

Assignment 1 – Predicting PM2.5

In this assignment, you will practice using Gradient Descent to predict PM2.5.

Announcement

10/13

- report 中 public + private 分數的意思是：testing中240筆的RMSE，也就是 square root $\{[(\text{public})^2 + (\text{private})^2]/2\}$
- [testing答案釋出!](#)
- hw1_best.sh：選擇的2筆kaggle分數中，private較好的那一個model (註：不用完全一模一樣重現，可接受誤差+0.2)
 - ex. (1) public:6.5/private:5.5 (2) public:6.1/private:5.7
 - 滿足hw1_best.sh 的 private分數 < 5.5 + 0.2 即可
- hw1.sh：public分數 > public simple baseline 即可
- kaggle的成績：只要2筆中任何1筆>某baseline，即得該baseline成績
- github死線：今晚午夜

10/6

- 10/5有通過public simple baseline名單 [📄](#)
 - 因 **kaggle名稱**...等等原因沒有被登記到的同學，請填寫 [📄](#)

重要連結

- 投影片連結 [📄](#)
- 老師講解投影片 [📄](#)
- Kaggle 連結 [🔗](#)
- Github Repo 表單 [📄](#)
- report template [📄](#)
- 遲交表單 [👁](#)
- 10/6(五)上課時間，助教會釋出hw1的[Sample Code](#)以及[Supplementary Slide](#)，同時會邀請小老師帮大家解決程式問題。
- (已截止)小老師教學表單 [👉](#)

The requirements of this assignment are as follows:

- hw1.sh
 - **Python3.5+** required
 - Only (1)numpy (2)scipy (3)pandas are allowed
 - numpy.linalg.lstsq is forbidden.
 - Please handcraft "linear regression" using **Gradient Descent**
 - beat public simple baseline
 - For those who wish to load model instead of running whole training process:
 - please upload your training code named **train.py**
 - as long as there are Gradient Descent Code in **train.py**, it's fine
- hw1_best.sh
 - **Python3.5+** required
 - any library is allowed
 - meet the higher score you choose in kaggle

Data 簡介

- 下載 [train.csv](#) : 每個月前20天每個小時的氣象資料(每小時有18種測資)。共12個月。
- 下載 [test.csv](#) : 排除train.csv中剩餘的資料, 取連續9小時的資料當feature, 預測第10小時的PM2.5值。總共取240筆不重複的test data。
- 下載 [sampleSubmission.csv](#)

作業修正&講解

- report第五題題目修正: $X = [x^1 \ x^2 \ \dots \ x^N]$ 改為 $X = [x^1 \ x^2 \ \dots \ x^N]^T$
- 第1-3題請都以題目給訂的兩種model來回答

FAQ

Q1. 為了回答report(1)-(3)是不是要上傳kaggle 8次,這樣會浪費kaggle上傳的coda

Ans. 同學可以先把model訓練好的答案做好, 等到kaggle死線之後便可以無限上傳看error了

Q2. 如果預先對training data做normalization, 那我可以上傳train.csv然後在hw1.sh中自己讀進來嗎

Ans. 可以



Posted by: [ntumlta](#)

Contact information: ntu.mlta@gmail.com .

Course information: [Machine Learning \(2017, Fall\) @ National Taiwan University](#).

