

學號: R05942114 系級: 電信一 姓名: 方敬勻

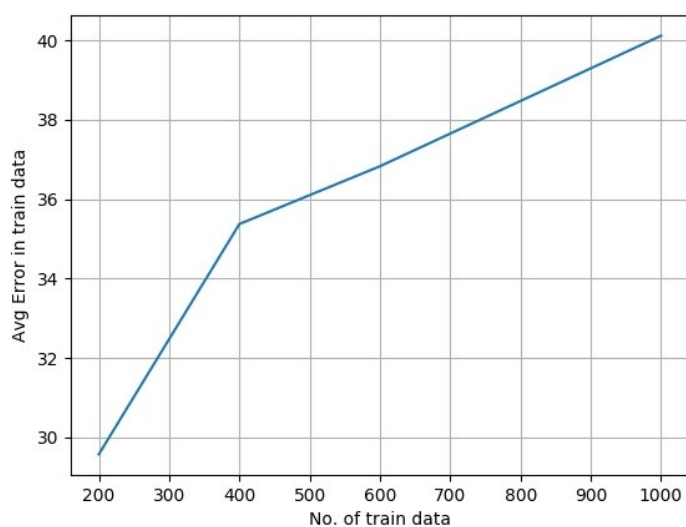
1. 請簡明扼要地闡述你如何抽取模型的輸入特徵 (feature)

答:

AMB_TEMP	前 1~9hour	CH4	前 1~9hour	CO	前 1~9hour
NMHC	前 1~9hour	NO	前 1~9hour	NO2	前 1~9hour
NOx	前 1~9hour	O3	前 1~9hour	PM10	前 1~9hour
PM2.5	前 1~9hour	RAINFALL	前 1~9hour	RH	前 1~9hour
SO2	前 1~9hour	THC	前 1~9hour		

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:

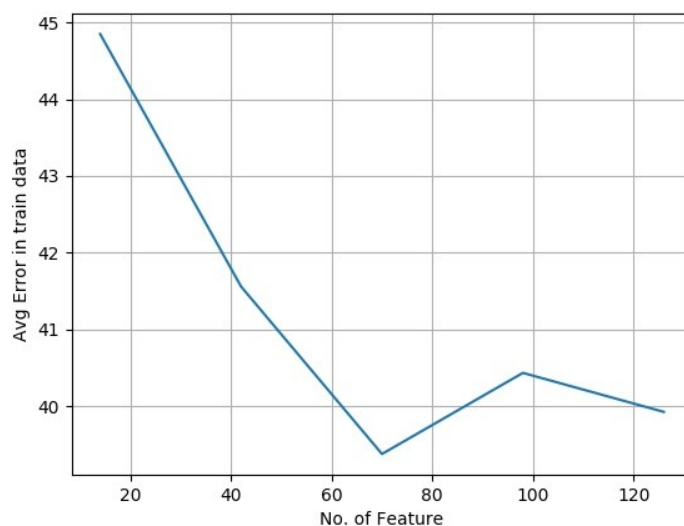


固定 Iteration=2000 的條件下, train data 數量以及對 train data 的 avg Error 分別為:

No. Train data	200	400	600	800	1000
Avg Error	29.572	35.375	36.831	38.476	40.121

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:



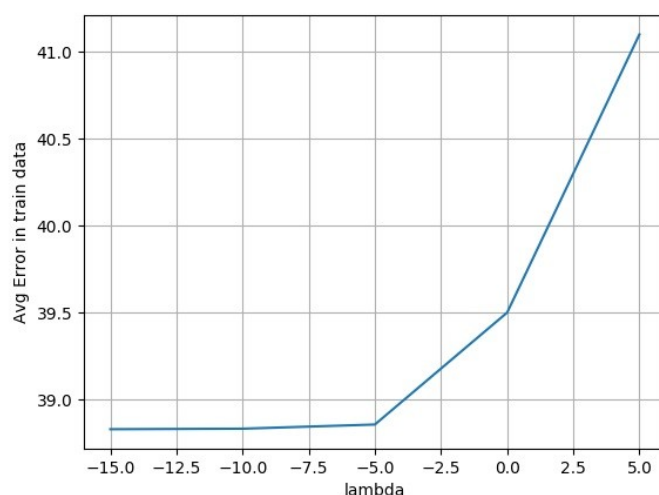
固定 iteration=2000,train data=800(有依情況調整 learning rate),feature 分別為 1.所述 feature 取前:

前 N 小時	1hour	1~3hour	1~5hour	1-7hour	1-9hour
Avg Error	44.848	41.56	39.380	40.436	38.831

紅字部份可能因為 iteration 或 learning rate 設定不佳導致

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

答:



固定 iteration=2000,train data=800

lambda	0	$10^{-10}$	$10^{-5}$	1	$10^5$
Avg Error	38.831	38.834	38.858	39.502	41.102

推測 Error 不會增加太多原因為 feature 都為一次式

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  $\sum_{n=1}^N (y^n - w \cdot x^n)^2$ 。若將所有訓練資料的特徵值以矩陣  $X = [x^1 \ x^2 \ \dots \ x^N]$  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，請以  $X$  和  $y$  表示可以最小化損失函數的向量  $w$ 。

答:  $(X^T X)^{-1} X^T y = w$