

1. 超參數選擇以及比較

表 1 各個學習率以及神經網路層數在測試集的準確率

學習率				
神經網路的層數		0.001	0.005	0.007
	3	77.41935	80.64516	83.87097
	4	70.96774	77.41935	80.64521
	5	67.74194	80.64516	77.41935
	6	67.74194	80.64516	77.41935

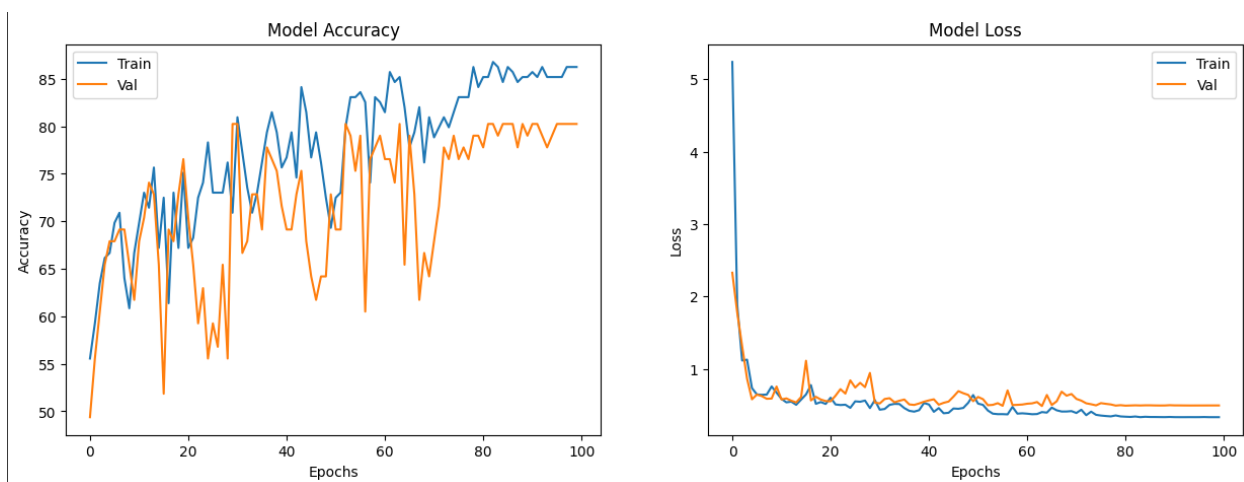


圖 1 學習率 0.001 和 3 層神經網路的準確率以及 Loss 圖

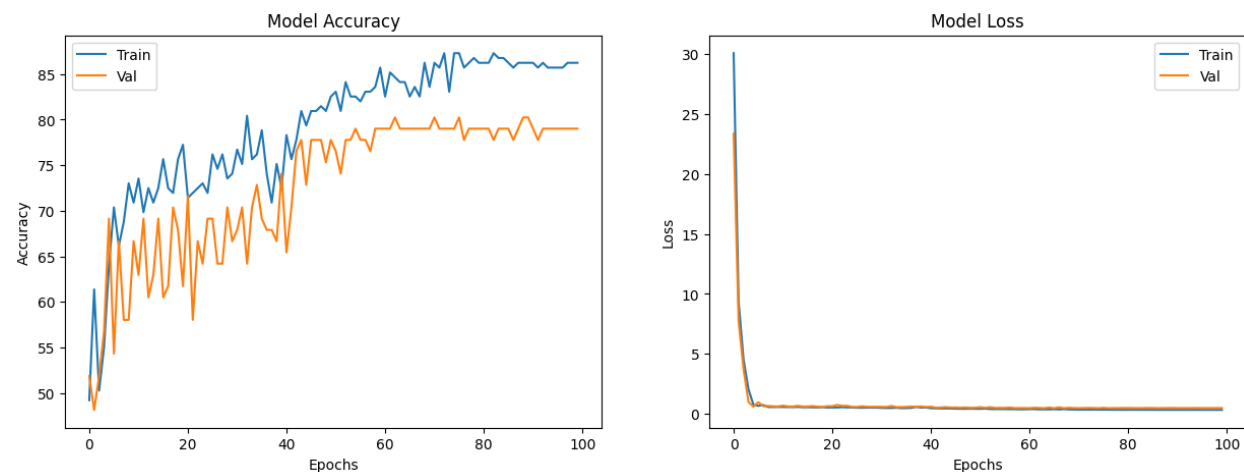


圖 2 學習率 0.005 和 3 層神經網路的準確率以及 Loss 圖

