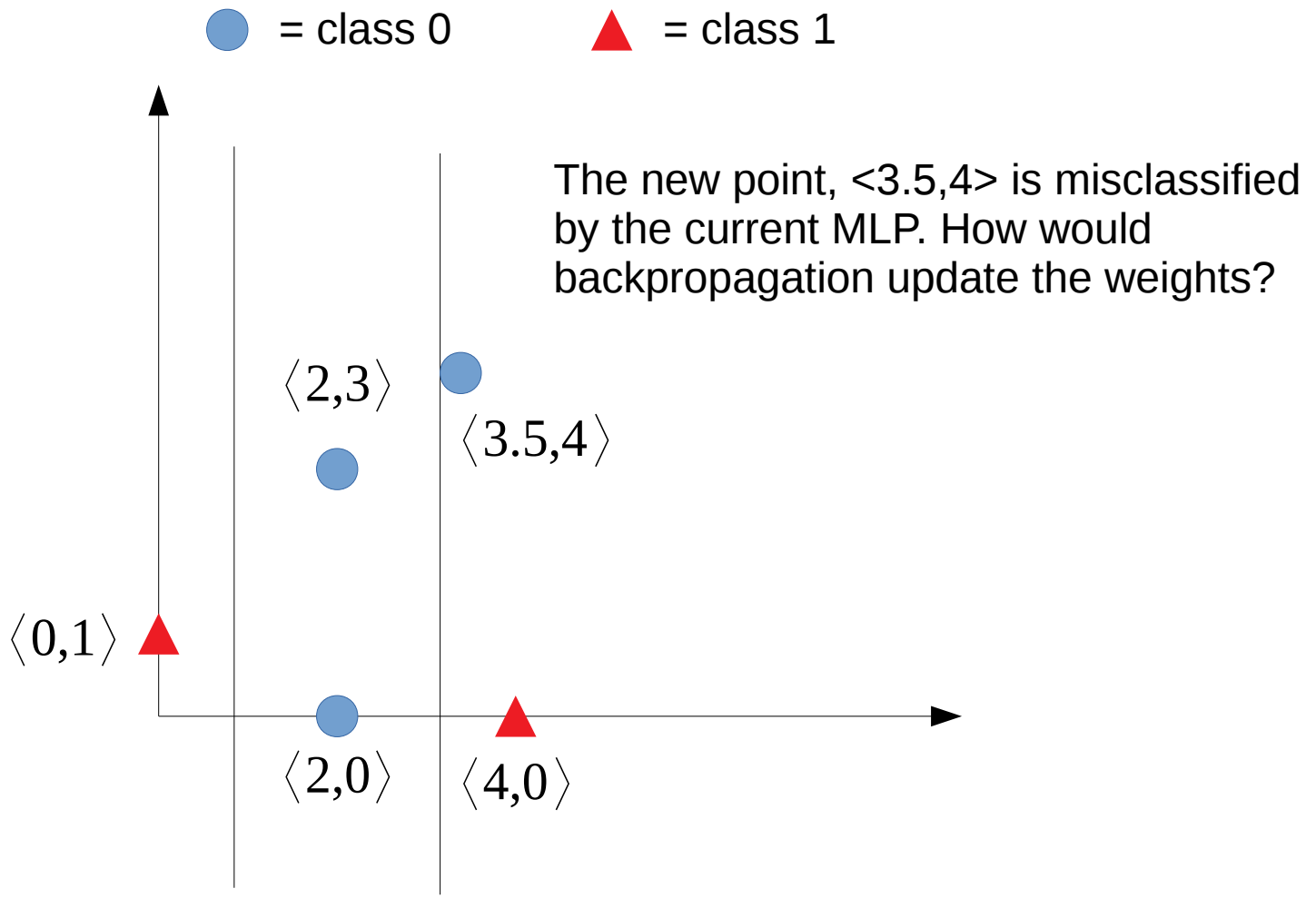
$$\boldsymbol{w} = \begin{bmatrix} -0.5 \\ 1 \\ 1 \end{bmatrix}$$



