

類神經網路訓練不起來怎麼辦 (四)：損失函數 (Loss) 也可能有影響

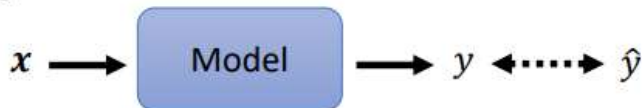
Create at 2022/06/08

- 類神經網路訓練不起來怎麼辦 (四)：損失函數 (Loss) 也可能有影響
 - Create at 2022/06/08
 - Classification (Short Version)
 - 課程網頁
- 上課資源：
 1. 類神經網路訓練不起來怎麼辦 (四)：損失函數 (Loss) 也可能有影響
(<https://www.youtube.com/watch?v=O2Vkp8dJ5FE>).
- 延伸參考資源：
 1. ML Lecture 4: Classification (<https://www.youtube.com/watch?v=fZAZUYEeIMg>).
 2. ML Lecture 5: Logistic Regression (<https://www.youtube.com/watch?v=hSXFuyyLukA>).

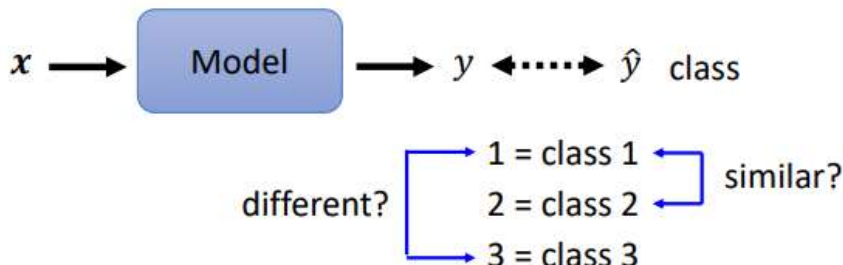
Classification (Short Version)

Classification as Regression?

