**判别学习算法和生成学习算法**

判别学习算法：直接学习p(y|x)，即直接通过输入特征空间x去确定目标类型{0,1}，比如Logistic Regression和Linear Regression以及感知学习算法都是判别学习算法。

生成学习算法：不直接对p(y|x)建模，而是通过对p(x|y)和p(y)建模。比如，y表示目标是dog(0)还是elephant(1)，则p(x|y=1)表示大象的特征分布，p(x|y=0)表示狗的特征分布。

生成学习算法通过学习p(y|x)和p(y)，一般都要通过贝叶斯公式转化为p(x|y)来进行预测。

IMG_256

最大释然估计也可以转换为联合概率的最值。

IMG_257

以评论正负面情感的判断为例（只考虑“开心”这一个词的特征情况下）

判别学习算法：从样本中直接学习到‘开心’一词出现在评论，则该评论属于正面评论

生成学习算法：从样本中分别统计正负面评论的概率，以及出现“开心”一词评论的概率，由贝叶斯公式计算出如果一条评论中出现了“开心”一词，分别属于正负面的概率，取最大的一类

**高斯判别分析（Gaussian Discriminant Analysis）**

高斯判别是一种生成学习算法。对于输入特征x是连续值的随机变量，使用高斯判别分析模型非常有效，它对p(x|y)使用高斯分布建模。

IMG_258，其中p为先验概率

IMG_259

IMG_260

依据前面对生成学习算法的分析，求联合概率的最大似然估计，

IMG_261

这个公式比较复杂，但可类比上面贝叶斯公式中的“最大释然估计”，跟概率密度相关；

公式求解意义在于比较y0和y1情况下哪个的值最大，分类的时候取值最大的那一类；

公式的求解依赖下面几个重要参数。

求得4个参数值及其直观解释为：

IMG_262

直观含义：类目1的样本数占总样本数的比例，即先验概率，类目0的先验概率刚好是 1 − ϕ

1代表正面评论，0代表负面评论的话，这里ϕ就是求正面评论在样本中出现的概率，正负面概率相加为1，所以负面的概率是1 − ϕ

IMG_263

直观含义：类目0每个维度特征的均值，结果是nx1的向量，n为特征维度

可以分开来看，分母是负面评论总数；分子是nx1的向量，意思是我们使用n维的特征时负面中每个特征出现的次数

IMG_264

直观含义：类目1每个维度特征的均值，结果是nx1的向量

与u0相似

IMG_265

这一步是比较直观的矩阵运算，意思是计算一条评论中所有特征与正负面类别间“距离”

总的来说，高斯判别的时候就是看输入的向量与正、负哪类更接近，以及根据贝叶斯公式，还需要考虑到正负两类本身发生的概率，综合比较他们的大小，哪类大就归属于那一类中。