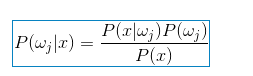
朴素贝叶斯分类器

贝叶斯分类器就是基于这条公式发展起来的，之所以这里还加上了朴素二字，是因为该分类器对各类的分布做了一个假设，即不同类的数据样本之间是相互独立的。这样的假设是非常强的，但并不影响朴素贝叶斯分类器的适用性。1997年，微软研究院的 Domingos 和 Pazzani 通过实验证明，即使在其前提假设不成立的情况下，该分类器依然表现出良好的性能。对这一现象的一个解释是，该分类器需要训练的参数比较少，所以能够很好的避免发生过拟合（overfitting）。



分类器的训练分两步：

计算先验概率；

计算似然函数；

应用过程只需利用训练过程中得到的先验概率和似然函数计算出后验概率即可。

所谓先验概率，其实就是每个类出现的概率，这个是个简单的统计问题，即把训练数据集中不同类所占的比值都计算出来即可。

训练似然函数与此类似，就是看各个特征对应的值属于某个类的概率值。

https://blog.csdn.net/yunduanmuxue/article/details/39693917