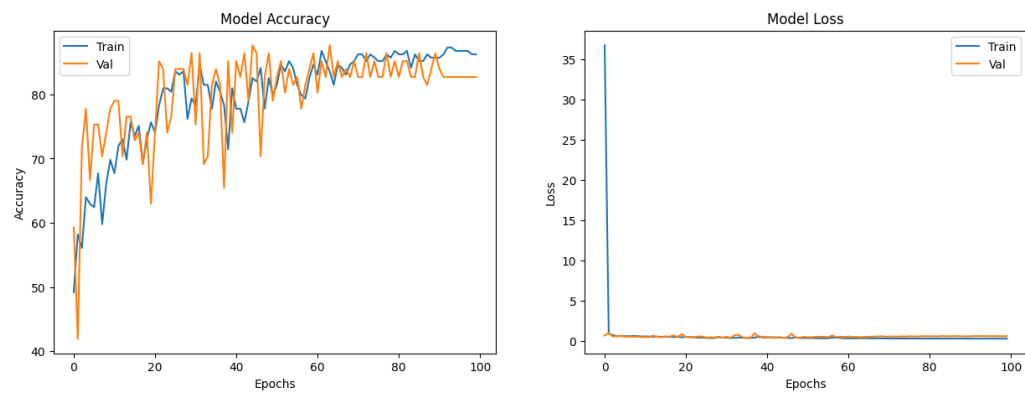
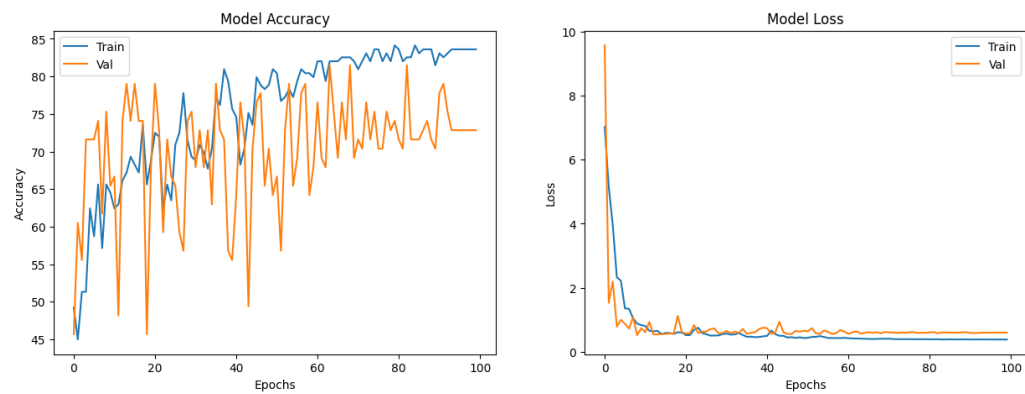


## 1. 調整 Learning Rate

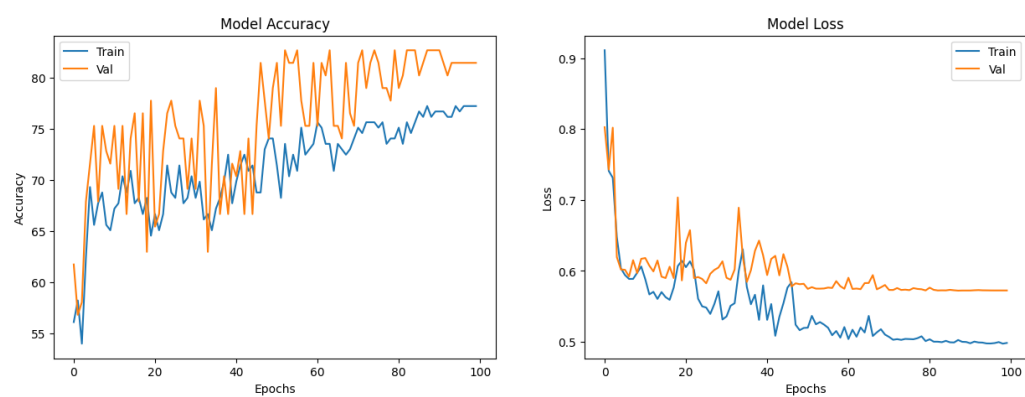
➤ Learning Rate= $1e-2$ , Test\_accuracy=83.871



