**题目：**Classification Restricted Boltzmann Machine for comprehensible credit scoring model

**领域：**信用评分+信用卡还款

**核心创新点：**方法创新:分类受限玻尔兹曼(ClassRBM)

**论文结构与实现方法：**

总述：利用“分类受限玻尔兹曼”来建立信服力强的信用评分模型

1. 方法介绍——ClassRBM

优点及介绍：

1.ClassRBM有三层，其中有一个隐层，层内没有连接，只允许层与层之间有连接。

2.ClassRBM的一个重要优点是，对于足够数量的隐藏单位，该模型可以表示二进制向量上的任何分布，并且可以通过添加新的隐藏单元来改进其可能性，除非生成的分布已经等于训练分布

3.对于所考虑的信用还款问题，二进制输入的向量x表示描述申请人的特征，输出向量y保留在信用决策变量的后面。 因此，隐藏单元的向量允许在代表信用申请人的整个空间上近似分配。

缺点：

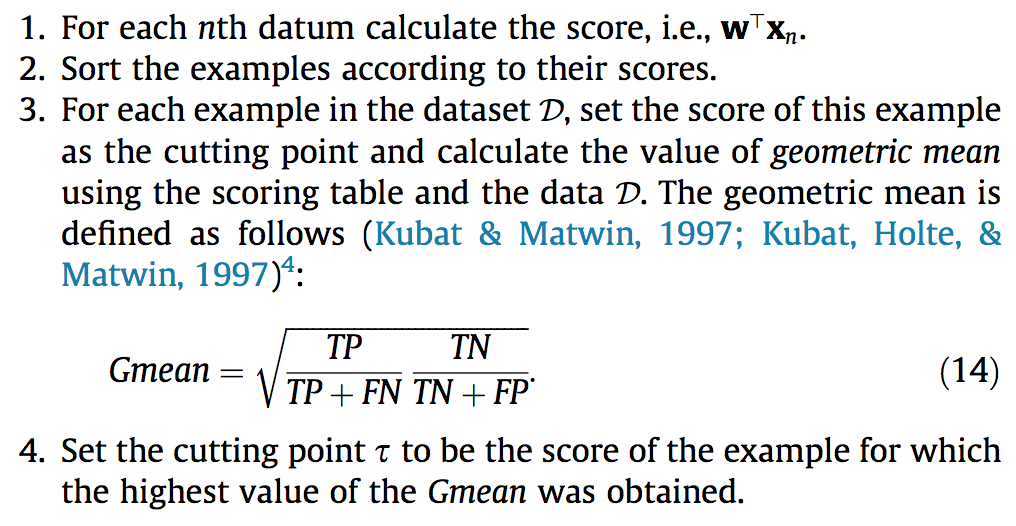
ClassRBM可以用作独立分类器来预测信用还款状态。 然而，由于它是难以解释的“黑匣子”模型，所以不太可能被用作信用评分模型。

1. 信服力强的信用评分表

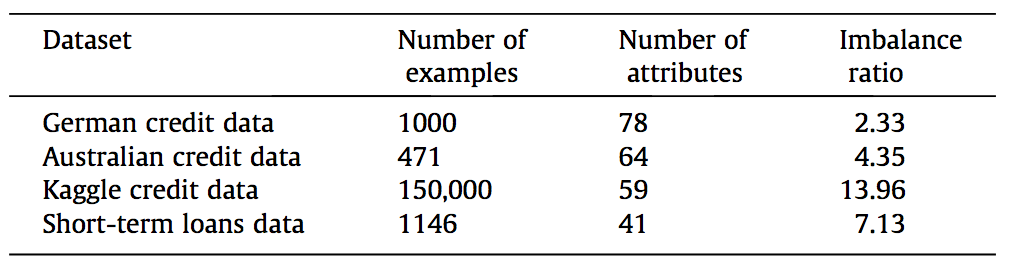
评分表是一个可理解的模型，因为它的简单解释。 每个第个输入可以获得等于wi的固定数量的点，并且临界点对应于在作出信用消费者的积极决定的每个点上的最小点数。 此外，如果拒绝给申请人的信用，则可以通过权重和切入点容易地解释决定的原因。

1. 确定临界点

切割点的确定是为了区分授信申请人是必不可少的。 建立阈值的方式可能受到训练数据集的不足之处的影响，例如不平衡的数据。 通常，偿还信贷的申请人人数远远高于失败人数。 任何不包括这个问题的学习方法都会产生有偏见的估计。 因此，对于确定的重量，我们提出以下确定切割点的步骤：



**实验数据：**四个高度不平衡数据集



**优秀之处：**比较了ClassRBM与其他的不同算法，数据集多样，并提供了其中一个数据集用该算法不足的地方

**改进之处：**

1.对输入变量进行二值化的需要可能导致丢失数据中的信息。 然而，在学习ClassRBM和评分表时，输入二值化是必要的。

2.构建评分表中必不可少的ClassRBM由相当大的参数组成，这可能导致学习过程中模型过度配套。 因此，应用某种正则化技术很重要。

3.可以结合代价矩阵建立评分表