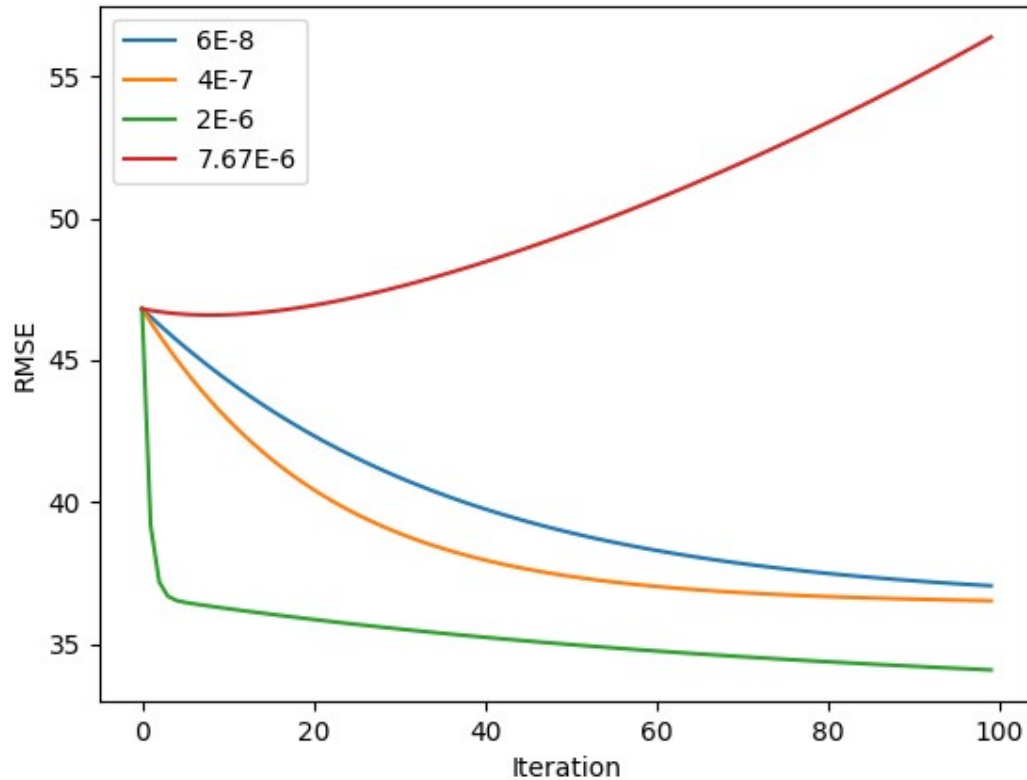


# Homework 1 Report - PM2.5 Prediction

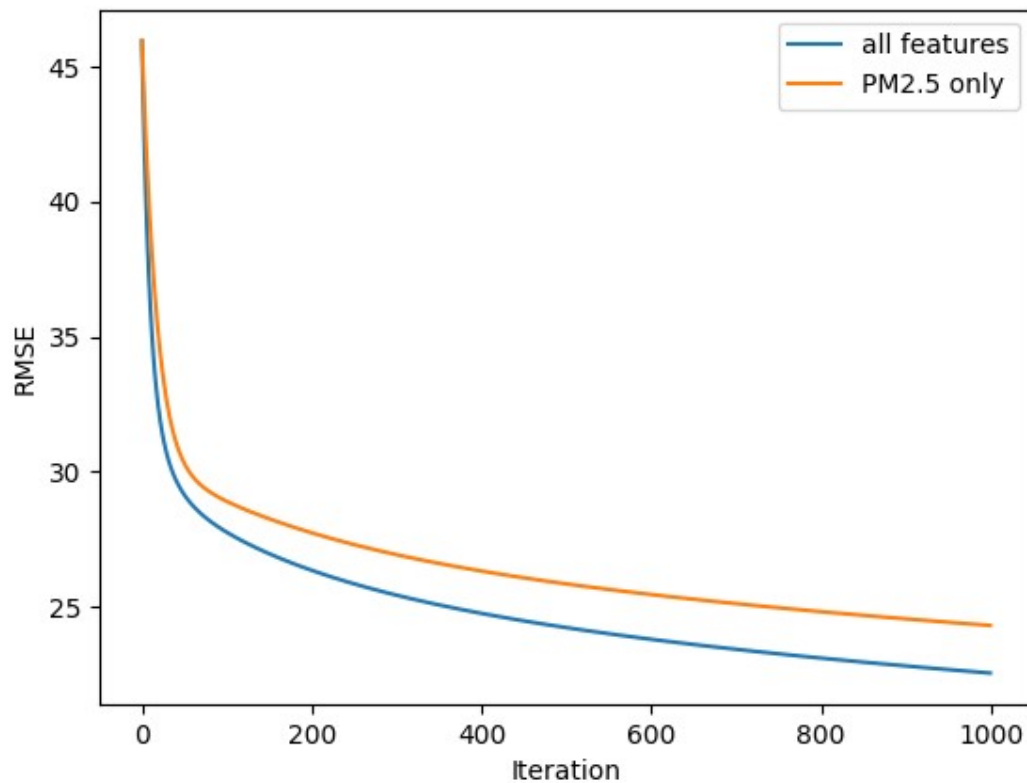
學號: B06901031 系級: 電機二 姓名: 楊宗賢

1. (1%) 請分別使用至少 4 種不同數值的 learning rate 進行 training (其他參數需一致), 對其作圖, 並且討論其收斂過程差異。



$\eta=6 \times 10^{-8}$  和  $\eta=4 \times 10^{-7}$  都太小, 以致於雖然藍線與黃線皆嚴格遞減, 斜率仍然太小;  $\eta=7.67 \times 10^{-6}$  太大了, 因此很快就向外發散;  $\eta=2 \times 10^{-6}$  一下就把 RMSE 從初值 46 減至 36, 且在 iteration 100 處亦有不錯的斜率, 是較剛好的 learning rate。

2. (1%) 請分別使用每筆 data9 小時內所有 feature 的一次項（含 bias 項）以及每筆 data9 小時內 PM2.5 的一次項（含 bias 項）進行 training，比較並討論這兩種模型的 root mean-square error（根據 kaggle 上的 public/private score）。



所有 feature: training RMSE=22.537, public score=10.32572, private score=10.37691;  
只用 PM2.5: training RMSE=24.302, public score=12.30281, private score=12.51373。

兩種 model 的學習曲線形狀類似，但只用 PM2.5 的 model 總是略遜一籌。我認為使用較多的 feature 才能得到最佳的預測結果。

3. (1%)請分別使用至少四種不同數值的 regularization parameter  $\lambda$  進行 training (其他參數需一至), 討論及討論其 RMSE(traning, testing) (testing 根據 kaggle 上的 public/private score) 以及參數 weight 的 L2 norm。

4~6 (3%) 請參考數學題目，將作答過程以各種形式（ $\text{latex}$  尤佳）清楚地呈現在 pdf 檔中（手寫再拍照也可以，但請注意解析度）。