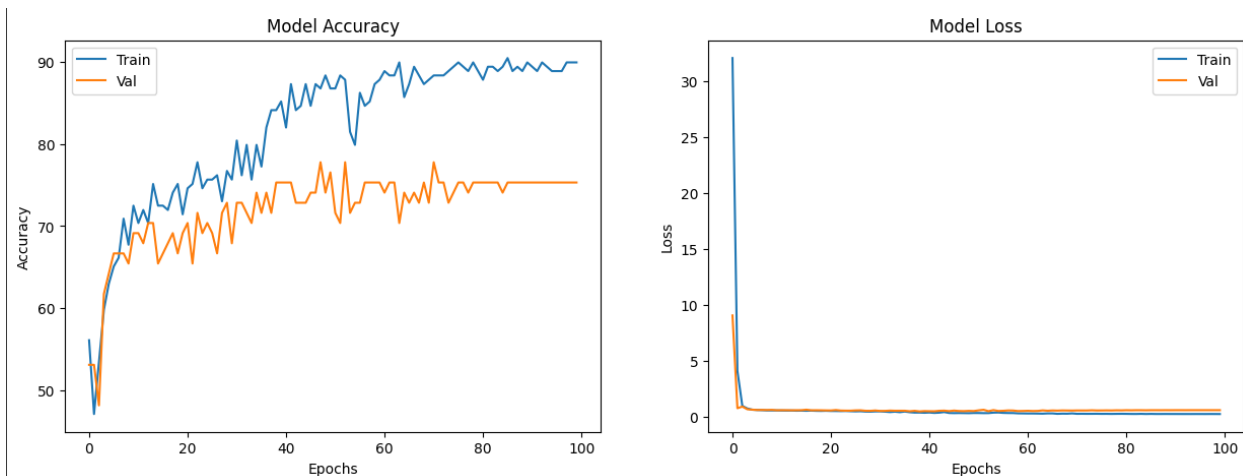


圖 3 學習率 0.007 和 3 層神經網路的準確率以及 Loss 圖

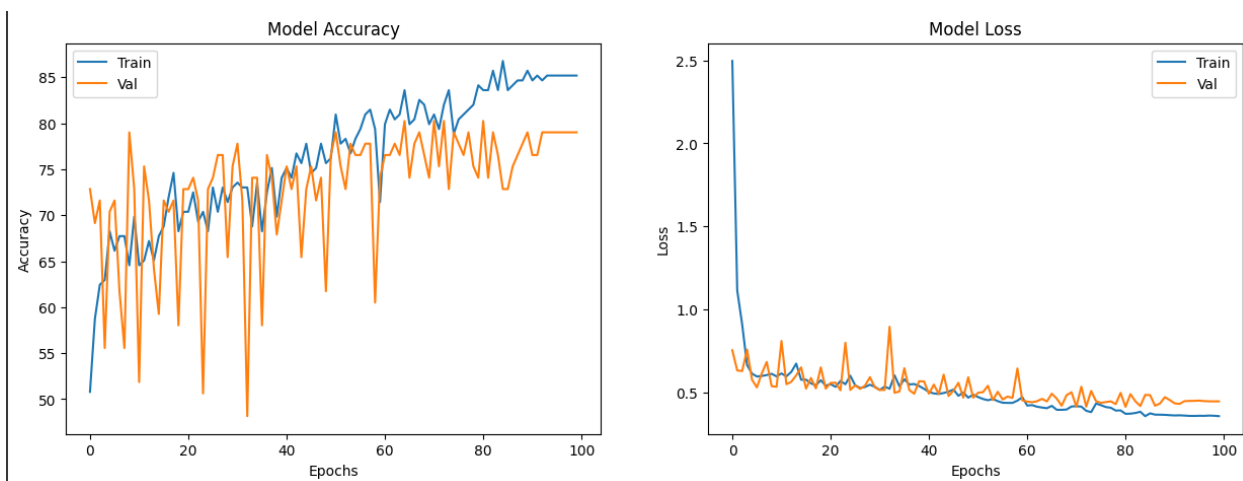


圖 4 學習率 0.001 和 4 層神經網路的準確率以及 Loss 圖

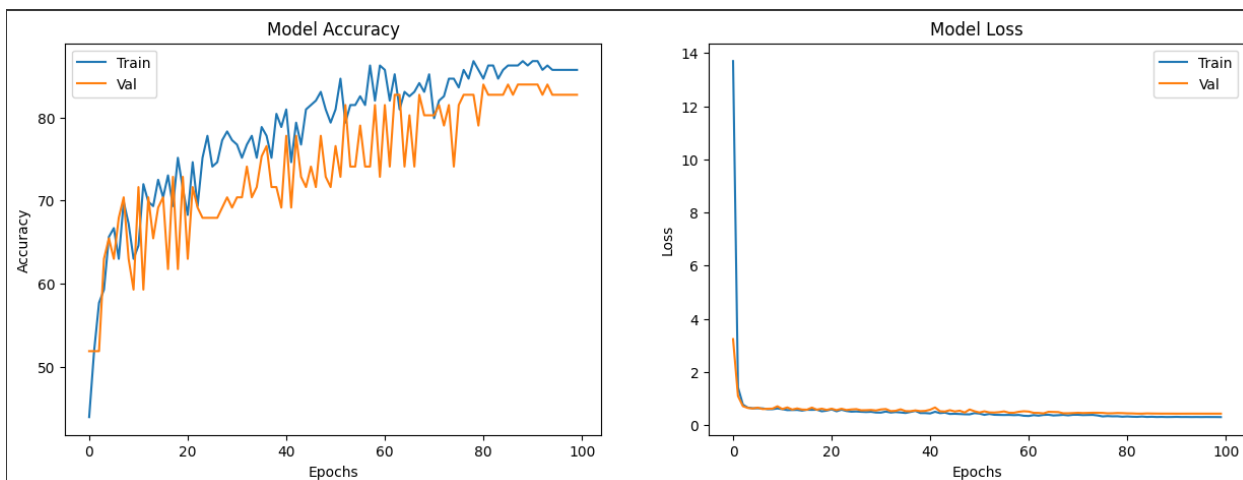


圖 5 學習率 0.005 和 4 層神經網路的準確率以及 Loss 圖

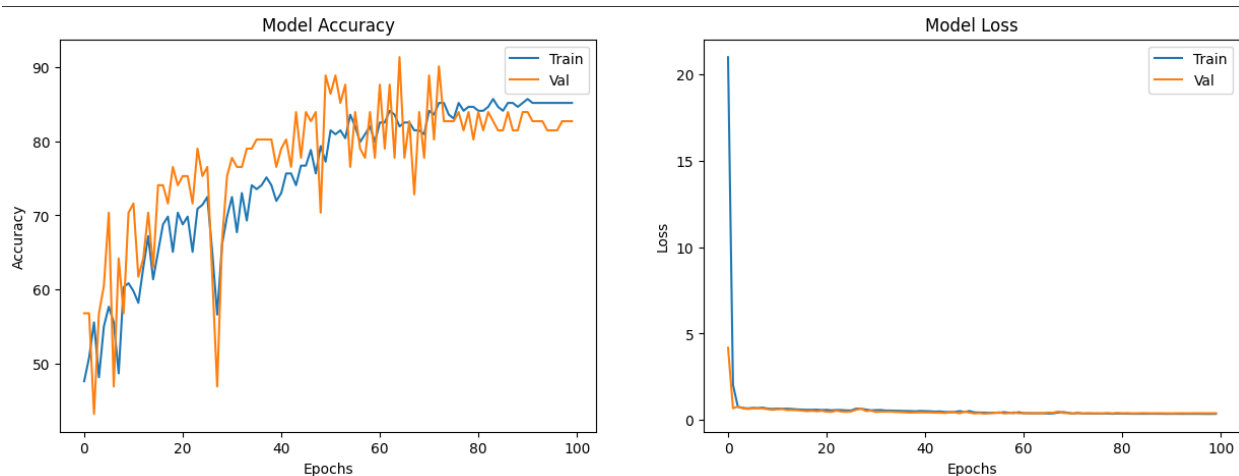


圖 6 學習率 0.007 和 4 層神經網路的準確率以及 Loss 圖

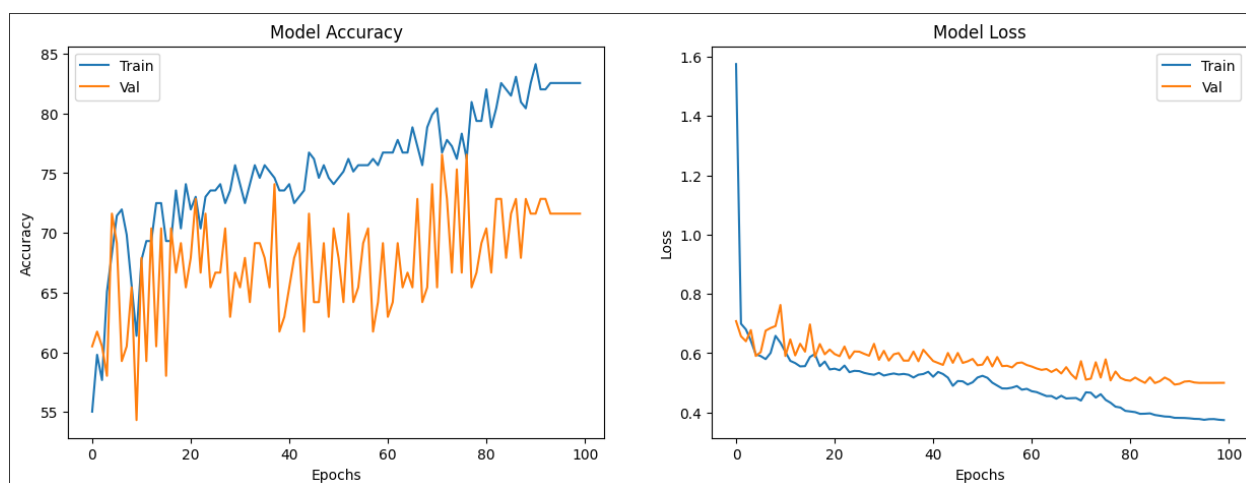