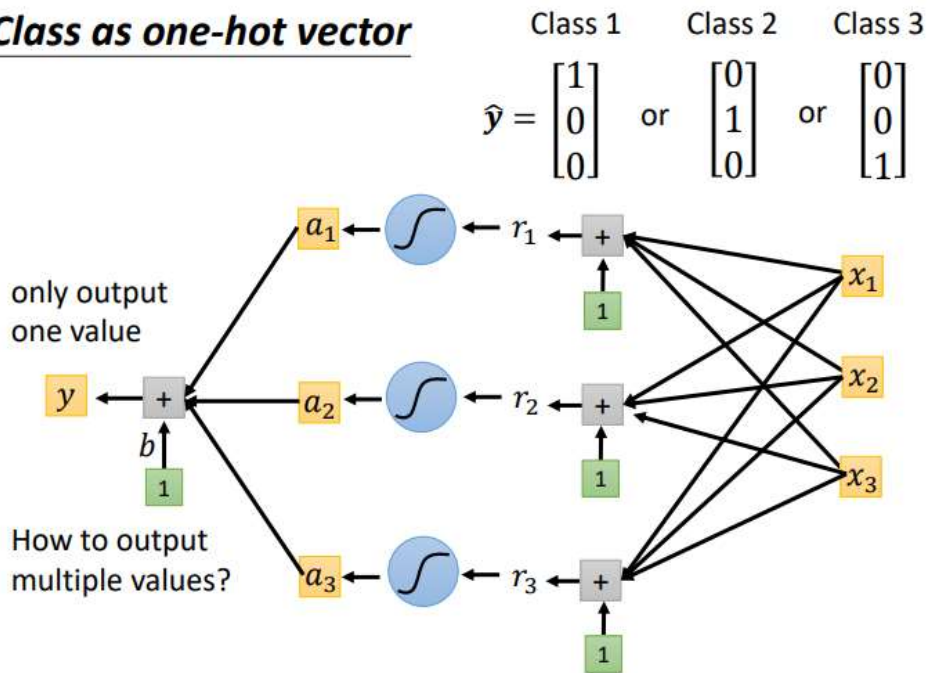
• Regression



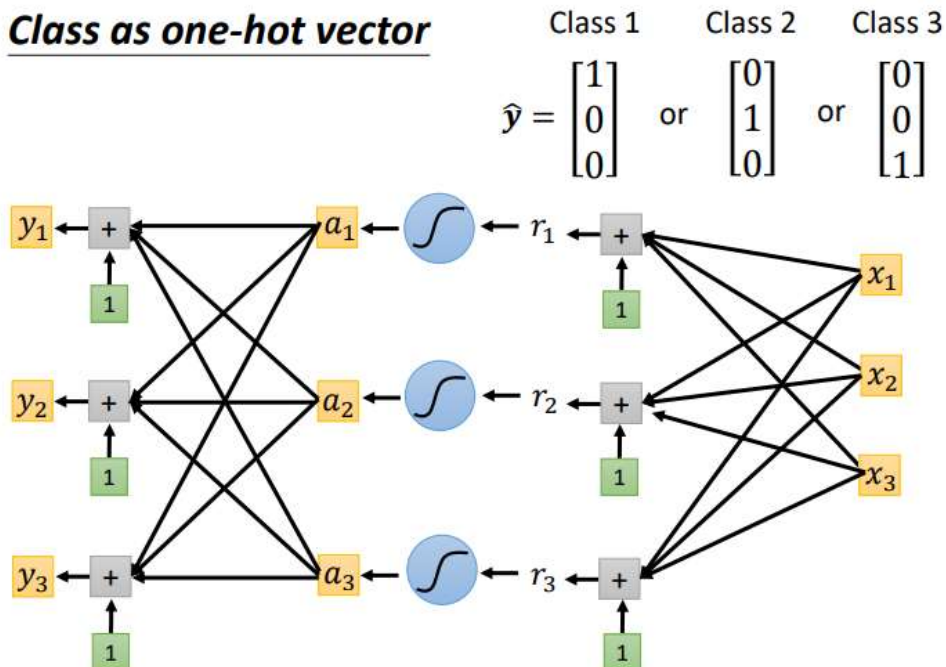
• Classification as regression?



- Regression 就是輸入一個向量，然後輸出一個數值
- 希望輸出的數值跟某一個 label 越接近越好

Class as one-hot vector

- 把 class 用 one-hot vector 表示
 - 不會有 class 1 跟 class 2 比較接近，class 1 跟 class 3 比較遠的問題
 - class 之間的距離都是一樣的
- 把原本 regression 只 output 一個數值，改成可以 output 三個數值
 - 如下頁投影片所示

Class as one-hot vector

- 輸入一個 feature 的 vector，輸出一個 vector
- 期待輸出離目標越接近越好

Regression

label

$$\hat{y} \longleftrightarrow y = b + c^T \sigma(b + Wx)$$

feature

Classification

feature

$$y = b' + W' \sigma(b + Wx)$$

label

$$\hat{y} \longleftrightarrow y' = \text{softmax}(y)$$

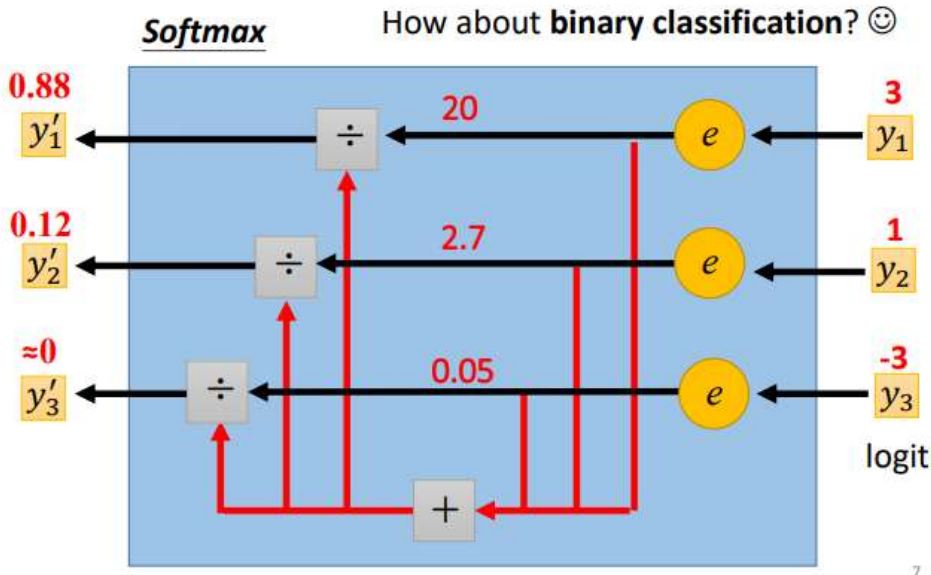
0 or 1 Make all values between 0 and 1 Can have any value

