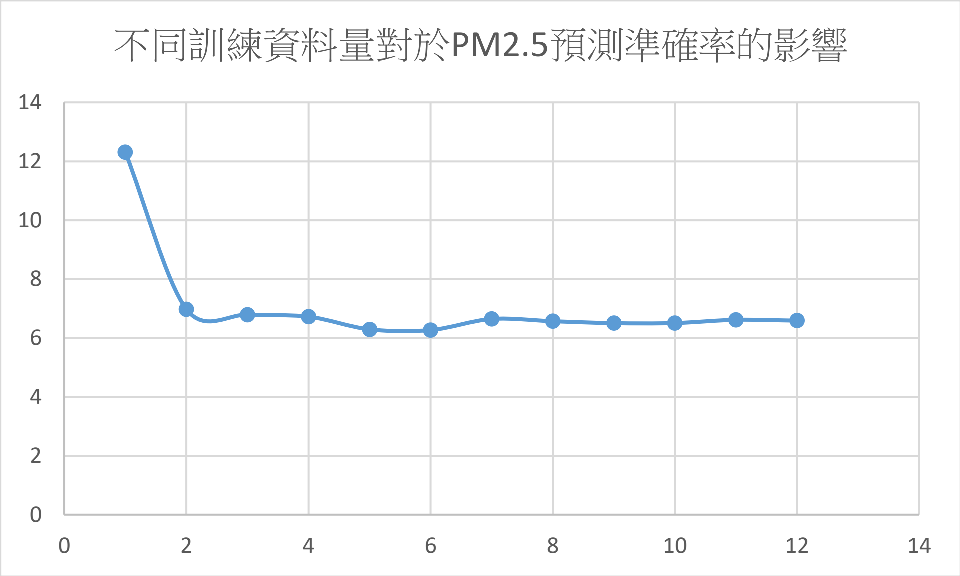
學號：B05902086　系級：資工一　姓名：周逸

1. 請簡明扼要地闡述你如何抽取模型的輸入特徵 (feature)
答：

對於每九個小時*十八類(按照)的資料，我抽取的為 feature 為底下的陣列(越接近前面的時間代表，越靠近預測目標的時間點)
然後第 15~18 類重新定義成
New 15=sin(15)*18
New 16=sin(16)*17
New 17=cos(15)*18
New 18=cos(16)*17
以下是一次項的矩陣
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 1, 1, 0, 0, 0],
[1, 1, 1, 1, 1, 1, 1, 1, 1],
[1, 1, 1, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 1, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 1, 1, 0, 0, 0],
[1, 1, 1, 1, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 0, 0, 0, 0, 0]
以下是二次項的矩陣
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],

[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 0, 0, 0, 0, 0, 0, 0],
[1, 1, 1, 1, 1, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[1, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0],
[0, 0, 0, 0, 0, 0, 0, 0, 0]

2.請作圖比較不同訓練資料量對於 PM2.5 預測準確率的影響
答：



當數據量不夠時會噴掉，當數據量夠多時就差不多

3. 請比較不同複雜度的模型對於 PM2.5 預測準確率的影響
答：

當純只有一次項的時候大概會卡在 6.0 附近就上不去了，但是加入了二次項後就能到達 5.8 附近，但是如果把三次項加進去反而又會跑到 5.9 左右

4. 請討論正規化(regularization)對於 PM2.5 預測準確率的影響

答：

在線性的情況下沒有很大的影響，但是當加入高次方項的時候，可以讓高次方項一開始的影響不會過大

5. 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 n=1Nyn-wxn2 。若將所有訓練資料的特徵值以矩陣 $X = [x^1\ x^2\ \ldots\ x^N]$ 表示，所有訓練資料的標註以向量 $y = [y^1\ y^2\ \ldots\ y^N]^T$ 表示，請以 X 和 y 表示可以最小化損失函數的向量 w 。

答：

$$w = (X^TX)^{-1}X^Ty$$