读书报告

71117203-袁佳怡

1. （必填）自己提出的问题的理解：
2. 提出的问题1：怎么理解当训练数据集线性可分时，感知机学习算法存在无穷多个解，其解由于不同的初值或不同的迭代顺序而可能有所不同。

讨论后的理解：初值或者迭代顺序不同时会得到不同的超平面去对数据进行二分类，可能就是在一个区域范围内，那些超平面有一定的符合区间，相当于进行微调也是满足分割的。

1. （必填）别人提出的问题的理解
2. 问题2：为什么“损失函数的一个自然选择是误分类点的总数”？

自己的理解：因为这是自然情况下对损失函数估计最直观的判断方式，即利用误分类点的数目衡量划分的准确度，通过减少误分类点的数目减小损失函数，提高准确度。但因为，这样的损失函数不是参数w, b的连续可导函数，不易优化，所有没能被采用。

1. 问题3：如何理解实例点更新次数越多，它距离分离超平面的距离越近？

自己的理解：因为每一次更新w和b，都是为了让平面离这个实例点更近，来让它满足分割，当一个实例点离超平面越近，它被扰动的可能性就越高，然后它被调整的次数就越多。

1. （必填）读书计划

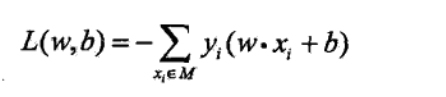
1、本周完成的内容章节：《统计学习方法》第二章

2、下周计划：《统计学习方法》第三章

四、（选做）读书摘要

1、 所谓感知机，其实就是一个在n维空间内的超平面（n-1维），这个超平面将整个空间分为两部分，超平面S又称为分离超平面。

2、 感知机学习策略（线性可分）a.对于一个线性可分的训练集，我们能够找到（至少）一个分离超平面，该超平面能把训练集的正、负实例点完全地分离开。为了找到该平面，需要确定一个学习策略，即定义经验损失函数。b. 感知机采用的损失函数是：误分类点到超平面的距离，即：。而又因为其是误分类点，yi与w\*xi+b异号，即，所以误分类点到超平面的距离又为：。

3、 若超平面S的误分类点集合为M，则所有误分类点到超平面的总距离为：，如果如考虑1/||w||（对损失函数乘上一个非0常数并不会影响极值点的取值），那么就得到感知机的损失函数：

1. 感知机学习算法：该算法使用的是梯度下降，但在极小化的过程中并非一次是的M中的误分类点的梯度下降，而是随机选取一个误分类点进行梯度下降。
2. 算法的收敛性：在训练集线性可分的时候，感知机学习算法的原始形式会收敛，而不可能一直迭代下去。即经过有限次迭代，可以得到一个将训练集完全正确划分的分离超平面及感知机模型。