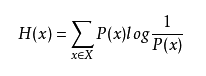
Active Basic Model:

The Kullback-Leibler divergence：KL散度（相对熵），是描述两个概率分布P和Q差异的一种方法。非对称性，不满足三角不等式。

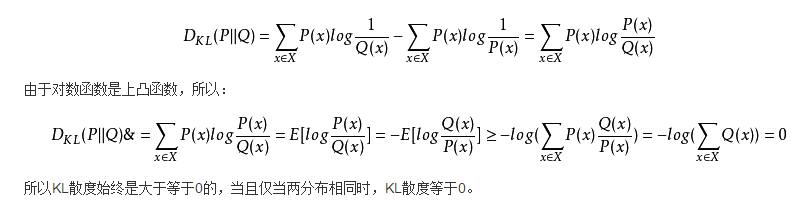
KL散度是用来度量使用基于Q的编码来编码来自P的样本平均所需的额外的比特个数。

给定一个字符集的概率分布，我们可以设计一种编码，使得表示该字符集组成的字符串平均需要的比特数最少。假设这个字符集是X，对x∈X，其出现概率为P(x)，那么其最优编码平均需要的比特数等于这个字符集的熵：



在同样的字符集上，假设存在另一个概率分布Q(X)。如果用概率分布P(X)的最优编码，来符合分布Q(X)的字符编码，那么表示这些字符就会比理想情况多用一些比特数。KL散度就是用来衡量这种情况下平均每个字符多用的比特数 。

http://d.hiphotos.baidu.com/baike/s%3D57/sign=98b20b874836acaf5de096fb7dd9d4a3/8694a4c27d1ed21b5da9665bae6eddc451da3f2f.jpg 为字符编码长度



ROC Curve and AUC Score

主要用来评价二值分类器的优劣。

ROC(Receiver Operating Characteristic)

AUC（Area Under Curve）

AOG：非监督学习中pervasive ambiguity如何解决

EM算法就是通过迭代地最大化完整数据的对数似然函数的期望，来最大化不完整数据的对数似然函数