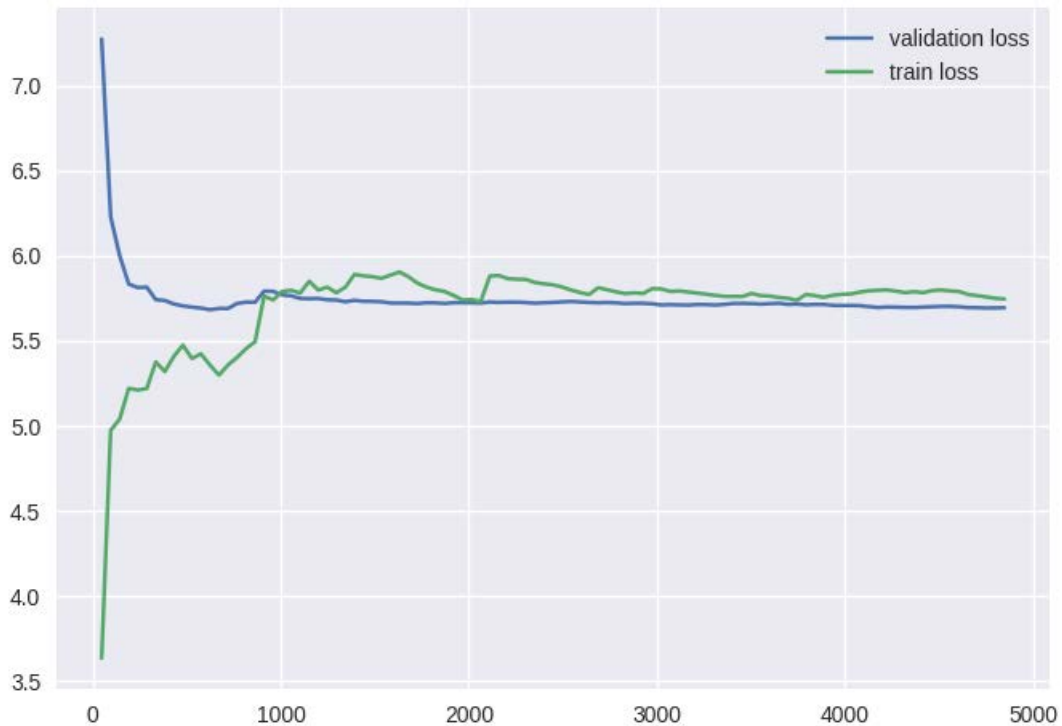


1. 請簡明扼要地闡述你如何抽取模型的輸入特徵 (feature)

答：只用 PM2.5 發現效果不錯後，再逐個參數試驗加入後的 loss，因為自然界中很多參數是在 log 下計算的，試驗後發現 PM2.5+log(SO2)效果最好

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：

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：只用一次 PM2.5 的效果就很好了，如果複雜到 lagrange 多項式則會跑出-164 之類不合理的預測值

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

答：在 cross validation 優化時，lambda 幾乎降到精度誤差範圍，在我所用的簡單 model 下正規化沒有太大的作用

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 $\sum_{n=1}^N (y^n - x^n \cdot w)^2$ 。若將所有訓練資料的特徵值以矩

陣 $X = [x^1 \ x^2 \ \dots \ x^N]$ 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： $(X^T X)^{-1} X^T y$