學號: B08123456 系級: 資工所二 姓名: 張庭逸

執行方式:

確保hw1.py, hw1.sh, train.csv, test.csv皆在同個目錄下,並執行 sh hw1.sh 檔案名

請實作以下兩種不同feature的模型, 回答第 1 ~ 2 題:

- (1) 抽全部9小時内的污染源feature當作一次項(加bias)
- (2) 抽全部9小時内pm2.5的一次項當作feature(加bias)

備註:

- a. NR請皆設為0, 其他的非數值(特殊字元)可以自己判斷
- b. 所有 advanced 的 gradient descent 技術(如: adam, adagrad 等) 都是可以用的
- c. 第1~2題請都以題目給訂的兩種model來回答
- d. 同學可以先把model訓練好, kaggle死線之後便可以無限上傳。

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

表格呈現的皆為未預處理過的測試資料,將其中1000組取出作為Validation data , minibatch的epoch為2000, 其餘參數照助教第一版本的code, 數據如下。

Condition	Kaggle		
	Public	Private	
All feature	407.59	401.95	
PM 2.5	7.93	6.74	

推測選用所有feature時,一是參數太多,model太過簡單,二是2000的epoch可能太少,都可能導致得到的function跟target function差異甚大。

2. (1%)解釋什麼樣的data preprocessing可以improve你的training/testing accuracy, e.g., 你怎麼挑掉你覺得不適合的data points。請提供數據(RMSE)以佐證你的想法。

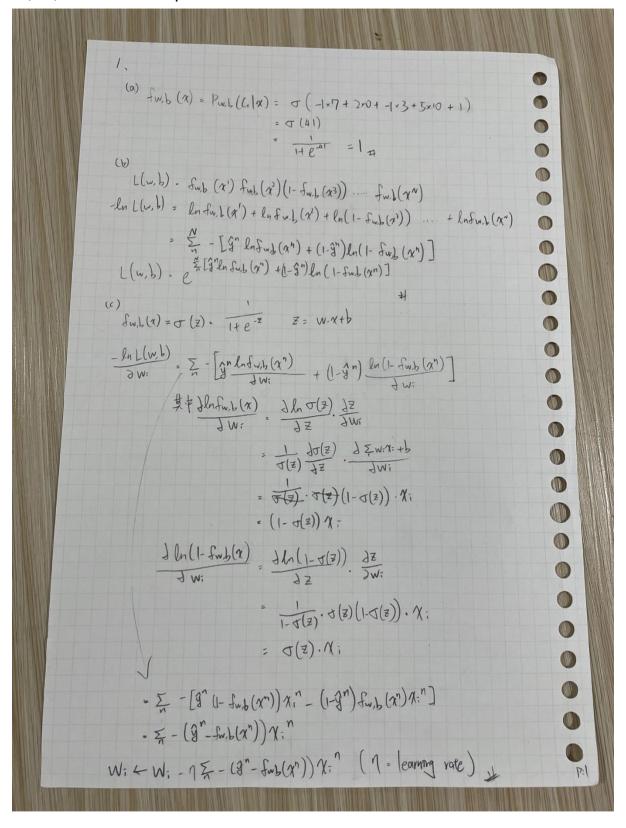
重要的preprocessing是去除雜訊。

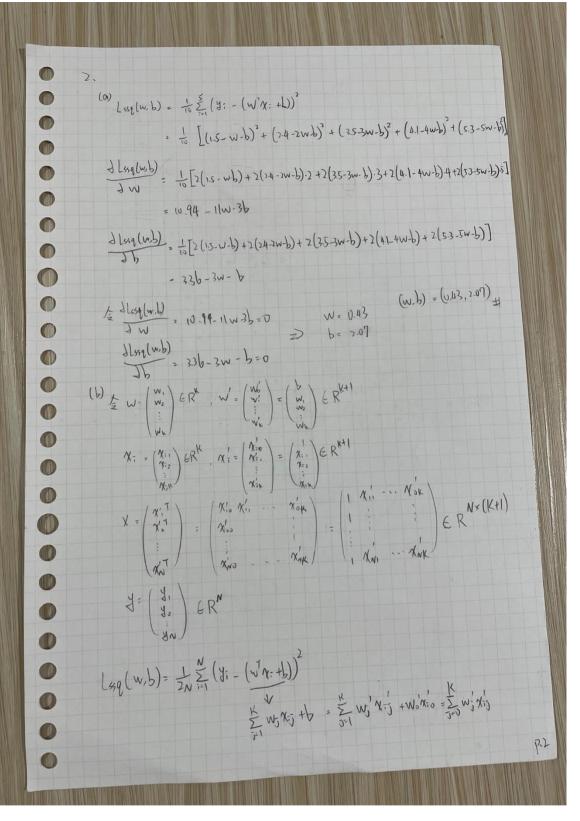
去除雜訊的方法為:從training data中,針對一個feature,計算所有數字出現的頻率並排序,極端值(極大或極小)中,出現次數太少或與原始資料前後數值相差太大,則判斷為outlier。針對各個feature,在outlier與正常data間取一個合理的border。

