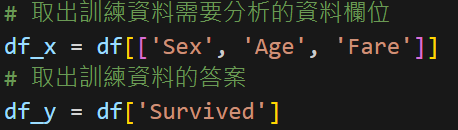
Report

* The way to improve:

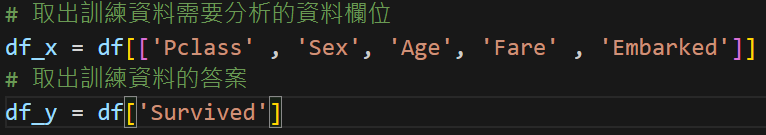
1. 增加輸入特徵

原本:





增加Pclass與Embarked(但Embarked需跟Sex一樣先做LabelEncoder)

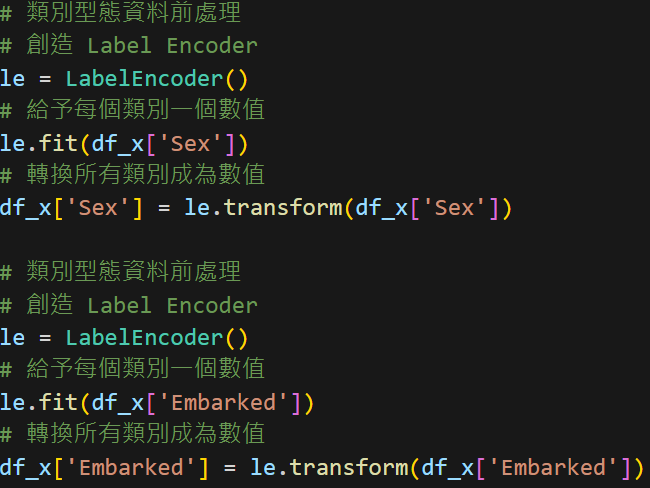


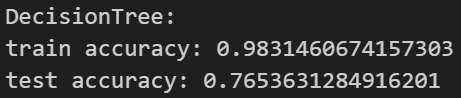


1. 使用不同的前處理方法

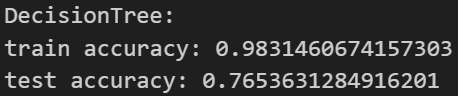
此處增加將Embarked進行LabelEncoder，因為Embarked與Sex一樣是string，而模型中是利用數學計算，string餵進去會爆掉，因此做與Sex一樣的動作，比較不一樣的是我新增兩個欄位，分別是sex與embarked，讓新增的欄位的type不為object，此做法是因為第二題會用到XGBClassifier，XGBClassifier的訓練資料不可為object，因此事先轉換，再將原本的那兩個欄位刪掉，且此作法不影響DecisionTree的表現

原本:

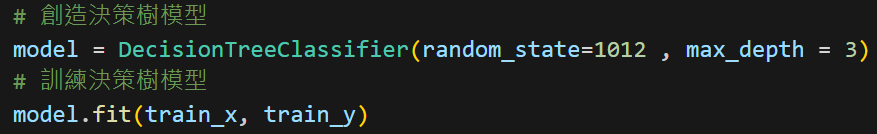


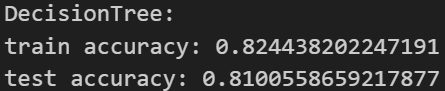


更改後:



1. 調整超參數





更改的地方為設定樹深，避免他overfitting(可從上面的看出train accuracy從0.98掉到0.82，但test accuracy卻從0.76升到0.81)

* Different model comparison:

