

# 海大資工 AI 機器學習作業報告

## (一)實驗結果

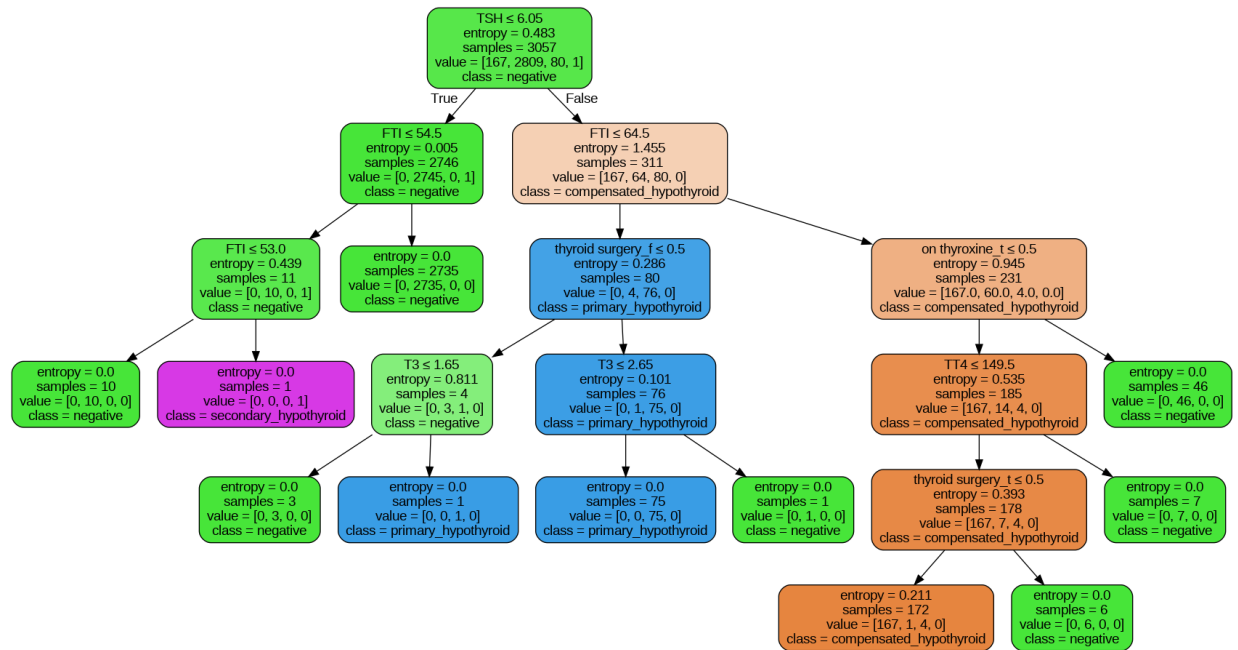
系統	分類器	系統設定	正確率
A	SVM	svm.SVC, kernel='poly'	97.64%
B	SVM	svm.SVC, kernel='linear'	96.76%
C	SVM	svm.SVC, kernel='rbf'	91.47%
D	Naive Bayes	MultinomialNB	93.52%
E	Decision Tree	criterion = 'entropy' max_depth=5	99.11%
F	Random Forest	n_estimators=100 criterion = 'gini'	99.11%
G	XGBoost	n_estimators=100 learning_rate= 0.7	98.52%
H	Tensorflow	3 層 hidden layer，每一層後面都有接 leakyReLu 和 batchnorm	97.94%

## (二)系統比較

系統 A、B、C 都屬於 SVM 分類器，差別在於 kernel（核函數）的不同，kernel='poly'的效果最佳。傳統機器學習模型（Decision Tree、Random Forest、XGBoost）與深度學習模型，深度學習模型的效果沒有傳統機器學習模型來得好。

## (三)結論

Decision Tree 和 Random Forest 的效果是最好的，我把 Decision Tree 展開來：



資料集的資料特徵與類別可能較為明確，適合樹狀結構分割，決策樹和隨機森林善於處理這類結構化資料，並能自動發現特徵之間的重要規則，因此表現特別突出。Tensorflow 模型雖然有三層 hidden layer 並加入 leakyReLU 與 batchnorm，理論上能處理更複雜的非線性關係，但由於資料集規模或特徵維度有限，深度學習未必能完全發揮優勢，反而可能因參數過多導致訓練效果略低於傳統方法。