**2019/03/08读书报告**

学号 71117228 姓名 李泓烨

读书进度：统计机器学习第1章

1. **读书报告内容**

**1. 自己提出问题的理解**

① 1.7中，生成方法和判别方法分别适用于什么情况？

**讨论后的理解**：在监督学习中，两种方法各有优缺点，适合于不同条件的学习问题。

生成方法的特点：上面说到，生成方法学习联合概率密度分布P(X,Y)，所以就可以从统计的角度表示数据的分布情况，能够反映同类数据本身的相似度。但它不关心到底划分各类的那个分类边界在哪。生成方法可以还原出联合概率分布P(Y|X)，而判别方法不能。生成方法的学习收敛速度更快，即当样本容量增加的时候，学到的模型可以更快的收敛于真实模型，当存在隐变量时，仍可以用生成方法学习。此时判别方法就不能用。

判别方法的特点：判别方法直接学习的是决策函数Y=f(X)或者条件概率分布P(Y|X)。不能反映训练数据本身的特性。但它寻找不同类别之间的最优分类面，反映的是异类数据之间的差异。直接面对预测，往往学习的准确率更高。由于直接学习P(Y|X)或P(X)，可以对数据进行各种程度上的抽象、定义特征并使用特征，因此可以简化学习问题。

**参考资料**：https://www.cnblogs.com/Allen-rg/p/10270390.html

② 罚项J(f)有没有一个确定的公式呢？不太理解什么叫“定义在假设空间上的泛函”

**讨论后的理解**：它用来体现一个算法的复杂度，具体怎么体现书上并没有明确说明，不同场景对应的泛函公式并不相同

**2. 别人提出问题的理解**

1. **提出的问题1**：怎么理解，从贝叶斯估计的角度来看，正则化项对应于模型的先验概率，可以假设复杂的模型有较小的先验概率，简单的模型有较大的先验概率？

**讨论后的理解**：先验概率：根据以往的经验得到的概率；可以理解为复杂的模型被采用的可能性小

2. **提出的问题2**：怎么理解极大似然估计是经验风险最小化的一个例子？

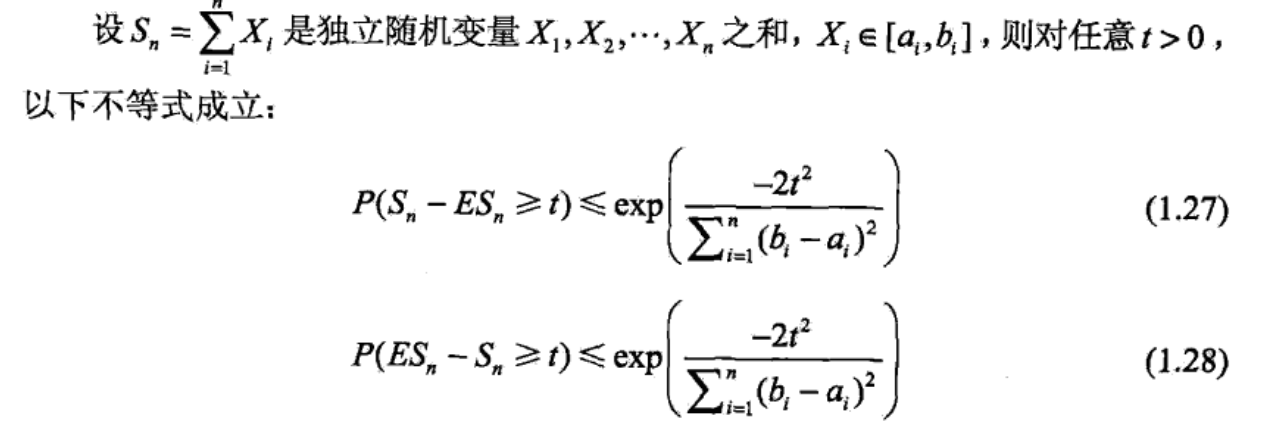
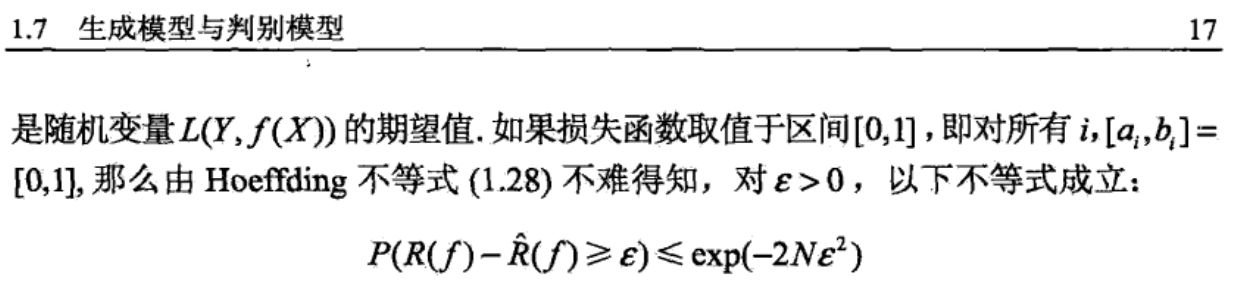
**小组成员的解释**：极大似然估计是建立在这样的思想上：已知某个参数能使这个样本出现的概率最大，我们当然不会再去选择其他小概率的样本，所以干脆就把这个参数作为估计的真实值

3. **提出的问题3**：能否简单解释一下联合分布具体指什么？

**自己的理解**：大一的同学提的，还没有学概率论，哈哈哈；比如说一个随机现象需要用几个变量去表述

4. **提出的问题4**：奥卡姆剃刀原理中，很好的解释已知数据评判标准是什么？

**自己的理解**：能解释数据的模型很多，同等情况下越简单越好；所以说在相同的泛化误差下，先验概率越小越好

5. **提出的问题5**：如何理解泛化公式上界的推导过程？

**自己的理解**：R就是1/N的S，可以把ai,bi变为ai/n,bi/n，右边放缩一下上面N2下面N最后就是N

**3. 读书计划**

① 本周完成的内容章节：看完了第1章

② 下周计划：看完第2章

**4. 读书摘要、理解&伪代码的具体实现**

1. 读书摘要

这里我以思维导图的形式记录了我的读书笔记（见下一页）

