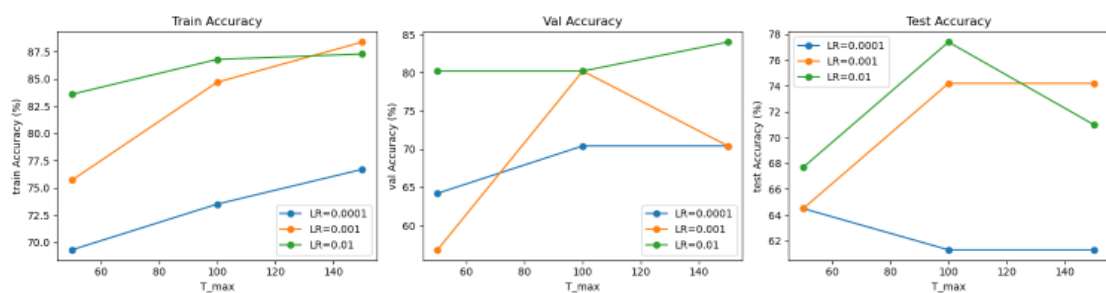


1. 挑選 Learning Rate 和 T_max 兩個參數

以下是挑選 Learning Rate 為 0.0001、0.001、0.01 和 T_max 為 50、100、150 的 Accuracy(%) 結果：

Learning Rate / T_max	0.0001			0.001			0.01		
	train	val	test	train	val	test	train	val	test
50	69.3	64.2	64.5	75.7	56.8	64.5	83.6	80.2	67.7
100	73.5	70.4	61.3	84.7	80.2	74.2	86.8	80.2	77.4
150	76.7	70.4	61.3	88.4	70.4	74.2	87.3	84	71

2. 從以下的圖來看，在 Learning Rate 為 0.01 時(綠色的線)，普遍有較高的準確度，而在 Train set 的時候，提高 T_max 的值有助於提升 Accuracy，但是在 Val set 和 Test set 則不明顯，甚至有可能會降低 Accuracy。



3. 在以上的結果中，都可以看到 Test 的準確度明顯低於 Train 的準確度，表示模型有 Overfitting (High variance) 的發生，對於沒看過的資料泛化能力較差，也有可能是因為 Test 的資料集與 Train 的資料集存在不同的 distribution 或是資料的來源不完全相同。

4. Variance Threshold: 刪除某些 Variance 小於特定閾值的特徵，像是如果某特徵的值都是一樣的，那麼這個特徵對模型的貢獻就較低，可以刪除。

優點：減少特徵數量來降低資料的維度，可以降低計算成本，也可以避免 Overfitting。另外，選擇與決策變數相關的特徵組合可以增加模型可解釋性，可以了解哪些特徵是比較重要的。

Reference: <https://medium.com/aimonks/a-comprehensive-guide-to-feature-selection-using-variance-threshold-in-scikit-learn-0b10146aa71f>

5. TabNet: 結合神經網路和決策樹的優點來解決 ANN 處理表格數據時的問題。

特點和優點：

- 利用 attention 的機制來做特徵篩選，模擬決策樹分裂的方式，不需要像傳統的機器學習做特徵工程。
- 提供每個特徵重要性的評分，使模型有更好的可解釋性，避免傳統神經網路的黑箱(Black Box)問題
- 對於含有不同資料類別的 tabular data ，像是含有圖片的資料，可以有效率的建立模型。

Reference: <https://medium.com/@kdk199604/tabnet-a-deep-learning-breakthrough-for-tabular-data-bcd39c47a81c>