11/22 AI homewok 7107029022 邱靖詒

畫出 BPN 向前傳遞(Forward pass)與向後傳遞(Backward pass)方法與計算。

向前傳遞是先將訓練資料丟進網絡去跑，在計算出輸出結果與對應目標之間誤差，而向後傳遞是依誤差值去調整網絡權重，經過多次訓練後，就會將網絡修正到誤差極小範圍內的輸出結果。

- 向前傳遞(Forward pass)

  1. 使用 sigmoid 為激活函數：

$$\boldsymbol{f(x) = \frac{1}{1 + e^{-x}}}$$

  2. 以均平方誤差(mean squared error,MSE)為計算誤差值方法：

$$\boldsymbol{MSE = \frac{(O_desired - O_actual)^2}{2}}$$

- 向後傳遞(Backward pass)

  1. 計算輸出層及隱藏層誤差值：

$$\boldsymbol{\delta_o = (C_i - u_i)u_i(1 - u_i)}$$

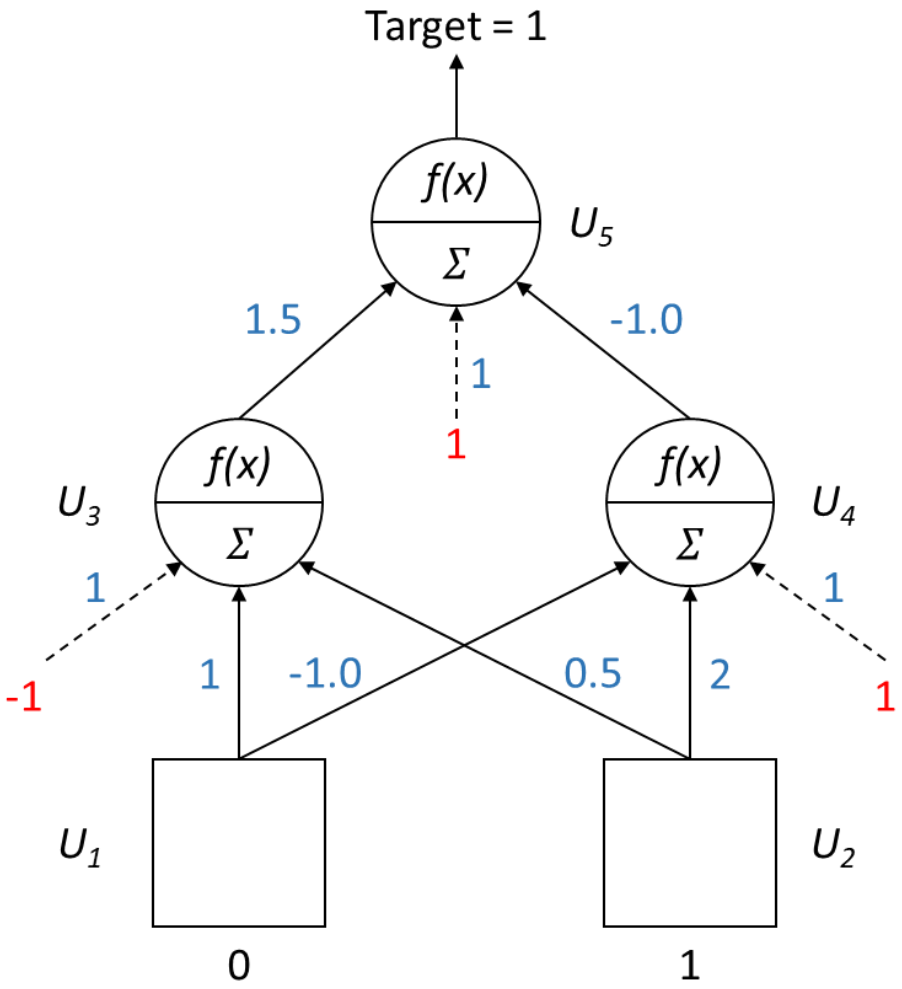
$$\boldsymbol{\delta_i = \left(\sum_{m\ :\ m\ >\ i} w_{m,j}\delta_o \right) u_i(1 - u_j)}$$

  2. 調整隱藏到輸出層和隱藏層到輸入層的權重：

$$\boldsymbol{w^*_{\ i,j} = w_{i,j} + \ \rho\delta_o u_i}$$

$$\boldsymbol{w^*_{\ i,j} = w_{i,j} + \ \rho\delta_i u_i}$$

以下圖作為例子，



■ **計算向前傳遞(Forward pass)**

計算各隱藏層的值和 MSE。

$u_3 = f(w_{3,1}u_1 + w_{3,2}u_2 + w_{3,b}*bias_3)$
$= f(1*0 + 0.5*1 + 1*-1)$
$= f(-0.5)$
$\fallingdotseq 0.377541$

$u_4 = f(w_{4,1}u_1 + w_{4,2}u_2 + w_{4,b}*bias_4)$
$= f(-1*0 + 2*1 + 1*1)$
$= f(3)$
$\fallingdotseq 0.952574$

$u_5 = f(w_{5,3}u_3 + w_{5,4}u_4 + w_{5,b}*bias_5)$
$= f(1.5*0.377541 + -1*0.952574 + 1*1)$
$= f(0.613738)$
$\fallingdotseq 0.648793$

$MSE = 0.5*(1.0 - 0.648793)^2 \fallingdotseq 0.0616733$

■ 計算向後傳遞(Backward pass)

計算出輸出層和隱藏層誤差值，在調整輸出層和隱藏各權重值及偏差值。(學習率為ρ=0.5)

δ_o = (1.0 − 0.648793) * 0.648793 * (1.0 − 0.648793)
= 0.080026

δ_{u4} = (δ_o* $w_{5,4}$) * u_4 * (1.0 − u_4)
= (0.080026 * −1.0) * 0.952574 * (1.0 - 0.952574)
= −0.00361531

δ_{u3} = (δ_o* $w_{5,3}$) * u_3 * (1.0 − u_3)
= (0.080026 * 1.5) * 0.377541 * (1.0 − 0.377541)
= 0.0282096

$w_{5,4}$ = $w_{5,4}$ + (ρ * δ_o * u_4)
= −1 + (0.5 * 0.080026 * 0.952574)
= −0.961885

$w_{5,3}$ = $w_{5,3}$ + (ρ * δ_o * u_3)
= 1.5 + (0.5 * 0.080026 * 0.377541)
= 1.51511

$w_{5,b}$ = $w_{5,b}$ + (ρ * δ_o * $bias_5$)
= 1 + (0.5 * 0.080026 * 1)
= 1.04001

$w_{4,2}$ = $w_{4,2}$ + (ρ * δ_{u4} * u_2)
= 2 + (0.5 * −0.00361531 * 1)
= 1.99819

$w_{4,1}$ = $w_{4,1}$ + (ρ * δ_{u4} * u_1)
= −1 + (0.5 * 0.00361531 * 0)
= −1.0

$w_{4,b}$ = $w_{4,b}$ + (ρ * δ_{u4} * $bias_4$)
= 1.0 + (0.5 * 0.00361531 * 1)
= 0.998192

$w_{3,2}$ = $w_{3,2}$ + (ρ * δ_{u3} * u_2)
= 1.0 + (0.5 * 0.0282096 * 1)
= 0.514105

$w_{3,1}$ = $w_{3,1}$ + (ρ * δ_{u3} * u_1)
= −1 + (0.5 * 0.0282096 * 0)
= −1.0

$w_{3,b}$ = $w_{3,b}$ + (ρ * δ_{u3} * $bias_3$)
= 1.0 + (0.5 * 0.0282096 * −1)
= 0.986895

經過調整後，再跑一次向前傳遞

$u_3 = f(w_{3,1}u_1 + w_{3,2}u_2 + w_{3,b}*bias_3)$
$= f(1*0 + 0.514105*1 + 0.985895*-1)$
$= f(-0.47179)$
$\fallingdotseq 0.384193$

$u_4 = f(w_{4,1}u_1 + w_{4,2}u_2 + w_{4,b}*bias_4)$
$= f(-1.0*0 + 1.99819*1 + 0.998192*1)$
$= f(2.99638)$
$\fallingdotseq 0.952411$

$u_5 = f(w_{5,3}u_3 + w_{5,4}u_4 + w_{5,b}*bias_5)$
$= f(1.51511*0.377541 + -0.961885*0.952574 + 1.04001*1)$
$= f(0.705995)$
$\fallingdotseq 0.669516$

$MSE = 0.5*(1.0 - 0.669516)^2 \fallingdotseq 0.0546099$

結果顯示，重新計算後 MSE 有明顯變低。