方骏-2020年3月15日-读书报告

1. 自己提出的问题的理解：
2. 提出的问题1：书中任一点到x0到超平面S的距离是怎么推导的？

讨论后的理解：首先，设点在平面S上的投影为，则有：。由于向量与S平面的法向量w平行，所以，，所以，那么距离公式就可以得到了。

1. 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：这只是一个显然的想法。就是最初的损失函数的想法，损失函数本来就应该实现这样的功能，即能够体现出误分类的情况就可以了，所以一个自然的想法就是误分类点的总数可以作为损失函数，这就是最初的显然想法而已。

1. 问题3： 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。

自己的理解：可能存在一定范围内，那个分割的平面都可行，只需要对得到的超平面解进行一定的微调，在这种微调中，就有无穷多个解，显然通过选择不同的初值和迭代顺序就可以得到不同的解。

1. 问题4： 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：考虑调整其它实例点的时候 重新对该实例点进行调整，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，如果该实例点离超平面越近 它被扰动的可能性就越高，然后调整的次数自然就越多。

1. 读书计划

1、本周完成的内容章节：《统计机器学习》第二章

2、下周计划：《统计机器学习》第三章

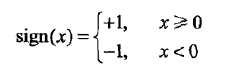
四、读书摘要及理解

1、感知机是二类分类的线性分类模型，其输入为实例的特征向量，输出为实例的类别，取+1和-1二值。感知机对应于输入空间中将实例划分为正负两类的分离超平面，属于判别模型。感知机学习旨在求出将训练数据进行线性划分的分离超平面，为此，导入基于误分类的损失函数，利用梯度下降法对损失函数进行极小化，求得感知机模型。

2、感知机假设输入空间是n元向量，输出空间是+1和-1。输入是表示实例的特征向量，对应于输入空间的点；输出表示实例的类别，函数为：



是两者的内积运算，sign是符号函数，即：



感知机是一种线性分类模型，属于判别模型。感知机模型的假设空间是定义在特征空间中的所有线性分类模型或线性分类器。感知机的几何解释为：线性方程，对应于特征空间中的一个超平面S，其中w是超平面的法向量，b是超平面的截距。这个超平面将特征空间划分为两个部分，分别为正负两类。感知机预测，通过学习得到的感知机模型，对于新的输入实例给出其对应的输出类别。

3、假设训练数据集是线性可分的，感知机学习的目标是求得一个能够将训练集正实例点和负实例点完全正确分开的分离超平面。为了找出这样的超平面，即确定感知机模型参数w,b，需要确定一个学习策略，即定义（经验）损失函数并将损失函数极小化。

损失函数被定义为误分类点到超平面S的总距离：



这显然是非负的，如果没有误分类点，损失函数值是0。而且，误分类点越少，误分类点离超平面越近，损失函数值就越小。

4、感知机学习算法是误分类驱动的，具体采用随机梯度下降法。首先，任意选取一个超平面，然后用梯度下降法不断地极小化目标函数。极小化过程中不是一次使M中所有误分类点的梯度下降，而是一次随机选取一个误分类点使其梯度下降。当训练集线性可分时，感知机学习算法是收敛的。

感知机学习算法还有一个对偶形式，基本想法是，将w和b表示为实例xi和yi的线性组合的形式，通过求解其系数而求得w和b。