➤ Learning Rate= $1e-3$ , Test\_accuracy=70.968

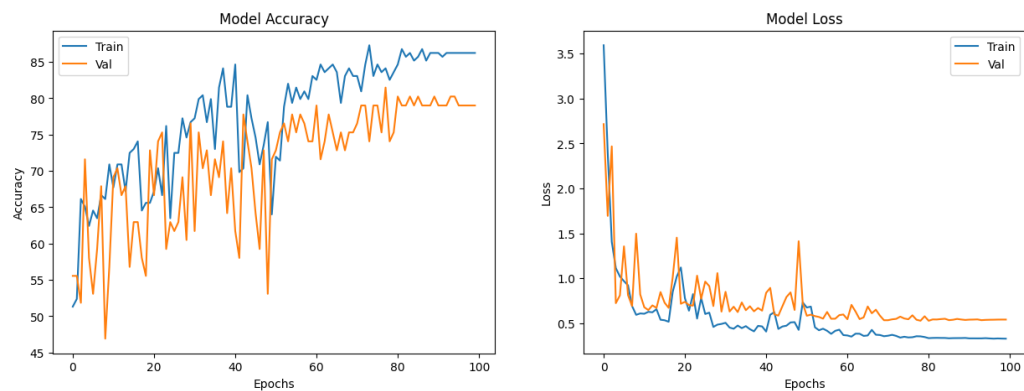


➤ Learning Rate= $1e-4$ , Test\_accuracy=64.516

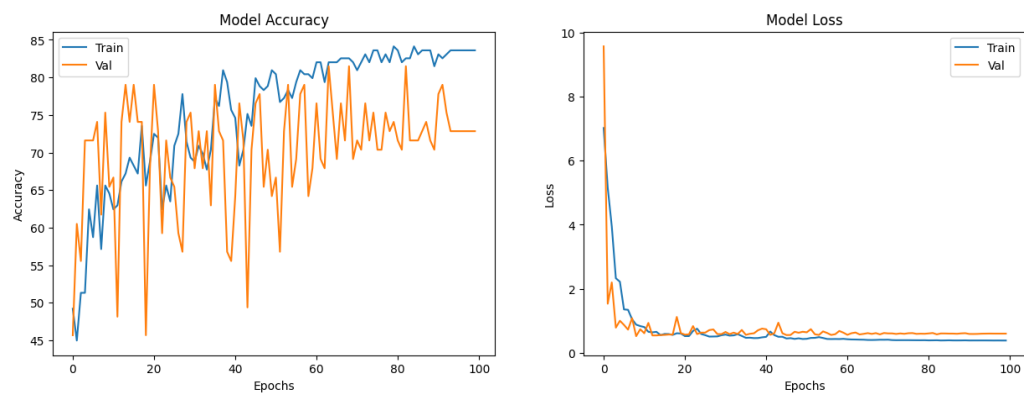


## 調整 Batch Size

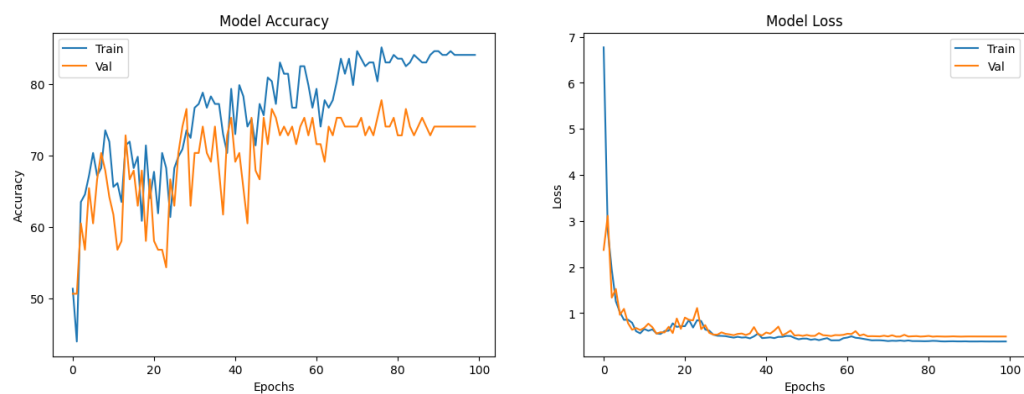
➤ Batch Size=16, Test\_accuracy=74.194



➤ Batch Size=32, Test\_accuracy=70.968



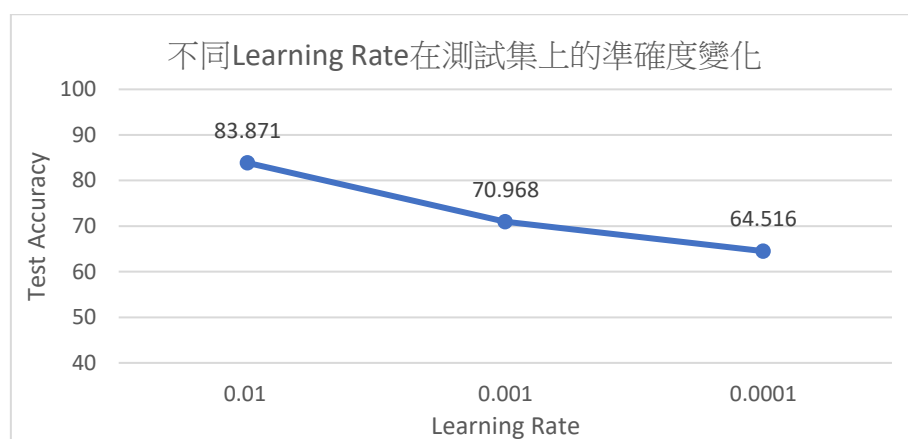
➤ Batch Size=48, Test\_accuracy=77.419



## 2. 調整 Learning Rate

由於 Learning Rate 決定了每次迭代時梯度下降的速度。Learning Rate 過大可能造成模型一直在最佳解附近徘徊但錯過而無法收斂；Learning Rate 過小則可能會造成收斂速度慢、或收斂於區域最小值。