圖 7 學習率 0.001 和 5 層神經網路的準確率以及 Loss 圖

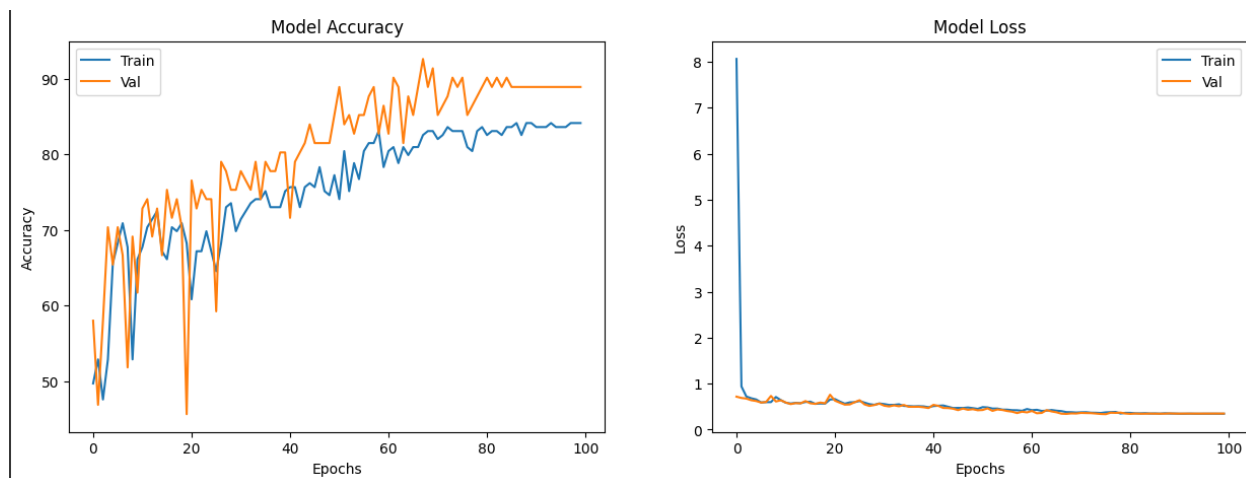


圖 8 學習率 0.005 和 5 層神經網路的準確率以及 Loss 圖

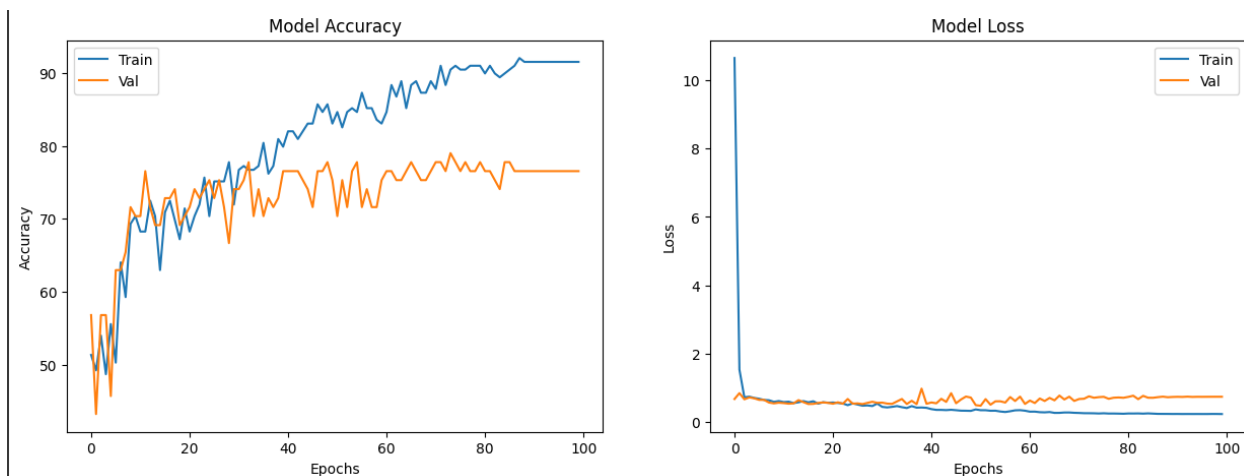


圖 9 學習率 0.007 和 5 層神經網路的準確率以及 Loss 圖

2. 分析超參數影響

在圖 1、4、7 中可以看到，學習率 0.001 的時候，在各種層數當中，驗證集的準確率和 Loss 值變動的幅度都比較大，模型比較不穩定。當神經網路層數從 3 層增加至 5 層時，模型的訓練準確率持續提升，但驗證集準確率呈現更明顯的波動，甚至在某些 epoch 出現大幅下滑。這顯示過深的模型可能導致過擬合。整體來看，3 層結構在準確率與 loss 表現最穩定，泛化能力較佳，可能是較合適的架構設定。

