***動機***

提案動機就是因為政府大力補助電動機車，在路上就越來越常看到gogoro的充電站，同時也讓我們想到幾年前的新聞，說政府要逐漸替換油車，讓台灣只剩下電車。 在假設電車真的環保的前提下，台灣的發電量跟交通環境是允許這麼大的改變嗎，因為這樣我們打算先從台灣的國民載具 機車來研究這件事情，在台灣最大宗的電動機車廠商是gogoro所以自然的我們的方向就是gogoro的充電站在現階段到底夠不夠有其gogoro的機族使用。

***資料處裡流程***

我們的資料處理過程有這些，我們先合併個年份的車輛登記資料跟比對gogoro的站點資料跟現實中的有沒有衝突，可能有些站沒了有些站還有，然後把這些資料欄位裡面的日期，地址，統一成一種格式再去用隨機森林做分析。

***決策樹圖***

從我們的決策樹的圖裡面我們可以看到每一個節點都有一個條件 像第一個就是電車登記數<=450 小於就會去左邊 大於就會去右邊這樣一層一層的判斷下去，吉尼係數越小代表資料越純，換句話說就是越小這個分類被分的越完整，不會有不同類的資料在裡面，samples代表這個傑典包含的樣本數，value代表不同類別的資料分布，依照順序代表新增、保持不變、減少(順序是因為我程式給他的標籤順序是這樣，數值跟樣本數部一樣是因為有加權的關係)，樹中有三種顏色，不同顏色代表不同的分類，橘色是新增，綠色是保持不變，紫色是減少，顏色越深的代表吉尼係數比較低。class代表他的分類。所以根據我們的決策樹圖可以發現我們的模型在分類減少跟新增站點這兩個類別的時候是滿精準的，保持不變的類別反而不太準。

***report***

這個是我們模型的report，我們以需要減少的類別來舉例，精確率代表這個類別中實際上真的屬於這類別的比例，召回率代表預測的準度，F1代表精確率跟召回率的平均數，越高代表預測(判斷)的越準確。Accuracy代表的是在所有的預測中有多少趴是正確的，宏平均跟加權平均就是預測結果的算術平均數，有或沒有加上加權。 下面三個不會參考樣本數，所以後面的support都是71。(訓練集80%測試20%)

***熱力圖***

這張熱力圖代表特徵之間的關聯，(用於衡量數據中變量之間的線性相關性。)紅色是正相關、藍色是負相關、白色是沒相關。正相關代表特徵成長成正比，負相關成負比。結

熱力圖的結論

(正相關)

就是電車登記數 和 站點小計 與 2024人口數 之間具有強烈的正相關性。

人口數量可能是影響電車設置和站點數量的重要因素。

(負相關)

土地面積 與 2024人口數 和 人口密度 之間呈現輕微負相關，反映了人口密集的區域往往土地面積較小。

***混淆矩陣的分析***



***總結***

透過上面幾種圖表我們可以發現新增跟減少的兩個分類誤判數量不多跟保持不變的這個分類相差甚遠。代表保持不變的這個類型容易被誤判。

要改善保持不變這個類型的分類準確度我們可以

1. 增加樣本數，但是台灣就那麼多行政區，所以沒辦法
2. 用更適合的分類模型或是調整參數
3. 還可以找出幫助區分保持不變類別的特徵