

1. 結果截圖

(1) Linear Regression

```
[Linear Regression] 只有考慮PM2.5的MAE = 2.6249621187905388  
[Linear Regression] 考慮18個屬性的MAE = 2.570843971346762
```

(2) Random Forest Regression

- 模型參數設定：
max_depth = 8
n_estimators = 200
- 實驗結果：

```
[Random Forest Regression] 只有考慮PM2.5的MAE = 3.0522847376367404  
[Random Forest Regression] 考慮18個屬性的MAE = 3.231471603296054
```

2. 說明與討論

從 Linear Regression 的結果可以看到，當放進模型 Input 中的 Feature (污染物) 因子增加，有使得模型評估的錯誤率下降，相較於只有考慮 PM2.5 單一因子的結果還要好。但在 Random Forest Regression 中，增加了污染物因子，卻讓模型的錯誤率提升。除此之外，整體看來 Linear Regression 的效果是比 Random Forest Regression 還要好，或許有可能是因為我還沒有 tune 到 Random Forest 模型參數的最佳狀態，但也有可能是因為在 PM2.5 數值預測的問題上，data 本身特性就是線性的。