Adding a new point

UNIVERSITY OF LEEDS

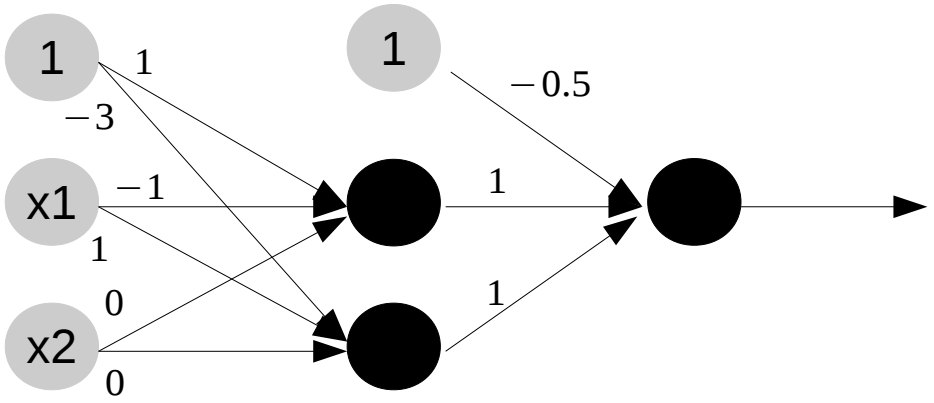
Construct an MLP able to classify the following dataset:



The new point, <3.5,4> is misclassified by the current MLP. How would backpropagation update the weights?

Forward pass

UNIVERSITY OF LEEDS



Forward pass: compute all the z

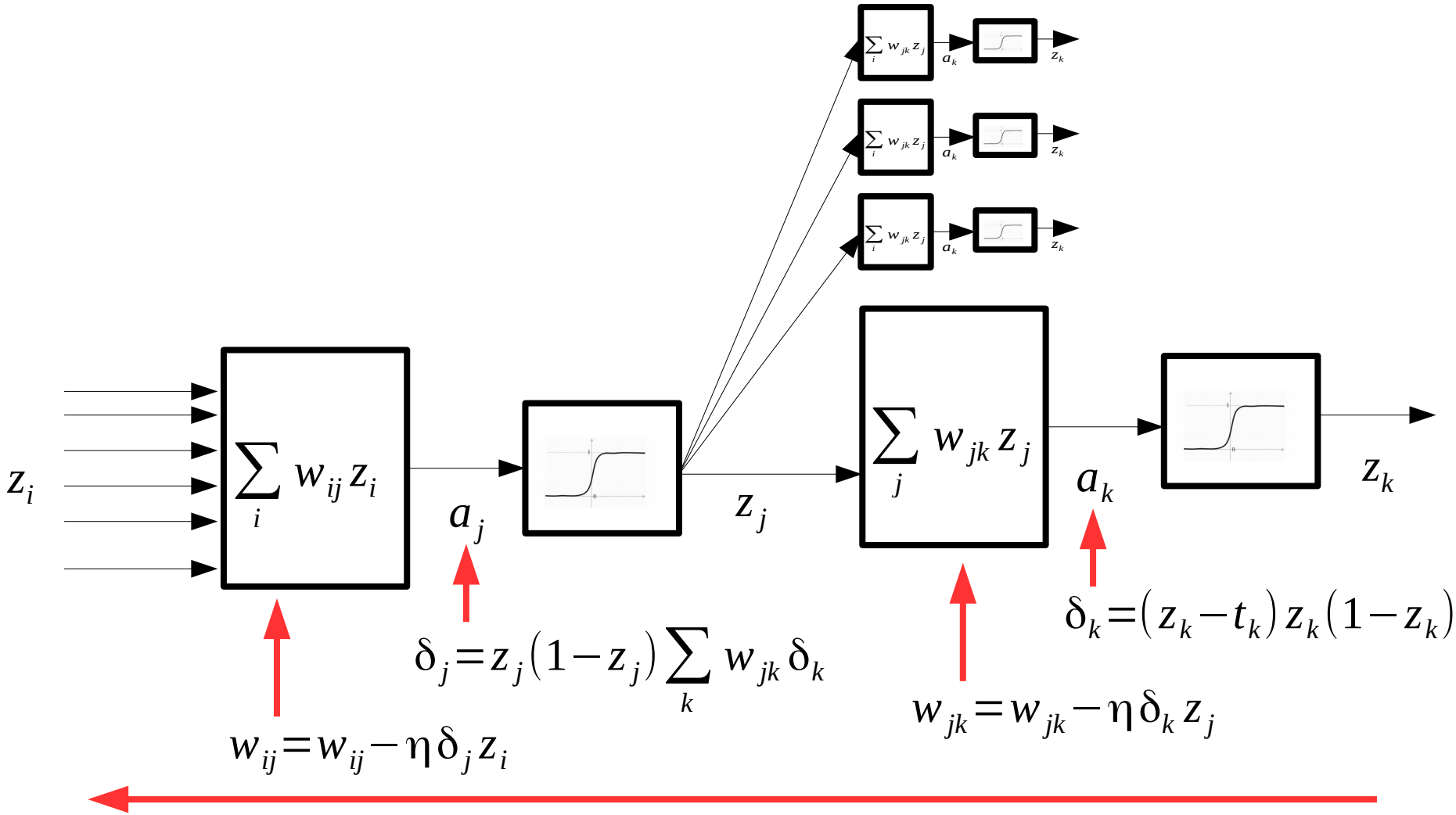
The misclassified point is <3.5,4> and is of class 1.

