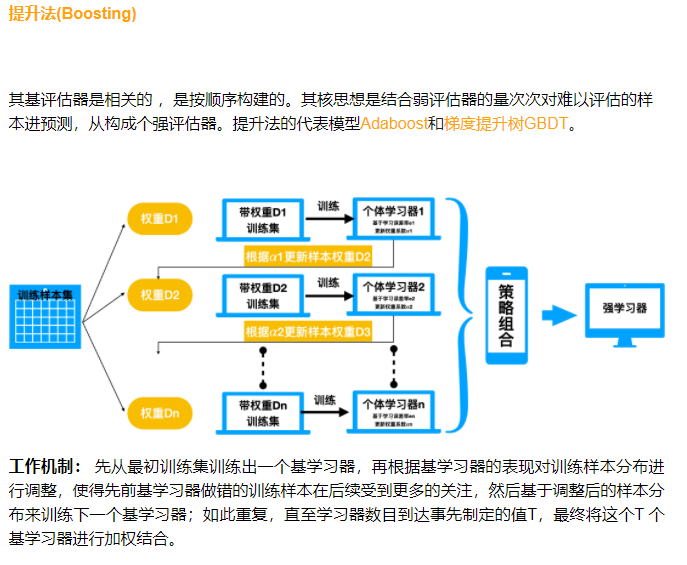
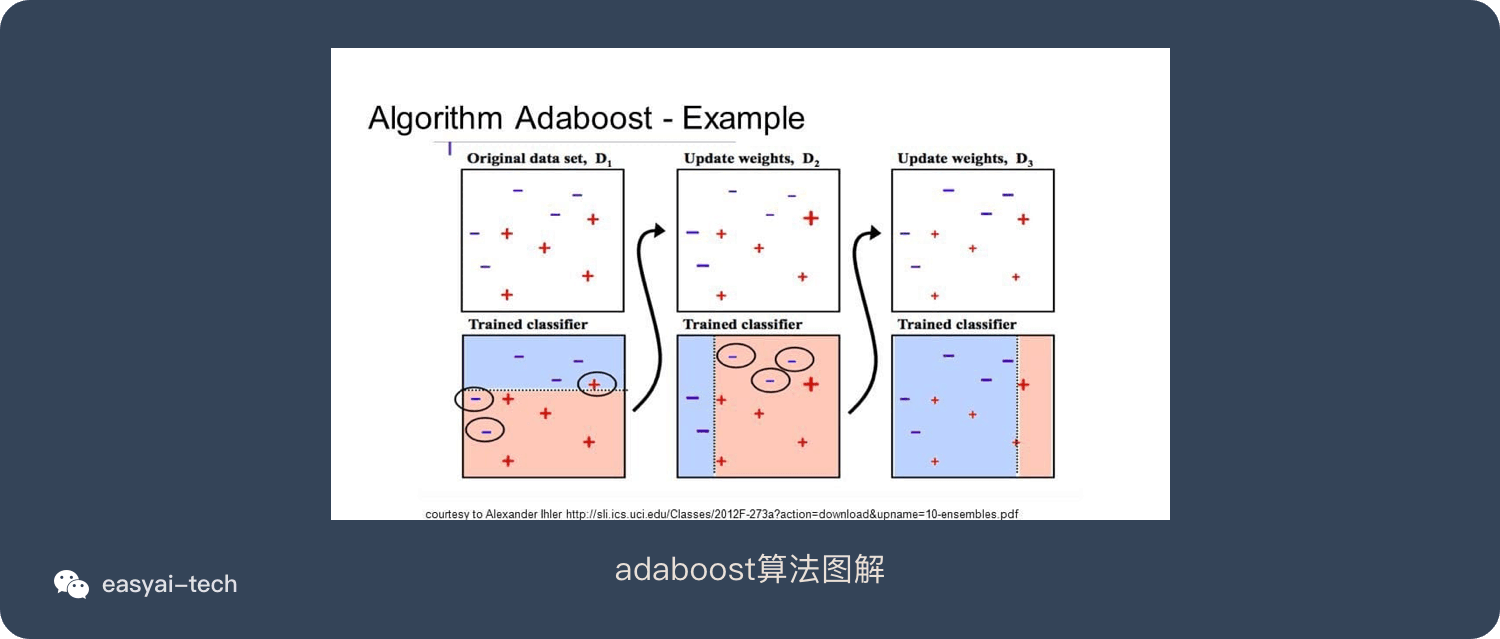
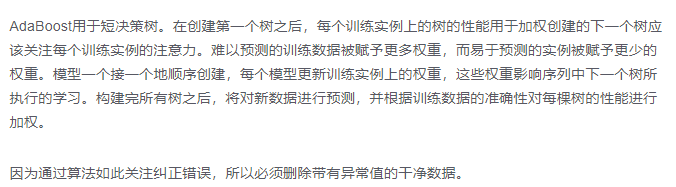
# 原理阶段

