

Hw3_report

110062238 潘茗脩

the difficulty I encountered:

我一開始寫 basic 的 min_max part 的時候把所有 X 的 data 的 max 跟 min 拿去計算導致整個 Accuracy 上不去。而在 advanced part 的時候，我對 softmax 的理解不是非常透徹，還有一開始切 validation data 時切錯導致 Accuracy 時常出現錯誤。

How to solve the difficulty:

經過與同學討論與查資料後，發現應該是要將每個 feature 個別做 min_max，而不是整個 X。以及後來發現 advanced 中的 X 應該是 6000 個 1、6000 個 2、...、6000 個 9，而照以前的切法(直接拆成 8 2)會造成根本沒有 9 的情況發生，導致訓練失敗。

Structure of my model:

	basic	advanced
Layers_dims	[30, 108, 54, 27, 1]	[784, 888, 444, 222, 111, 56, 10]
Activation_fn	["relu", "relu", "relu", "sigmoid"]	["relu", "relu", "relu", "relu", "relu", "sigmoid"]
Learning_rate	0.0993652	0.62211
Num_iterations	3269	500
Batches_size	nan	64
classes	2	10

How I improve my model:

使用老師上課教的 early stopping 的方法。我後來發現應該要適度的犧牲 training data 的準確度來讓 validation data 更加準確，其實就是避免 overfitting。

my reflections:

這次的作業讓我感受到訓練神經網路的難度，不過助教在討論區回覆問題的速度也十分之快，讓我少走了更多彎路。