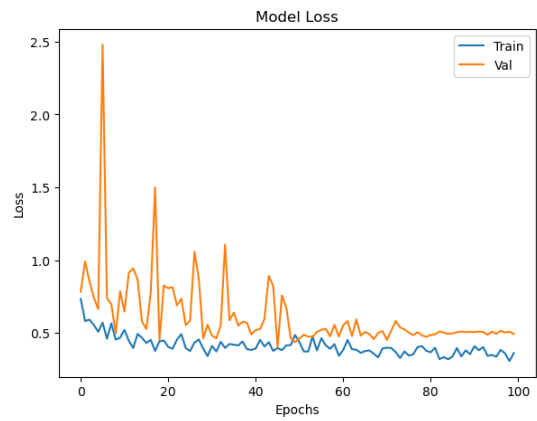
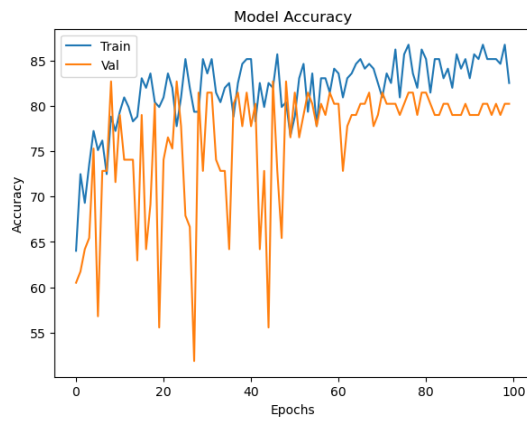
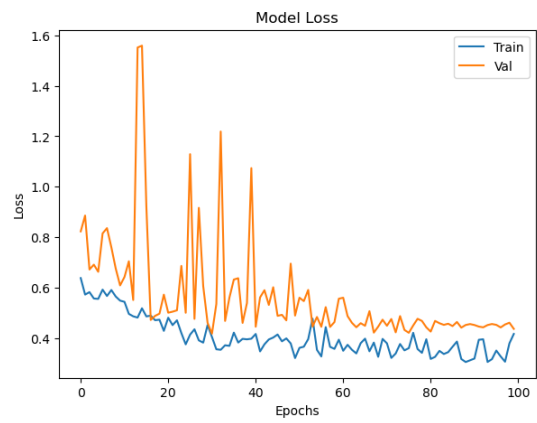
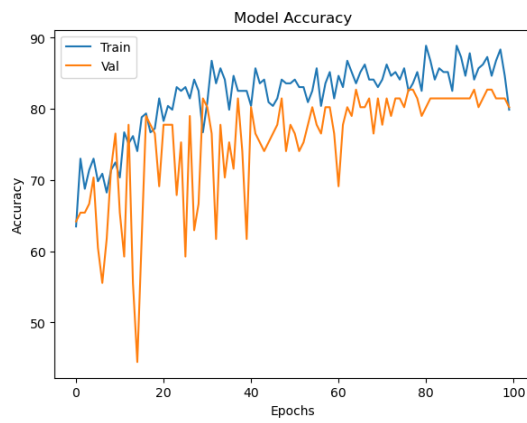


1.

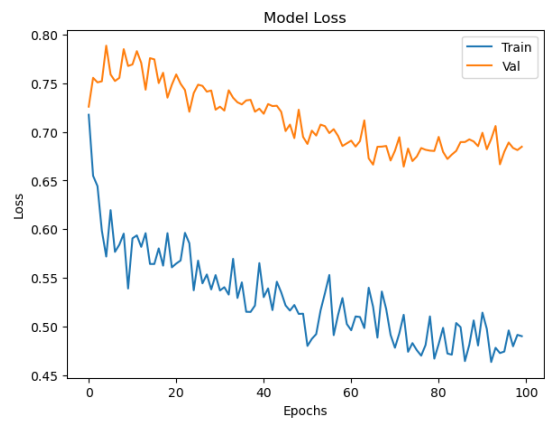
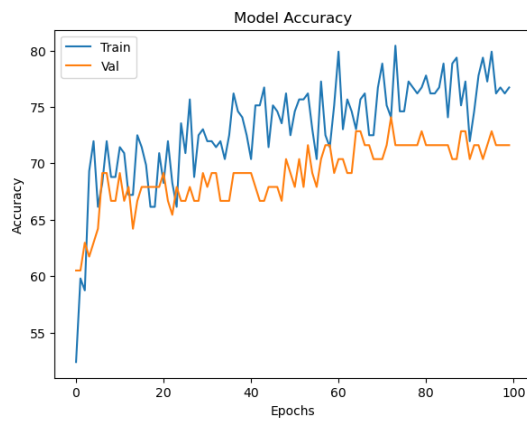
Lr0.01 node=128:



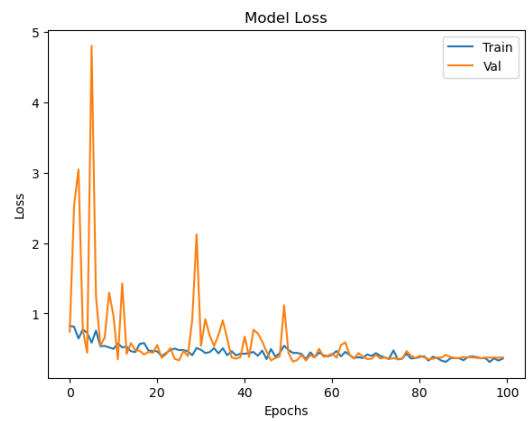
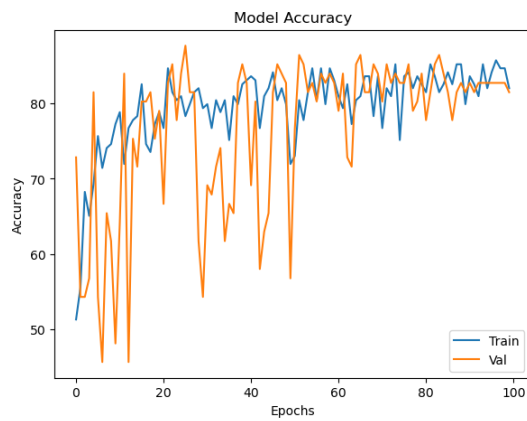
Lr0.001 node=128:



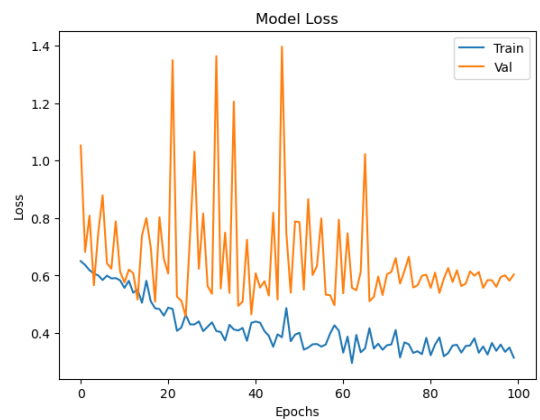
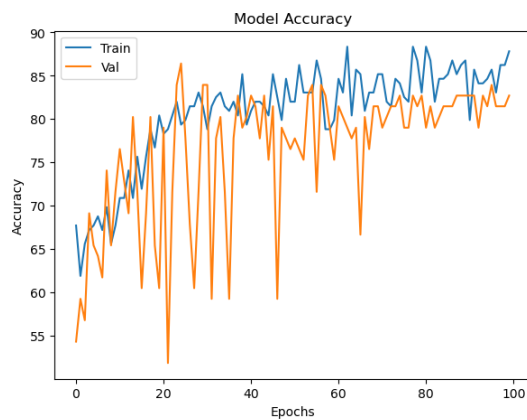
Lr0.0001 node=128:



