前面几篇笔记分别学习了线性回归、回归和这几种机器学习算法，不同的学习算法有不同的判别函数，如何衡量判别函数的优劣？一般我们就用损失函数衡量。

假设有一组训练数据，判别式为，输出：



令：



下面介绍下几种常用的

1. 

最理想的状态是：当预测出的值与不一样的时候，损失1，否则没损失为0，如下式所示（可以理解为在上犯了几次错，当然越小越好）：



明显，（1）式不可微分，这导致我们无法用梯度下降法求出最好的，既然我们没有办法最小化这个式子（不可微分），那我们就重新定义一个：



其中是什么形式，可以自己定义。

1. ：

上面讲到，是什么形式，我们可以自己定义。现在我们希望当时，越接近越好；时，越接近越好：

 

 



将上面两种的描绘出来如下图所示：

从上图就能看出有一个很大的不合理的地方：当很大的时候，

会出现一个很大的，这个不是我们期望的，我们期望的是当很大的时候，越小越好。

1. +：

现在对做个再输出，则我们希望当时，越接近越好；时，越接近越好：

 

 



4）+：

根据之前的笔记内容，我们可以知道 的代价函数为：



5）：

在中，我们用表示第个训练样本违背边界的大小，此时：



当训练样本在边界之外时，即，由于，得：；

当训练样本在边界上时，即，得：；

当训练样本在边界内部时，即，得：；

综上：



此时，误差被称为；



综上几种的如下图：



由上图可知：

当变化相同的长度时，的变化上面：

比下降的更快，即更合理一些；

和都是的上确界，都只能近似0/1误差的解，但是它们二者的优劣并没有确定的定论，不同情况下不一样。



2017.07.31.night