

统计学习第四次读书报告

报告人：71117415-喻泽弘

读书时间：3.23-3.30

读书进度：统计学习第四章

问题与解答

- 个人提出的问题：

- 朴素贝叶斯能否解决不独立同分布的情况？

无法解决，在朴素贝叶斯中，独立同分布是基本假设，如果不满足该假设，就无法使用朴素贝叶斯方法。可以考虑其它的贝叶斯方法解决该问题，不过模型会更加复杂。

- 别人提出的问题：

- 最大似然估计的直观理解？

最大似然估计的目的就是：利用已知的样本结果，反推最有可能导致这个结果的参数值。

原理：极大似然估计是建立在极大似然原理基础上的一个统计方法，是概率论在统计学中的应用。极大似然估计提供了一种给定观察数据来评估模型参数的方法，即：“模型已定，参数未知”。通过若干次试验，观察结果，利用试验结果得到某个参数值能够使样本出现的概率最大，则为极大似然估计。

- 条件假设等于是说用于分类的特征在类确定的条件下都是条件独立的，这一假设使朴素贝叶斯法变得简单，但为什么说有时也会牺牲一定的分类准确率？

现实情况下，属性之间并不一定能够满足条件独立，因此，此时不适用于贝叶斯模型，从而会对该类数据进行预测时，会牺牲一定的分类准确率

下周计划

完成统计学习第五章的阅读

学习总结

- 朴素贝叶斯的基本假设是条件独立，这是一个较强的假设。由于这一假设，模型包含的条件概率的数量大大减少，朴素贝叶斯的学习与预测大为简化。因而朴素贝叶斯高效，且易于实现。其缺点是分类性能不一定很高