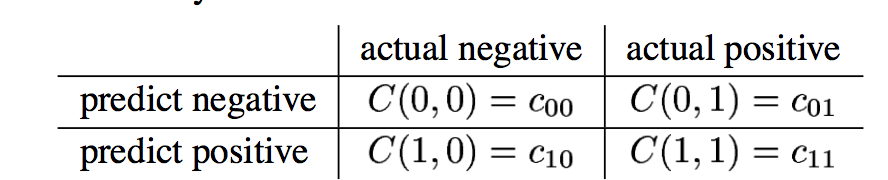
**类别：**理论类

**题目：The Foundations of Cost-Sensitive Learning**

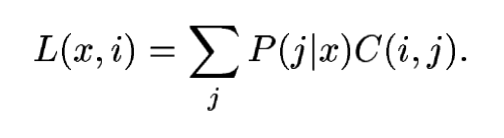
**作者：**Charles Elkan

**总述：**证明了一个定理，显示了如何改变训练集中的负面例子的比例，以便通过标准非成本敏感学习方法学习的分类器来做出最佳的成本敏感分类决策。但是，在具有不同错误分类成本的域中应用这些方法之一的推荐方法是从给定的训练集中学习分类器，然后使用分类器给出的概率估计明确地计算最优决策。

1. **基于代价矩阵进行决策**

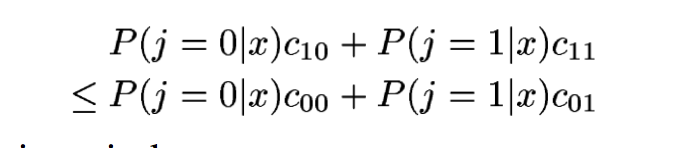
一个代价矩阵C是

对一个样本的最佳预测是以下函数的最小化



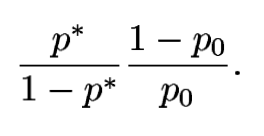
如果i=j则代表预测是预测是正确的，反之是错误的。

当达到以下条件取等号时，达到临界阈值，当p>p\*时，最佳预测取到，



1. **如何调整训练集产生出基于p\*的分类器？**

定理一：为了使目标概率阈值与给定的概率阈值相对应，训练集中的负面示例的数量应该乘以



在这些情况下，如果我们将稀有类称为正向类，则定理1直接说明如何更改常见示例的数量，而不丢弃或复制任何罕见的示例。

1. **基于两种分类器（朴素贝叶斯与决策树）来观察基准率的影响（没看懂）**

