

1.

Epochs\α	0.001	0.01	0.1
100	Epoch 100/100, Train loss: 0.3689 Train acc:85.1852% Val loss: 0.4905 Val acc: 79.0123% Test accuracy is 70.96774193548387%	Epoch 100/100, Train loss:0.3014 Train acc:86.7725% Val loss: 0.4989 Val acc: 79.0123% Test accuracy is 74.19354838709677%	Epoch 100/100, Train loss:0.6880 Train acc:55.0265% Val loss: 0.6903, Val acc: 53.0864% Test accuracy is 48.38709677419355%
500	Epoch 500/500, Train loss: 0.2542 Train acc:88.8889% Val loss: 0.4568, Val acc: 82.7160% Test accuracy is 77.41935483870968%	Epoch 500/500, Train loss:0.0440 Train acc:99.4709% Val loss: 1.7211 Val acc: 74.0741% Test accuracy is 87.09677419354838%	Epoch 500/500, Train loss:0.6906, Train acc:53.4392% Val loss: 0.6869, Val acc: 56.7901% Test accuracy is 48.38709677419355%

2.

從結果可以發現，**learning rate** 在三個水準下的模型表現大相逕庭。當 **learning rate = 0.1** 時，可能是因為參數更新的步長較大，導致模型會有震盪的現象，無法收斂至最低點。當 **learning rate = 0.001** 時，可能是因為參數更新的步長較小，模型收斂速度較慢，在 **epochs** 不足下，無法快速收斂至最低點。相較之下，**learning rate = 0.01** 時，模型的表現最好。當 **epochs** 從 100 調至 500 時，模型的表現幾乎都較原先好，因為模型訓練的輪次變多。不過，在 **learning rate = 0.01** 時，因為 **epochs** 調高導致了 **overfitting** 的狀況發生。

3.

訓練集與測試集在 **accuracy** 指標差異大的原因來自於訓練模型時發生 **overfitting** 的狀況，導致模型訓練時看似能準確地將資料分類，但實際上只是因為模型過度擬合資料，在跑測試集後會發現模型的 **accuracy** 會與在訓練集時差很多，造成 **high variance** 的現象。

4.

特徵選擇可以顯著影響模型的效能和泛化能力，決定了機器學習的上限，而模型只是逼近這個上限，可見特徵選擇的重要。以下為特徵選擇的方法：

過濾法：透過單獨評估每個特徵與目標變數之間的相關性進行特徵選擇。其優點為計算簡單、速度快。

包裝法：根據目標函數，每次選擇幾個特徵，或者排除幾個特徵，利用模型評估出的分數進行特徵的篩選，優點為預測較精確，缺點為成本高。

嵌入法：嵌入方法結合了過濾和包裝方法的優點，能夠在模型訓練過程中自動選擇最相關的特徵。

5.

TabNet 是專門處理表格數據的模型，具有可解釋性與高效的執行效能等優點。此模型可直接使用表格數據，不需要預先處理，使用基於梯度下降的最佳化方法，使它能方便地加入 end-to-end 的模型中。TabNet 採用順序注意力的機器學習技術，用戶可以了解模型中的每一個步驟，特徵選擇的原因，這個機制可以解釋模型獲得最終預測結果的過程，還可用來改進模型，增加精準度。

Reference:

- [ChatGPT](#)
- [ML100Days-030 特徵選擇](#)
- [\[資料分析&機器學習\] 第 2.4 講：資料前處理\(Missing data, One-hot encoding, Feature Scaling\)](#)
- [论文阅读：GBDT 能否被深度学习取代——TabNet](#)
- [Google 雲端 AI 平臺內建 TabNet 表格模型，可替代傳統決策樹演算法](#)