

## Machine Learning Lab1 Report

### I. Basic

#### A. Variables

- Number of trees : 100
- Number of features :  $\sqrt{total\ features}$
- Sample size 用 training size 的 0.8 建構 Random Forest  
(split case 8000 for training, use 6400 to build random forest)

### II. Conclusion

#### A. The difficulty and Solution

這個作業的計算量滿大的，剛好又卡在兩個期中考中間，測試的時間滿少的。另外，在資料處理的部分，因為結構稍微有點麻煩，因此直接省略了 str 判定的問題(直接把只有兩個數值的資料排除在 decision tree 中)，若有更多時間可以再針對這個地方優化。

#### B. Reflection

其實基本的 code 不難，實作的同時對結構更熟悉(藉由第一部份的 code 也可以判斷自己實作的 tree 正確度，覺得非常讚!!!!)。但如果作業結束後可以對優化方向和額外操作做更多講解或是一些經驗分享之類的，感覺可以對這個項目理解的更透徹!