参考blog:

<http://www.hankcs.com/ml/adaboost.html>

1.代码相关

（1）代码为：

AdaBoost1.py #简单数据集

AdaBoost2.py #复杂数据集

（2）子函数：

loadSimpleData(filename) #载入数据

plotData(dataMat,classLabels) #可视化数据

adaBoostTrain(dataMat,classLabels,M=40) #训练AdaBoost---给定分类器数目M,返回弱分类器weakClass和fx

stumpClassify(dataMat,dimen,thre,ineq) #已知特征维度dimen,分类边界thre,符号Ineq----求预测的label

bulidStump(dataMat,classLabels,D) #已知权值矩阵D和训练数据dataMat和classLabels----求最优的Gm(x)

adaClassify(datToclass,weakClass) #提升树分类----输入分类器和测试数据集datToclass,输出类别

evaluate(fx,classLabels) #查准率/召回率

train\_test(dataMat,classLabels,dataMatTest,classLabelsTest,M) #输出查准率/召回率 ---- 输入训练集和测试集，输出prTrain,prTest

2.注意点

（1）数据读取：

data = pd.read\_table(‘文件名’)

dataSet = data.iloc[:,:-1]

labelSet = data.iloc[:,-1]