說明：請各位使用此template進行Report撰寫，如果想要用其他排版模式也請註明題號以及題目內容（請勿擅自更改題號），最後上傳至cool前，請務必轉成PDF檔，並且命名為report.pdf，否則將不予計分。

備註 :

* 所有 advanced 的 gradient descent 技術(如: adam, adagrad 等) 都是可以用的
* 第1～2題請都以題目給訂的兩種model來回答

---------------------------------------------------------------------------

學號：R12945060 系級： 生醫電資所碩二 姓名：羅佳蓉

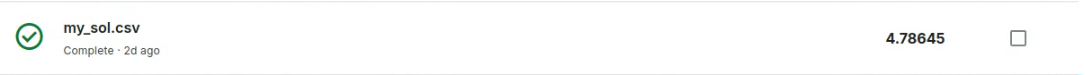
1. (1%) 解釋什麼樣的data preprocessing可以improve你的training/testing accuracy，e.g., 你怎麼挑掉你覺得不適合的data points。請提供數據(例如 kaggle public score RMSE)以佐證你的想法。

* 通過 valid 函數將 PM2.5 > 20 的數據排除，減少了極端異常值對模型的影響，從而提升了訓練和測試集的一致性，模型的學習過程會更加穩定

沒特別排除異常值:



排除異常值:

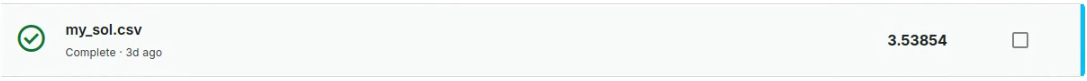


* 選擇了與預測 PM2.5 相關的特徵來進行訓練，避免雜訊特徵對模型的干擾

沒篩選相關特徵:



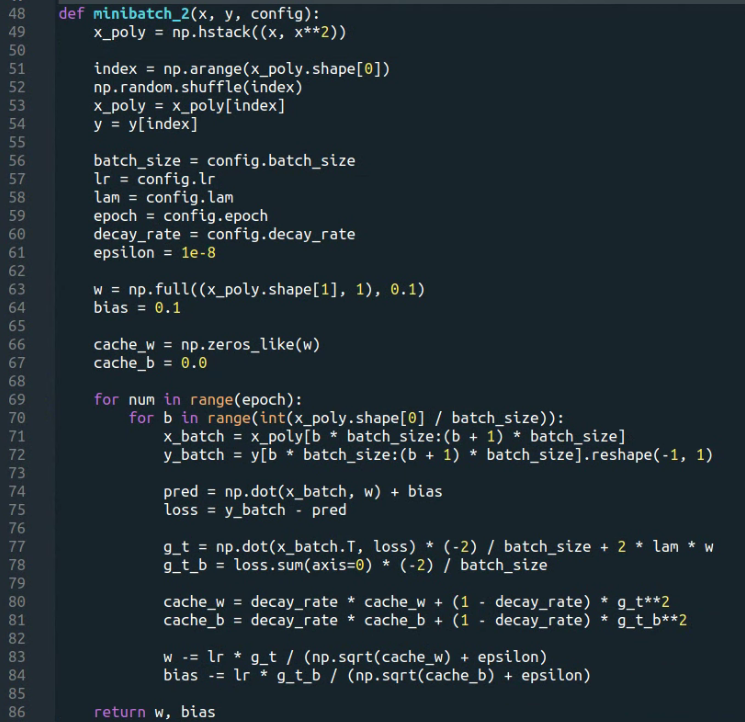
篩選相關特徵:



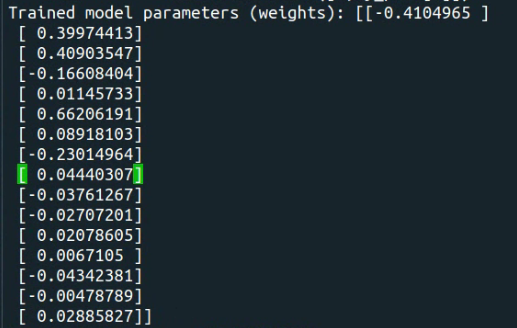
2. (1%) 請實作 2nd-order polynomial regression model (不用考慮交互項)。

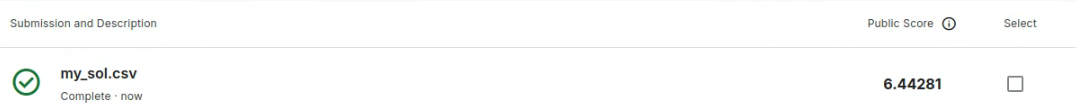


1. 貼上 polynomial regression 版本的 Gradient descent code 內容



1. 在只使用 NO 數值作為 feature 的情況下，紀錄該 model 所訓練出的 parameter 數值以及 kaggle public score.





3.(6%) Refer to math problem:

<https://drive.google.com/file/d/1cOathlUn3Gw4RbwGTttI4neX-tcI_qUo/view?usp=drive_link>