$$\boldsymbol{V}=\begin{bmatrix}1 & -3\\ -1 & 1\\ 0 & 0\end{bmatrix} \quad \boldsymbol{x}=\begin{bmatrix}1\\ 3.5\\ 4\end{bmatrix} \quad \boldsymbol{z}_h=\sigma(\boldsymbol{V}^T\boldsymbol{x})=\sigma\left(\begin{bmatrix}-2.5\\ 0.5\end{bmatrix}\right)=\begin{bmatrix}3.73\text{e-}06\\ 0.924\end{bmatrix}$$

$$\boldsymbol{w}=\begin{bmatrix}-0.5\\ 1\\ 1\end{bmatrix} \quad z_o=\sigma\left(\boldsymbol{w}^T\begin{bmatrix}1\\ \boldsymbol{z}_h\end{bmatrix}\right)=\sigma(0.424)=0.892$$

Computing delta

UNIVERSITY OF LEEDS



Forward pass

UNIVERSITY OF LEEDS

$$\boldsymbol{V}=\begin{bmatrix}1 & -3\\ -1 & 1\\ 0 & 0\end{bmatrix} \qquad \boldsymbol{z}_h=\begin{bmatrix}1\\ 3.73\text{e-}06\\ 0.924\end{bmatrix} \qquad \boldsymbol{w}=\begin{bmatrix}-0.5\\ 1\\ 1\end{bmatrix} \qquad z_o=0.892$$

$$\delta_o=\beta\left(z_o-t\right)z_o\left(1-z_o\right)=5\left(0.892-0\right)0.892\left(1-0.892\right)=0.427$$

$$w=w-\eta\,\delta_o\,z_h=\begin{bmatrix}-0.5\\ 1\\ 1\end{bmatrix}-0.5\cdot\left(0.427\right)\cdot\begin{bmatrix}1\\ 3.73\text{e-}06\\ 0.924\end{bmatrix}=\begin{bmatrix}-0.59\\ 1\\ 0.908\end{bmatrix}$$

$$\delta_h^{(2)}=\beta\,z_h\left(1-z_h\right)\sum_i w_i\,\delta_o=5\cdot0.924\cdot\left(1-0.924\right)\cdot1\cdot\left(0.427\right)=0.12$$

$$\boldsymbol{v}_2=\boldsymbol{v}_2-\eta\,\delta_h^{(2)}\boldsymbol{x}=\begin{bmatrix}-3.06006\\ 0.789778\\ -0.240254\end{bmatrix}$$