3. 訓練與測試集準確率不一致的原因

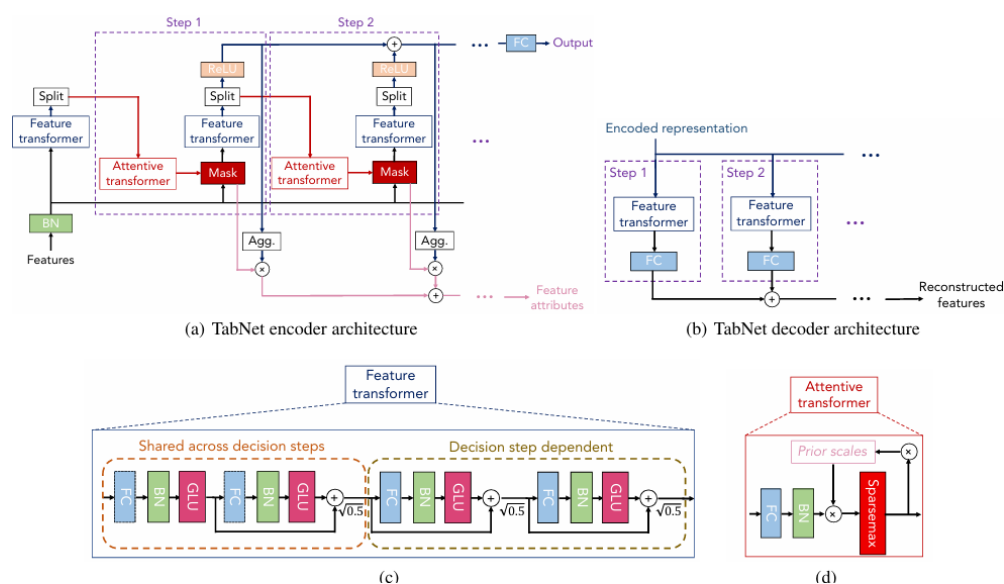
針對學習率 0.001 和 5 層神經網路的模型，可以猜測雖然模型在訓練集上達到約 85% 的準確率，但在測試集上的準確率僅為 67.74%，顯示模型存在過擬合現象。可能原因包括：模型過度學習訓練資料中的細節，導致泛化能力下降；此外，驗證集準確率波動大，可能代表模型在不同 epoch 間對參數變動敏感，使得模型選擇不穩定。也可能是測試資料的分布與訓練資料略有差異，進一步影響模型表現。

4. 特徵選擇方法與重要性

Lasso 透過 L1 正則化在模型訓練時引入懲罰項，會將不重要特徵的權重壓縮至零，達到「自動選擇特徵」的效果。這使模型只保留與預測結果最相關的資訊，移除噪聲與冗餘特徵，降低模型複雜度，進而減少過擬合的風險，提升泛化能力，特別適合高維、特徵冗餘的資料集。

參考論文: “A survey on feature selection methods”

5. 介紹更適合 tabular data 的深度學習模型



TabNet 是一種適用於表格資料的深度學習架構，解決傳統神經網路在 tabular data 上效能與可解釋性不足的問題。TabNet 結合了逐步特徵選擇、稀疏性控制與深度表示學習，能夠針對每筆資料選出最具判別力的特徵，提升模型效能並實現可視化解釋。此外，作者也首次在表格資料上導入自監督預訓練，顯著改善低標記情境下的學習表現與收斂速度。

參考論文: “TabNet: Attentive Interpretable Tabular Learning”