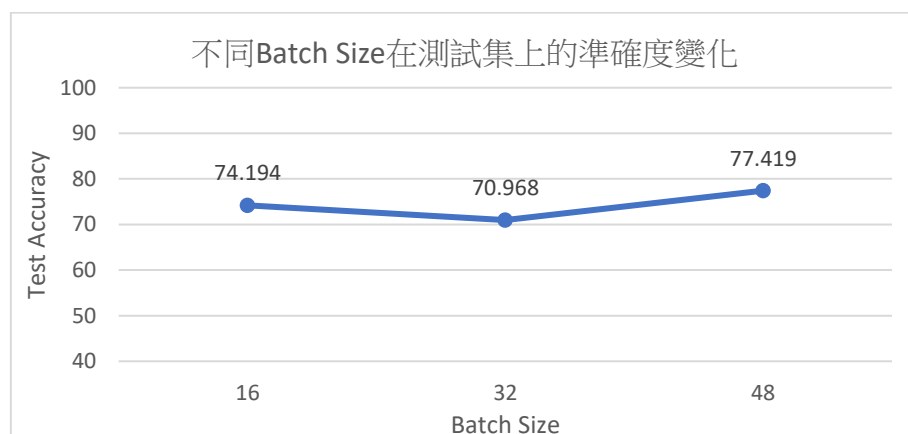
在本實驗的三個數據中，Learning Rate 與測試集的準確度呈現正相關。當 Learning Rate=0.01 時，可以發現 loss 下降的非常快，圖形幾乎是呈現直角，代表 Learning Rate 可能太大，導致模型收斂太快。而 Learning Rate=0.0001 時，可能是因為收斂速度太慢，導致模型還沒到達最佳解，所以準確度較低。



## 調整 Batch Size

Batch Size 指的是一次訓練所抓的訓練樣本數。Batch Size 大的時候，因為考量了較多數據，修正的方向會比較準確，但因為每次 Iteration 要算比較久，往 Minima 前進的次數就相對小；Batch Size 小的時候，因為只有考量局部數據，所以修正的方向會出現偏差，但因為每次 Iteration 要計算的資料量比較少，就有機會在同樣的時間成本下，修正比較多次。

在本實驗的三個數據中，Batch Size 與測試集的準確度之間並沒有明確的關聯性。推測可能是因為 Batch Size 需要配合 Learning Rate 等其他超參數進行調整。



3. 造成訓練與測試資料準確度差異的原因可能為：

- 過擬合：當模型過度學習訓練集的資料，而沒有正確觀察到資料規律時，將導致模型在訓練集的準確度很高，但測試集的準確度卻很低。
- 訓練集與測試集資料不符：若訓練集與測試集之間的切分不適當，可能造成訓練集與測試集的資料無法代表彼此，導致準確度有落差。

4. 特徵選擇是指為了建構模型而選擇相關特徵子集的過程，可以簡化模型、縮短訓練時間、增強可解釋性、改善通用性並避免過擬合。特徵選擇的方法可以分為以下三種：

- Wrapper methods 包裝類方法：每個新子集都被用來訓練一個模型，然後用驗證數據集來測試打分數。
- Filter methods 過濾類方法：採用相關係數等代理指標，而不根據特徵子集的錯誤率計分。
- Embedded methods 嵌入類方法：考慮特徵間的關係及預期要使用的模型，在進行模型建立的過程同時也做特徵選擇。

Reference：維基百科 - 特徵選擇 <https://zh.wikipedia.org/zh-tw/%E7%89%B9%E5%BE%81%E9%80%89%E6%8B%A9>

5. TabNet 是一個適合處理表格資料的深度學習模型，基本原理包括：

- Attention Mechanism：模型能夠關注特定數據特徵，同時忽略不相關的特徵。
- Decision Trees：使用一組決策樹對特徵和目標變數之間的關係進行建模，使模型能夠捕捉資料中複雜的交互作用和模式。
- Sparse Feature Selection：在訓練過程中採用稀疏特徵選擇，也就是在每個決策步驟中僅使用特徵子集進行預測。這有助於提高模型的效率和可擴展性，特別是對於具有大量特徵的資料集。
- Sequential Decision-Making：在決策樹的每一步做出順序決策，使模型能夠學習輸入資料的層次表示，從而獲得更好的泛化效能和可解釋性。

Reference：<https://medium.com/@turkishtechnology/deep-learning-with-tabnet-b881236e28c1>