

108062608 黃柏翰

Problem A

1. nn.linear 的output 維度 為 (batch, 1), 但是label的維度為(1, batch), 讓一開始train的結果很不好
2. learning 大一點比較有效率, 因為iteration 太少

Problem B

1. 常常有GPU out of memory 的問題, 利用restart kernel來解決
2. 求的loss 如果用累加在除上平均有問題, 因為是batch load data, 所以用 np.mean(error list), 才能求得over all average error in one epoch