A case-based reasoning model that uses preference theory functions

for credit scoring

本文主要介绍了 a case-based reasoning (CBR) model，输入为案例的信息，输出为正或反，代表在此案例中能否同意客户的贷款等请求。

主要使用方法：

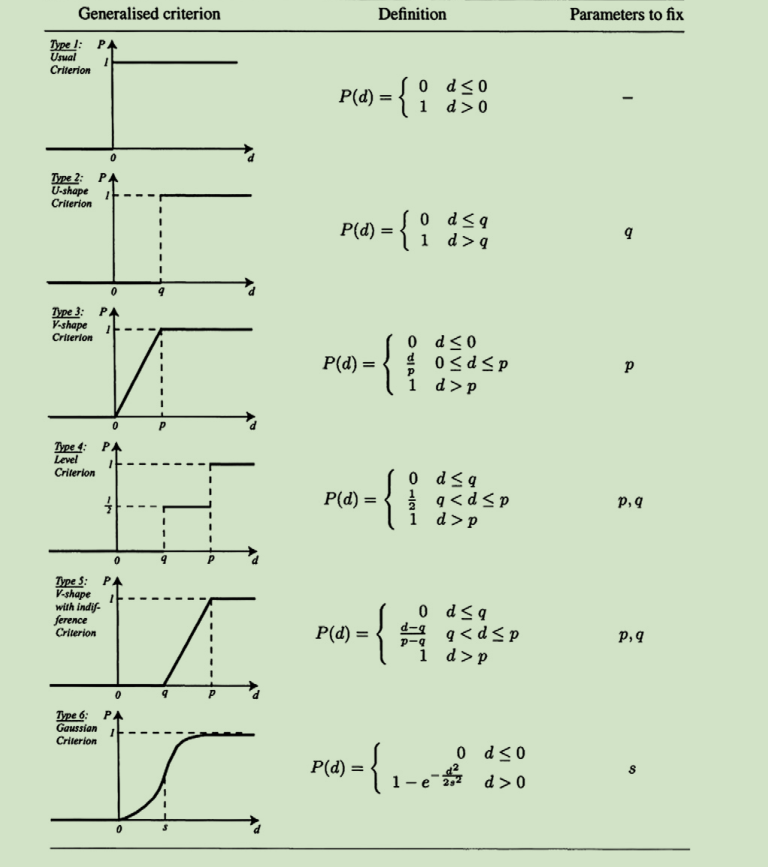
KNN，偏好函数（preference functions），遗传算法

论文主要内容：

首先介绍了使用KNN进行分类的方法，将之前的案例储存起来，输入新的案例时，根据KNN算法，找到最相似的K案例中所属的类较多的一类的类别作为输出。在衡量案例之间的相似度的的过程中，不同的feature之间的weight是不同的，需要将weight当做待定参数由遗传算法习得。同时，在衡量距离的过程中，对于两个案例间的同一个feature要经过preference function的转化，设feature的距离（此处一般是欧几里得范数计算出的距离）为d，则把d当做函数的自变量，将转化后的值参与两个案例之间的相似度的判定，偏好的函数的参数同样也是由遗传算法学习得到，文中采用的是第五种偏好函数。

此种方法的好处是缓解了传统的CBR中的过泛化（over-generalization）的问题，从而提升效果。

常用的偏好的函数为：



论文可取之处：

对几种方法进行了结合，改善了过泛化（over-generalization）的问题，在一定的程度上提升了系统的表现。此论文发表于2012年，内容较为初级和简单。