Regression 跟 Classification
個別的運算流程

- softmax 簡單解釋成：把 y Normalize 移到 0 1 之間

Soft-max

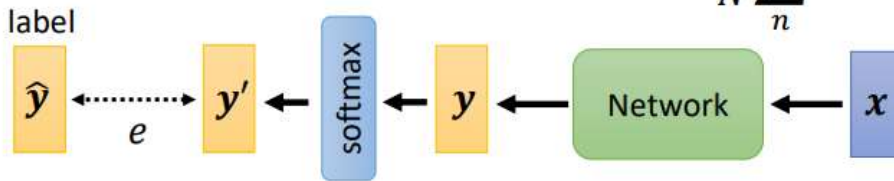
$$y'_i = \frac{\exp(y_i)}{\sum_j \exp(y_i)} \quad \begin{cases} 1 > y'_i > 0 \\ \sum_i y'_i = 1 \end{cases}$$



- softmax 的 block
 - 可以讓大的值與小的值差距更大
 - softmax 的輸入會叫他 logit
- 當 2 個 class 時用 sigmoid 跟用 softmax 是等價的

Loss of Classification

$$L = \frac{1}{N} \sum_n e_n$$



Mean Square Error (MSE) $e = \sum_i (\hat{y}_i - y'_i)^2$

Cross-entropy  $e = - \sum_i \hat{y}_i \ln y'_i$

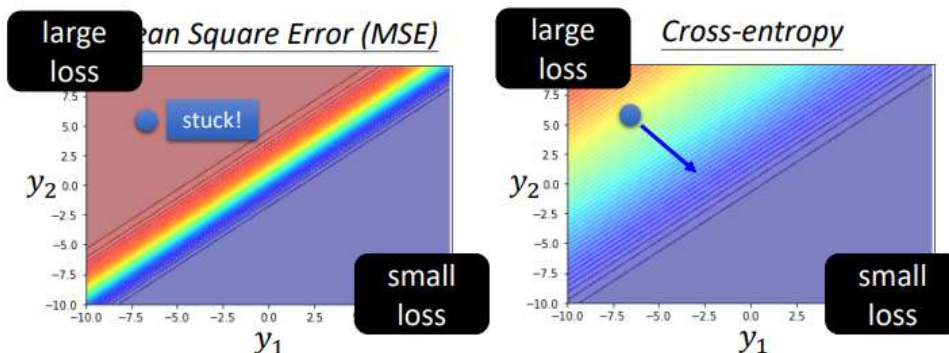
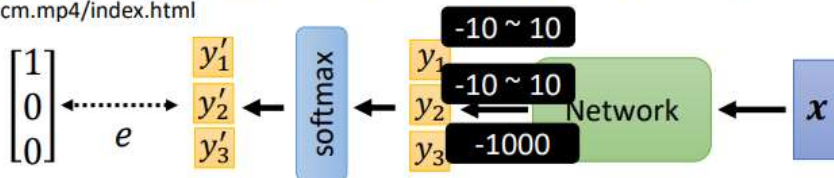
Minimizing cross-entropy is equivalent to maximizing likelihood.

計算 y' 跟 \hat{y} 之間的距離有多種方式

1. MSE
2. Cross-entropy

Minimizing cross-entropy is **equivalent** to maximizing likelihood.

[http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Deep%20More%20\(v2\).ecm.mp4/index.html](http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Deep%20More%20(v2).ecm.mp4/index.html)



Changing the loss function can change the difficulty of optimization.

證明 Cross-entropy 比 MSE 更加適合用在分類上
pytorch 把 softmax 綁在 cross-entropy 裡面

- Loss function 的定義也可能影像 training 容不容易
- 改 loss function 可以改變 optimization 的難度

課程網頁 (<https://speech.ee.ntu.edu.tw/~hylee/ml/2022-spring.php>).

tags: 2022 李宏毅_機器學習