

LR(逻辑斯蒂回归)和 SVM(支持向量机)都是分类算法;

如果不考虑核函数, LR 和 SVM 都是线性分类算法, 也就是它们的分类决策面都是线性的;

两者本质区别是其损失函数不同, LR 使用的是交叉熵损失函数(cross entropy loss), SVM 使用的是铰链损失函数(hinge loss).

这两个损失函数的目的都是增加对分类影响较大的数据点的权重, 减少与分类关系较小的数据点的权重。SVM 只考虑支持向量, 即和分类最相关的点, 而 LR 通过非线性映射, 大大减小了离分类平面较远的点的权重。而且 SVM 的损失函数就自带正则(损失函数中的 $\frac{1}{2}||w||^2$ 项), 而 LR 必须另外在损失函数上添加正则项,

SVM 只考虑局部的边界线附近的点, 而 LR 考虑全局(远离的点边界线的确定也起作用)。

Adaboost 则属于集成学习方法的一种, 其主要思想是把多个弱分类器结合起来, 得到的结果会比一个强决策器更好。其实现是一个迭代渐进的过程, 从一个最基础的分类器开始, 每次寻找一个最能解决当前错误样本的分类器。用加权取和的方式把这个新分类器结合进已有的分类器中。

Adaboost 的好处是自带了特征选择, 只使用在训练集中发现有效的特征。这样就降低了分类时需要计算的特征数量, 也在一定程度上解决了高维数据难以理解的问题。