**23280673810823**

**Things that need more brainstorming:**

Need more different methods to to predict the data

(ie. (mse和cross entropy, sgd和adam, ))(requirement ***1***)

Detect under- and over-fitting(network too deep??vanishing gradient proble->use RELU Training error high->make model more complex), how to address them (requirement ***4***)

Ps: simple model: high bias , low variance

Complex model: low bias, high variance

**Things I am already working on or trying to figure out how to do:**

Feature selection, design (requirement ***3***): PCA(too many features->dimension reduction, create feature(can visualize it))

Exploration of at least one or two techniques on which we did not spend significant time in class (requirement ***2***) eg: neural network, CNN

Performance validation(requirement ***4***):

n-fold cross validation->do model selection (plot error rate and choose),

combine models(adaboost), etc.

還沒有空做的: 把preprocessed的data存起來這樣之後就不用每次跑都preprocess

xq