为简单起见，做如下假设：假设样本空间χ和假设空间 H 是离散的。令$P(h|X,\mathcal{L}_a)$代表算法$\mathcal{L}_a$基于训练数据 X 产生假设 h 的概率。令f代表我们希望学得的真实函数，对于二分类问题，假设真实目标函数可以是任何函数$\chi \mapsto \{0,1\}$，且所有可能的f符合均匀分布。

令 $l(h(x),f(x))$为性能度量函数，用分类误差的和作为函数性能的评估，则有：

$$\sum_f E_{ote}(\mathcal{L}_a|X,f) = \sum_f \sum_h \sum_{x\in\chi-X} p(x)l\big(h(x),f(x)\big)p(h|X,\mathcal{L}_a)$$

$$= \sum_{x\in\chi-X} p(x) \sum_h p(h|X,\mathcal{L}_a) \sum_f l\big(h(x),f(x)\big)$$

对于评估函数 l，这里假设是二分类问题，所以对应样本空间χ所有的f可能为$2^{|\chi|}$。由于f为均匀分布，所以通过算法学得的$h(x)$中，必有一半的f与$h(x)$的预测值相等。因此：

$$\sum_f l\big(h(x),f(x)\big) = \frac{1}{2}2^{|\chi|}l\big(h(x)=f(x)\big) + \frac{1}{2}2^{|\chi|}l\big(h(x)\neq f(x)\big)$$

$$= 2^{|\chi|-1}(l\big(h(x)=f(x)\big) + l\big(h(x)\neq f(x)\big))$$

对于二分类问题，任何评估函数 l 的正确分类与错误分类的得分应该是固定的，即：

$$l(0,0) = l(1,1), l(0,1) = l(1,0)$$

$$l(0,0) + l(0,1) = l(1,1) + l(1,0)$$

因此，$l\big(h(x)=f(x)\big) + l\big(h(x)\neq f(x)\big)$为常数，设为 A，则：

$$\sum_f E_{ote}(\mathcal{L}_a|X,f) = \sum_{x\in\chi-X} p(x) \sum_h p(h|X,\mathcal{L}_a) \sum_f l\big(h(x),f(x)\big)$$

$$= \sum_{x\in\chi-X} p(x) \sum_h p(h|X,\mathcal{L}_a) \cdot 2^{|\chi|-1}$$

$$= \sum_{x\in\chi-X} p(x) \cdot 1 \cdot 2^{|\chi|-1}$$