

學號：R07922108 系級：資工碩一 姓名：陳鎰龍

1. (2%)記錄誤差值 (RMSE)(根據 kaggle public+private 分數)，討論兩種 feature 的影響

	Public	Private
所有污染源	5.67149	7.22350
PM2.5	5.93475	7.25222

用所有污染物的數值當成 features 會比起只看 PM2.5 來的好，代表 PM2.5 與其他因素還有關聯，才會得到比較低的誤差。

2. (1%)將 feature 從抽前 9 小時改成抽前 5 小時，討論其變化

	Public		Private	
抽樣時間	9 小時	5 小時	9 小時	5 小時
所有污染源	5.67149	5.99288	7.22350	7.17488
PM2.5s	5.93475	6.23907	7.25222	7.24860

取 9 個小時的 features 會比 5 個小時還要好，但 private 上沒有太大的變化，我也不太了解原因。

3. (1%)Regularization on all the weight with $\lambda=0.1$ 、 0.01 、 0.001 、 0.0001 ，並作圖

	0.1		0.01		0.001		0.0001	
	Public	Private	Public	Private	Public	Private	Public	Private
所有污染源	5.80936	7.26402	5.75136	7.24145	5.68662	7.21740	5.83369	7.27887
PM2.5	5.93034	7.24746	5.93159	7.24638	5.95194	7.25538	6.12905	7.43391

4. (1%)在線性回歸問題中，假設有 N 筆訓練資料，每筆訓練資料的特徵 (feature) 為一向量 x^n ，其標註(label)為一純量 y^n ，模型參數為一向量 w (此處忽略偏權值 b)，則線性回歸的損失函數(loss function)為 $\sum_{n=1}^N (y^n - x^n \cdot w)^2$ 。若將所有訓練資料的特徵值以矩陣 $X = [x^1 \ x^2 \ \dots \ x^N]^T$ 表示，所有訓練資料的標註以向量 $y = [y^1 \ y^2 \ \dots \ y^N]^T$ 表示，請問如何以 X 和 y 表示可以最小化損失函數的向量 w ？請選出正確答案。(其中 $X^T X$ 為 invertible)

- (a) $(X^T X)X^T y$
- (b) $(X^T X)yX^T$
- (c) $(X^T X)^{-1}X^T y$
- (d) $(X^T X)^{-1}yX^T$

我選擇 C