學號：R06723025 系級： 財金碩一 姓名：林耘寬

請實做以下兩種不同 feature 的模型，回答第 (1) ~ (3) 題：

(1) 抽全部 9 小時內的污染源 feature 的一次項(加 bias)

(2) 抽全部 9 小時內 pm2.5 的一次項當作 feature(加 bias)

備註 :

a. NR 請皆設為 0，其他的數值不要做任何更動

b. 所有 advanced 的 gradient descent 技術(如: adam, adagrad 等) 都是可以用的

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

	(1)	(2)
Private	5.55600	5.84518
Public	7.57300	7.46231
Average	6.5645	6.653745

可以發現第一個模型的 public 分數較高，但 private 分數較低，平均以後也較第二個模型低，推測可能為使用的 feature 較多，所以較為準確。

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

	(1)	(2)
Private	5.41989	5.80979
Public	7.77428	7.57955
Average	6.597085	6.69467

相較於第一題，兩個模型的平均分數都呈現上升，推論可能是拿掉的 feature 當中，有些對於我們估計還是有所幫助。

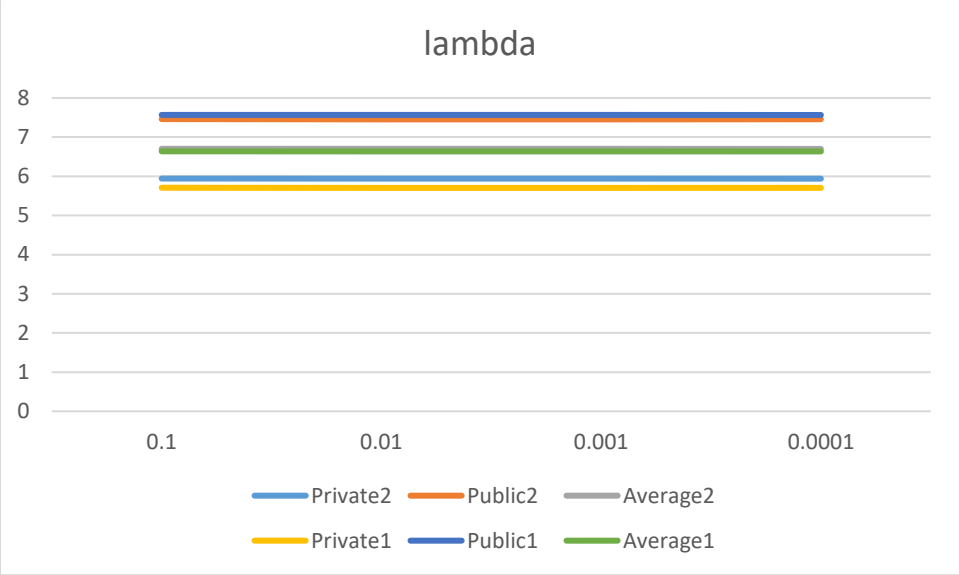
3. (1%)Regularization on all the weight with λ=0.1、0.01、0.001、0.0001，並作圖

(1)

	0.1	0.01	0.001	0.0001
Private	5.70817	5.70534	5.70520	5.70515
Public	7.56748	7.56512	7.56496	7.56491
Average	6.637825	6.63523	6.63508	6.63503

(2)

	0.1	0.01	0.001	0.0001
Private	5.94291	5.94027	5.94000	5.93998
Public	7.46075	7.45681	7.45642	7.45638
Average	6.70183	6.69854	6.69821	6.69818



並未因為正規化而導致分數有所差異，推論是因為原本僅為一次項的函數，overfitting 的狀況並未太嚴重。

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)為一存量 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$\sum_{n=1}^{N} (y^n - w \cdot x)^2$ 。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 ？請寫下算式並選出正確答案。(其中 X^TX 為 invertible)

(a) $(X^TX)X^Ty$
(b) $(X^TX)^{-0}X^Ty$
(c) $(X^TX)^{-1}X^Ty$
(d) $(X^TX)^{-2}X^Ty$

Suppose that X is a N * (d+1) matrix, where each row contains d features and bias, y is a N * 1 matrix and thus w is a (d+1) * 1 matrix.
Then, we can rewrite the loss function into the matrix form.
Loss(w) = $\frac{1}{N}\|Xw - y\|^2 = \frac{1}{N}(w^TX^TXw - 2w^TX^Ty + y^Ty)$
And we find the gradient of the loss function, which will be $\frac{2}{N}(X^TXw - X^Ty)$.
We can solve the equation $\frac{2}{N}(X^TXw - X^Ty) = 0$. Thus, w = $(X^TX)^{-1}X^Ty$, which is the vector that can minimize the loss function and (c) will be the answer.