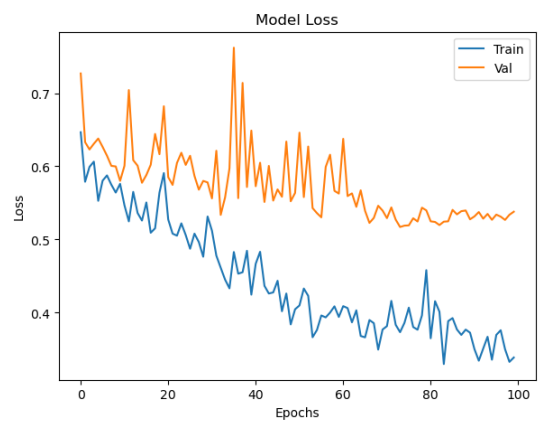
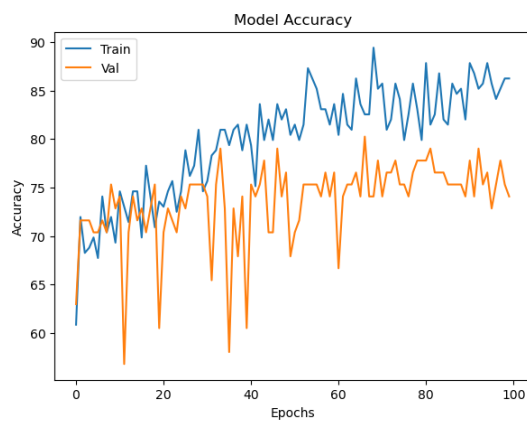
Lr0.01 node=256:



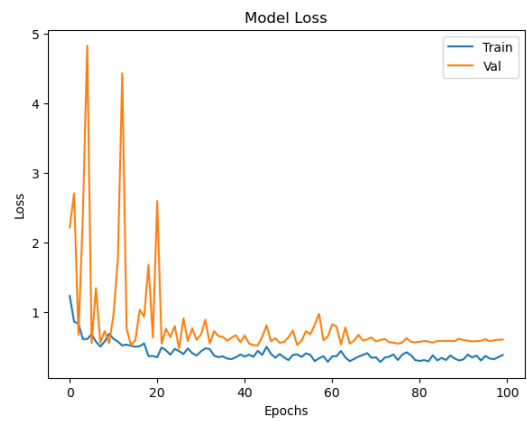
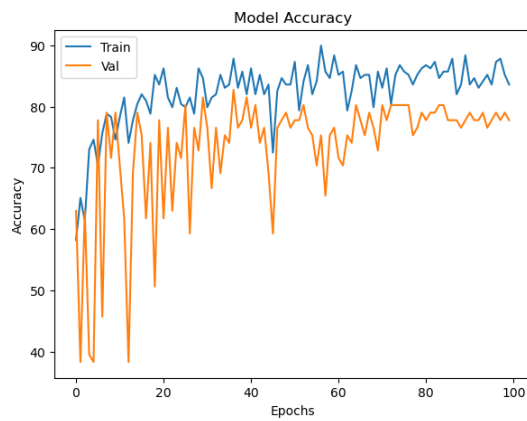
Lr0.001 node=256:



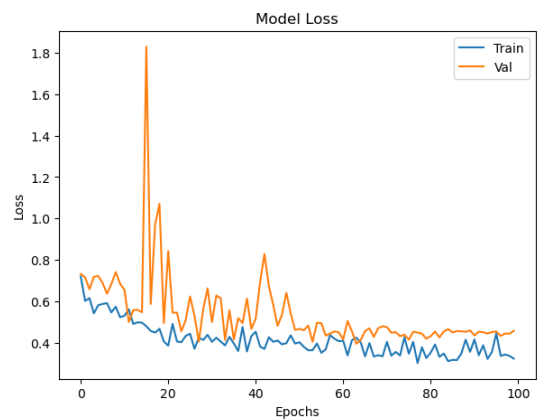
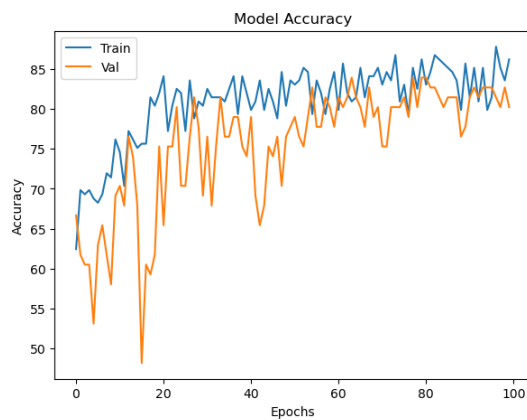
Lr0.0001 node=256:



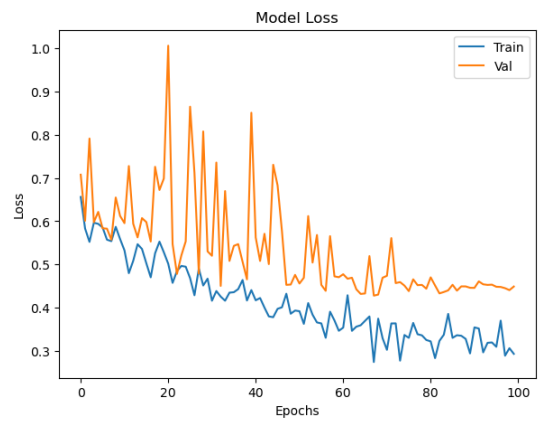
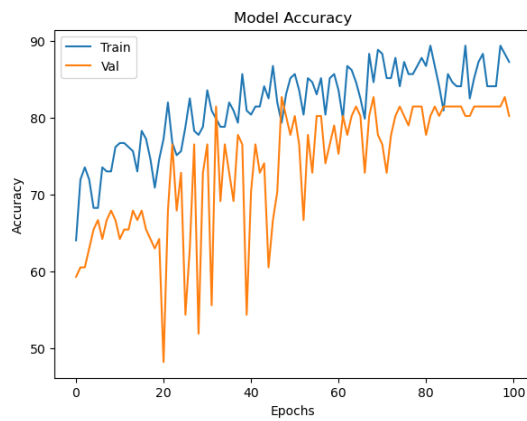
Lr0.01 node=512:



Lr0.001 node=512:



Lr0.0001 node=512:



Train acc:

train acc	10^{-2}	10^{-3}	10^{-4}
train acc nodes=128	86.77	88.89	80.42
train acc nodes=256	85.71	88.36	89.42
train acc nodes=512	89.95	87.83	89.42

Train loss

train loss	10^{-2}	10^{-3}	10^{-4}
train loss nodes=128	0.3086	0.3051	0.4635
train loss nodes=256	0.3085	0.2958	0.3298
train loss nodes=512	0.2805	0.3013	0.2741

Val acc:

val acc	10^{-2}	10^{-3}	10^{-4}
val acc nodes=128	82.72	82.72	74.07
val acc nodes=256	87.65	86.42	80.25
val acc nodes=512	82.72	83.95	82.72

Val loss:

val loss	10^{-2}	10^{-3}	10^{-4}
val loss nodes=128	0.4056	0.4174	0.6643
val loss nodes=256	0.3138	0.4588	0.5171
val loss nodes=512	0.4741	0.3945	0.4278

Test acc:

test acc	10^{-2}	10^{-3}	10^{-4}
test acc nodes=128	80.6451613	80.64516	74.19355
test acc nodes=256	77.4193548	74.19355	74.19355
test acc nodes=512	83.8709677	80.64516	77.41935

2.

在其他超參數和模型架構不變的情況下，當 learning rate 降低但 nodes 數

不變時，大多數情況下準確率（ACC）會下降，而損失值（loss）則會上升。相反地，在相同 learning rate 的情況下，當 nodes 數設為 256 時，不論是準確率還是損失值表現都最好，其次是 nodes 為 128，再來是 512。因此，最佳的參數組合為 learning rate = 10^{-2} 、nodes = 256。

3.

訓練集與測試集準確率差異通常是過擬合所造成。模型過度學習訓練資料中的細節與雜訊，導致無法良好泛化至未見資料。此外，模型過於複雜、資料量不足、未正確使用正則化（如 Dropout 或 L2）也可能造成此現象。若訓練與測試資料分布不同或發生資料洩漏，也會導致表現差距。解法包括交叉驗證、簡化模型結構與正則化技巧的應用。

4.

特徵選擇的目的在於找出對模型預測最有幫助的欄位。像是我會先看每個欄位跟目標變數的關係，例如數值型的我會用相關係數，類別型的可以看分布情況。也可以透過模型本身去判斷哪些特徵比較重要，例如用樹模型觀察 feature importance。選對特徵能讓模型學得更快、準確率更高，也比較不容易過擬合。

5.

雖然人工神經網路很常見，但它不一定適合處理表格資料，因為表格資料不像影像或語音那樣有明確的結構。我覺得比較適合的模型是那種能根據不同欄位特性去學習的，例如有注意力機制的模型，它會自己判斷哪些欄位比較重要，動態調整學習方向。這樣不只提升準確率，也讓模型在資料稀疏或異質性高的情況下還是有穩定表現。另外，這種模型通常也有比較強的可解釋性，對於需要知道模型怎麼判斷的情境很有幫助。