1. **Boosting是什么？**



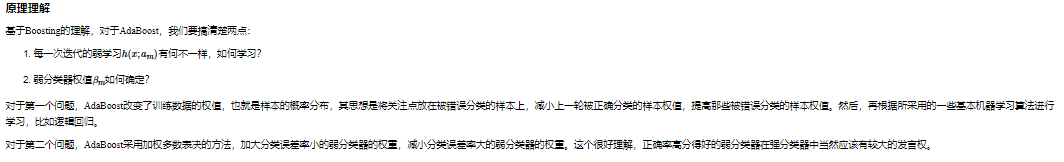
1. **Adaboost是什么？**



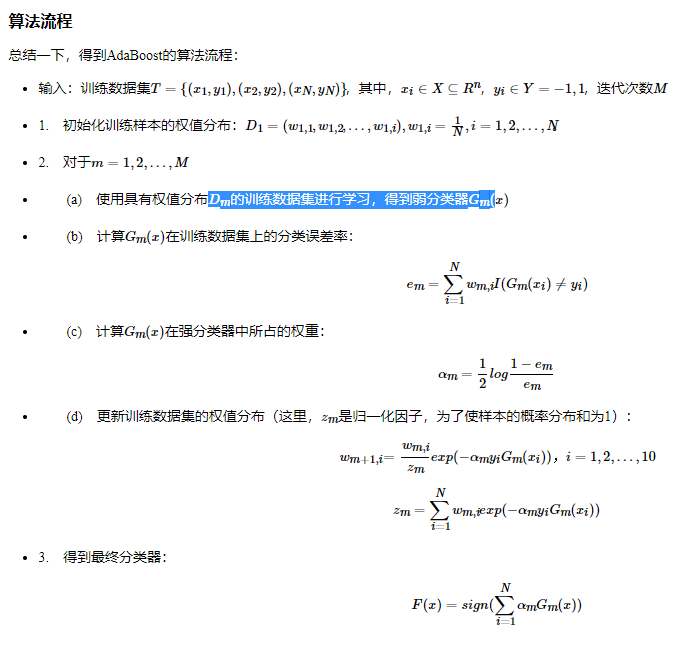


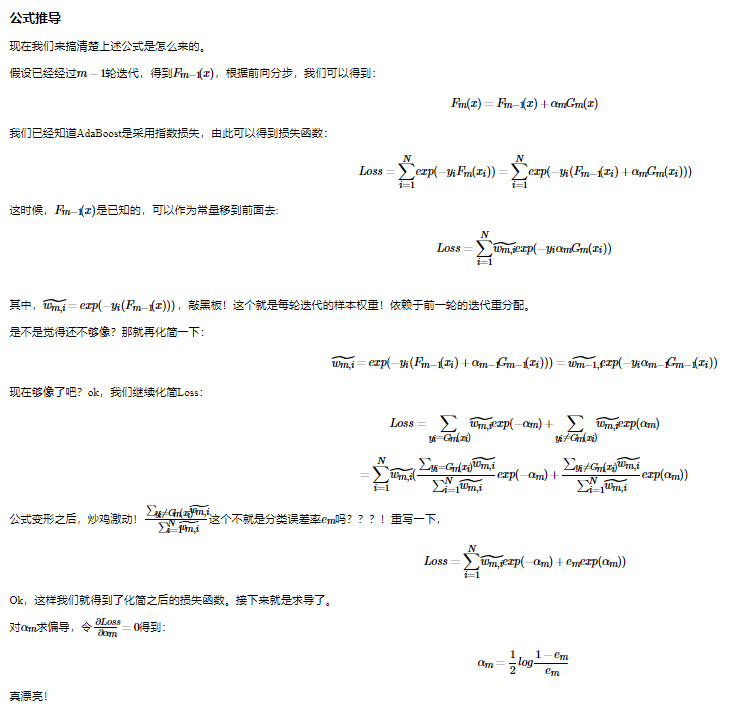
1. **Adaboost原理**

https://www.cnblogs.com/ScorpioLu/p/8295990.html









1. **Adaboost优缺点**

1、 优点

（1）很好的利用了弱分类器进行级联

（2）可以将不同的分类算法作为弱分类器

（3）AdaBoost具有很高的精度

（4）相对于bagging算法和Random Forest算法，AdaBoost充分考虑的每个分类器的权重

2、缺点

（1）AdaBoost迭代次数也就是弱分类器数目不太好设定，可以使用交叉验证来进行确定

（2）数据不平衡导致分类精度下降

（3）训练比较耗时，每次重新选择当前分类器最好切分点

1. **Bagging和Boosting二者之间的区别**

1）样本选择上

Bagging：训练集是在原始集中有放回选取的，从原始集中选出的各轮训练集之间是独立的。

Boosting：每一轮的训练集不变，只是训练集中每个样例在分类器中的权重发生变化。而权值是根据上一轮的分类结果进行调整。

2）样例权重

Bagging：使用均匀取样，每个样例的权重相等

Boosting：根据错误率不断调整样例的权值，错误率越大则权重越大。

3）预测函数

Bagging：所有预测函数的权重相等。

Boosting：每个弱分类器都有相应的权重，对于分类误差小的分类器会有更大的权重。

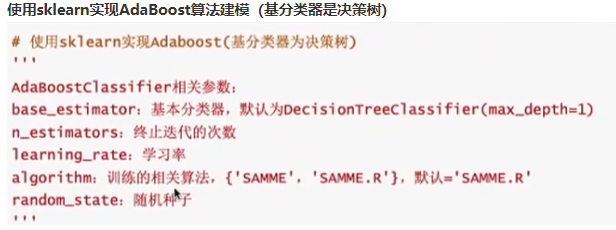
4）并行计算

Bagging：各个预测函数可以并行生成

Boosting：各个预测函数只能顺序生成，因为后一个模型参数需要前一轮模型的结果。

# 代码阶段

1. **AdaBoostClassifier参数**



1. **SAMME算法(头疼)**

https://zhuanlan.zhihu.com/p/82568979