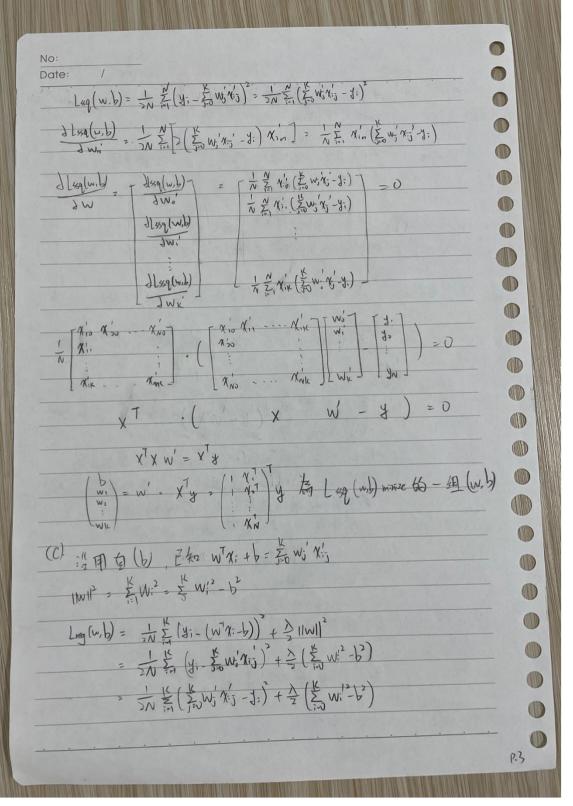
以下表格為選擇1,2,3,17項的feature,調整epoch以及調整是否有拿掉極端值。可看出,有去除雜訊的表現是有微微比只用原始的資料表現的要好。

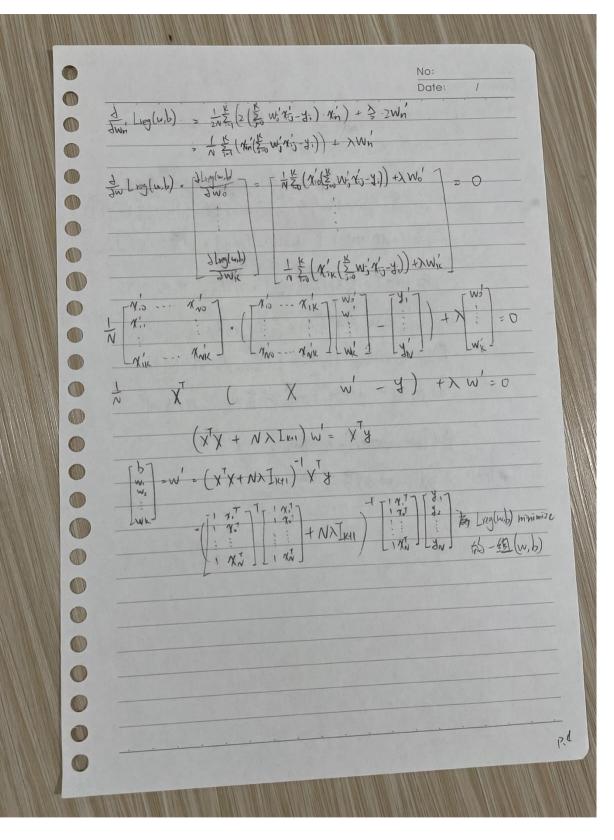
Condition		Kaggle	
		Public	Private
Epoch:300	Preprocess	7.59871	6.58600
	Original	7.59779	6.58770
Epoch:2000	Preprocess	7.59693	6.56804
	Original	7.58972	6.57709

3.(4%) Refer to math problem









Date: 3. [47 (wb)= E[] [(fw,b(1+1)-4)] = = = [N (fwh (x;)+y;)+ W7;)] $= E \left[\frac{1}{7N} \left(\sum_{i=1}^{N} \left($ = >N (= (fu,b(n;)-yi) + = zfu,b(n;) E[WTn;] + E[= (WTn;)]) 0 $E\left[\frac{x}{2} \left(w^{2} + 1\right)^{2}\right] = \sum_{i=1}^{N} E\left[\left(\frac{x}{2} + w_{2} + 1\right)^{2}\right] = \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}, w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{i \in J_{i}, i \in K} w_{i}' + 1\right] + \sum_{i=1}^{N} E\left[\sum_{j$ 0 $=\sum_{i=1}^{N}\left(\sum_{\substack{i,j,j'\neq k\\ (i,j,j'\neq k)}}W_{j}W_{j}'\cdot E\left[\gamma_{i,j}\gamma_{i,j'}\right]\right)=\sum_{\substack{i=1\\ (i,j,j'\neq k)}}^{N}\left(\sum_{\substack{i,j,j'\neq k\\ (i,j,j'\neq k)}}W_{j}W_{j}'\cdot S_{ii}S_{ij}'\sigma^{2}\right)$ 0 = \frac{N}{\sum_{\infty} \left(\frac{\z}{\sigma_{\infty}} \right) + \frac{\z}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \right(\frac{\z}{\sigma_{\infty}} \right) \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sigma_{\infty}} \right) \frac{N}{\sigma_{\infty}} \left(\frac{\z}{\sin $-\sum_{i=1}^{N} \left(\sum_{j=1}^{k} W_{j} \cdot W_{j} \cdot 1 \cdot 1 \cdot \sigma^{2} \right) = \sum_{i=1}^{N} \sigma^{2} \sum_{j=1}^{k} W_{j}^{2} = N_{0}^{2} ||w||^{2}$ = 1 (2 (fw.)(x=)-y=) + No [1w1]) [ssq (w,b) = \frac{1}{2N} \frac{\frac{1}{2} \left(\frac{1}{2} \left(\frac{1} \left(\frac{1}2 \left(\frac{1}2 \left(\frac{1}2 \left(\frac{1}2 \left(\frac{1}2 1 1 1 0 0 0

