hyperparameter tuning

gridsearchcv

randomsearchcv

parameter

model parameter and hyperparameter

hyperparamter tuning refers to the process of chooisng the optimum set of hyperparameters fora machine learning model, This process is also called hyperparamter optmization

gridsearchcv is identifying all combinations and finding which combination finds the best outcome

try hyperparameters with all models

task 1: Select a set of models that we will try with our dataset

for each of the model we will try different parameters

1 hyper parameter and vary the other 2

split the data into training and testing set 70% training 30% testing

x\_train x\_test, ytrain ytest = train\_test\_split(feature1, feature 2, test size = 0.3

model = svm.svc(kernel='rbf', (c=30, gamma='auto')

model.fit(x\_Train,ytrain)

model.score(xtest, ytest)

cross validation (define how many fold we will use)

build model, get a score

reshuffling do again get a new score

reshuffle do again and get a new score

do for each fold

we do this to avoid overfit and underfit

we take all scores and and take the average and we use it for the

cross\_val\_score(svm.svc(kernel='linear', c=10,gamma'auto'),data, target, cv=5)

we can use gridsearchcv to automatically try the c and kernel

from sklearn.model\_selection import GridSearchCV

clf = GridSearchCV(svm.SVC(gamma='auto), {

'C': [1,10,20],

'kernel': ['rbf', ]

}, cv=5, return\_train\_score=False)

clf.best\_params\_

clf.best\_score\_

we can add n\_iter=(int) for randomsearchCV and it will do the same thing but randomly select the int

we can get the parameters for each model by searching sklearn and look at the modules. find what we need to tune for our data.

2-3 slides to explain the first half of the project

we need to do the technical aspect

we need to summarize what we learned the 4 months learning and our project in 4-5 bullet points