

题目：On Discriminative vs. Generative classifiers: A comparison of logistic regression and naive Bayes

吴恩达大大的论文，主要讲生成式模型与判别式模型的收敛速度与准确性的比较，基本为数学证明。

无论是生成式模型还是判别式模型，都可作为分类器使用，分类器的数学表达即为：给定输入 X 以及分类变量 Y ，求 $P(Y|X)$ 。判别式模型直接估算 $P(Y|X)$ ，或者也可像 SVM 那样，估算出输入和输出之间的映射，与概率无关；判别式模型的典型代表是：logistic 回归；产生式模型的思想是先估计联合概率密度 $P(X,Y)$ ，再通过贝叶斯公式求出 $P(Y|X)$ ；生成式模型的典型代表则是：朴素贝叶斯模型。

一般认为判别式模型更受欢迎，“人们更应该直接去解决问题，永远不要把求解更复杂的问题作为中间阶段”（Vapnik），产生式模型（朴素贝叶斯）在少量样本的情况下，可以取得更好的精确率，并且速度更快，判别式模型（logistics 回归）在样本增加的情况下，逐渐逼近前者的性能，但速度相对较慢。