**AdaBoost**

boosting算法是一类将弱学习器提升为强学习器的[集成学习算法](https://blog.csdn.net/weixin_38629654/article/details/80515680" \t "_blank)，它通过改变训练样本的权值，学习多个分类器，并将这些分类器进行线性组合，提高泛化性能。

一个分类，如果存在一个多项式算法能够学习他，并得到很高的正确率，那么这个算法称为强学习器，反之如果正确率只是稍大于随机猜测（50%），则称为弱学习器。

AdaBoost，是英文”Adaptive Boosting“（自适应增强）的缩写，由Yoav Freund和Robert Schapire在1995年提出。

它的自适应在于：前一个弱分类器分错的样本的权值（样本对应的权值）会得到加强，权值更新后的样本再次被用来训练下一个新的弱分类器。在每轮训练中，用总体（样本总体）训练新的弱分类器，产生新的样本权值、该弱分类器的话语权，一直迭代直到达到预定的错误率或达到指定的最大迭代次数。

**算法原理**

（1）初始化训练数据（每个样本）的权值分布：如果有N个样本，则每一个训练的样本点最开始时都被赋予相同的权重：1/N。

（2）**训练弱分类器**。具体训练过程中，如果某个样本已经被准确地分类，那么在构造下一个训练集中，它的权重就被降低；相反，如果某个样本点没有被准确地分类，那么它的权重就得到提高。同时，得到弱分类器对应的话语权。然后，更新权值后的样本集被用于训练下一个分类器，整个训练过程如此迭代地进行下去。

（3）将各个训练得到的弱分类器组合成强分类器。各个弱分类器的训练过程结束后，**分类误差率小的弱分类器的话语权较大，其在最终的分类函数中起着较大的决定作用，而分类误差率大的弱分类器的话语权较小，其在最终的分类函数中起着较小的决定作用。**换言之，误差率低的弱分类器在最终分类器中占的比例较大，反之较小。