

# HW4 報告

## 一、問題描述：

分別計算出以下兩個模型預測結果的 Mean Square Error(MSE), Mean Absolute Error(MAE), Cross-Entropy。

Model 1		Probability			Answer		
Target		Dog	Cat	Other	Dog	Cat	Other
Data 1	Dog	0.4	0.3	0.3	1	0	0
Data 2	Cat	0.3	0.4	0.3	0	1	0
Data 3	Dog	0.5	0.2	0.3	1	0	0
Data 4	Other	0.6	0.2	0.2	0	0	1

Table 1. 模型一分類結果

Model 2		Probability			Answer		
Target		Dog	Cat	Other	Dog	Cat	Other
Data 1	Dog	0.8	0.1	0.1	1	0	0
Data 2	Cat	0.1	0.7	0.2	0	1	0
Data 3	Dog	0.7	0.1	0.2	1	0	0
Data 4	Other	0.4	0.3	0.3	0	0	1

Table 2. 模型二分類結果

## 二、方法介紹與結果：

### 1. 利用 Scikit-learn 函數庫：

從 sklearn.metrics 中，可以分別引入 mean\_squared\_error,

mean\_absolute\_error, log\_loss 三個函數，便可以分別算出兩個模型預測結果的損失，結果如 Table 3 所示。

Model	MSE	MAE	Cross-Entropy
1	0.208	0.417	1.038
2	0.09	0.25	0.535

Table 3. 使用內建的損失函數計算模型一跟模型二的 MSE, MAE, Cross-Entropy

## 2. 從公式出發，手刻損失函數：

根據損失函數的數學定義(如 Equation 1 ~ Equation 3 所示)，分別手刻出 MSE, MAE, Cross-entropy 這三個函數，計算出模型的損失(如 Table 4 所示)，並與 Table 3 的結果進行比對。

$$MSE = \sum_{i=1}^n (y_i - y_i^p)^2$$

Equation 1. MSE 的損失函數公式

$$MAE = \sum_{i=1}^n |y_i - y_i^p|$$

Equation 2. MAE 的損失函數公式

$$Cross - Entropy = - \sum_{i=1}^N y_i * \log(y_i^p)$$

Equation 3. Cross-Entropy 的損失函數公式

Model	MSE	MAE	Cross-Entropy
1	0.208	0.417	1.038
2	0.09	0.25	0.535

Table 4. 使用手刻的損失函數計算模型一跟模型二的 MSE, MAE, Cross-Entropy

### 三、 結論：

從 Table 1 與 Table 2 可以觀察到，雖然在此分類問題中，兩個模型預測出了相同的分類結果(最終分類結果為表格中的紅字部份)。但模型一卻明顯擁有較大的損失。這是因為模型一的預測結果中，對預測各類別發生的機率值相當接近，而模型二卻對正確的類別有較高的預測值，代表模型二對於正確的類別有更好的擬合能力，因此模型二不管是以 MSE, MAE, Cross-Entropy 都有較低的損失。

那既然不管使用哪個損失函數，較優異的模型二都能有較低的損失，為何分類模型上還需要 Cross-Entropy 作為常用的損失函數呢?那是因為根據公式，只有當  $y_i=1$ (表示為該資料點的答案)時，對應的  $y_{ip}$  才會被納入損失，分類模型的輸出為機率，因此對應的  $y_{ip}$  若是越大(越靠近 1，代表分類的很好)，則根據 Cross-Entropy 公式，取  $-\log$  之後的值會越小，這種計算標準可以更好的將注意力放在正確的類別上，而不受其他的類別所影響。而對於錯誤的分類，Cross-Entropy 的懲罰力度也更大(機率越靠近 0，取  $-\log$  之後的值也越大)，因此在分類問題中，我們會使用 Cross-Entropy 作為衡量的標準。