一种基于邻域粒的 LVQ 聚类算法

一.成果完成人及所在单位： 江海亮，计算机与信息工程学院 指导教师：姜红波

二.成果简介：

学习矢量量化(LVQ)是一种信息监督、结构简单、功能强大的聚类方法。LVQ 假设数据样本有标签，学习过程使用标签来辅助聚类。但 LVQ 对初始值敏感，导致聚类效果较差。为了克服这些缺点，本研究提出了一种采用邻域粒技术和 LVQ 结合的粒 LVQ 聚类算法。首先对数据集样本的某些特征进行邻域粒化，形成邻域粒向量。然后定义相邻粒向量的大小和运算，给出粒向量之间的相对距离和绝对距离。最后，证明这些粒距离度量，并设计出粒 LVQ 聚类算法。

三. 主要研究内容及创新点：

LVQ 算法属于原型聚类，原型聚类通常情况下先对原型进行初始化，然后对原型进行迭代更新求解。因此 LVQ 算法试图找到一组原型向量来描述数据，但 LVQ 算法训练前先假设样本有类别标签，训练过程使用类别标签来辅助聚类。原有 LVQ 算法对初始值较为敏感，甚至聚类不成功。而粒子含有样本空间的全局信息，我们将粒计算引入 LVQ 聚类，可以降低对于初始值的敏感度。本研究使用邻域粒化方法对 LVQ 聚类算法进行改进，在数据集中构建粒子和粒向量，并定义邻域粒向量的大小和运算规则，提出对邻域粒向量进行距离计算的相对距离和绝对距离度量，从而设计出邻域粒 LVQ 聚类算法（见图 1）。在多个 UCI 数据集上进行聚类实验，实验结果表明（见图 2），粒 LVQ 聚类算法能够成功对样本进行聚类，在合适的邻域参数和学习率参数情况下，能取得较好的结果。与传统 LVQ 聚类算法相比，邻域粒 LVQ 聚类算法可以得到更好的聚类结果。

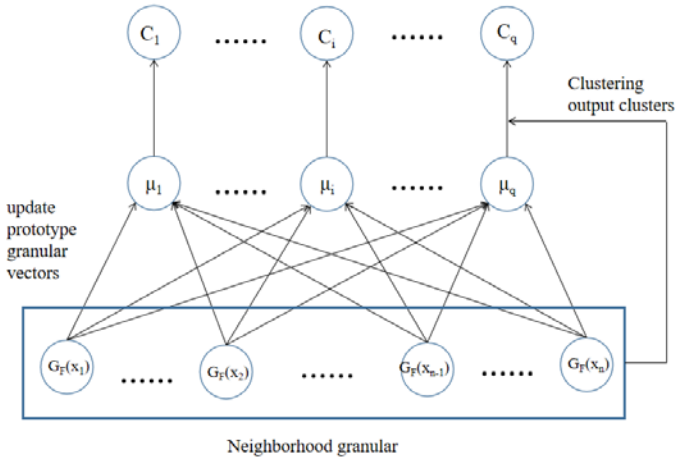


图 1 粒 LVQ 聚类算法流程图

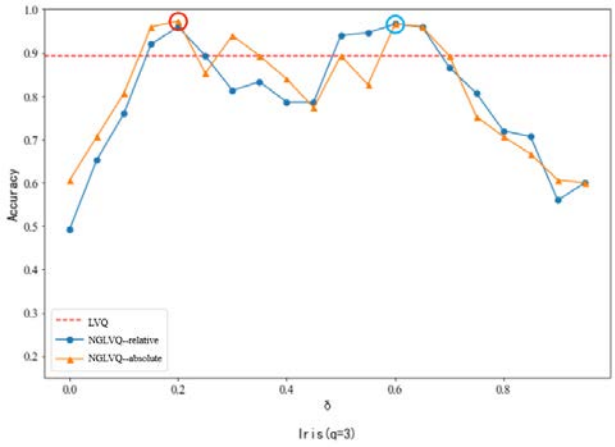


图 2 粒 LVQ 聚类算法精度示例图

四. 主要研究成果及应用情况

在《Journal of Intelligent & Fuzzy Systems》期刊发表论文 1 篇。

联系人：江海亮， 电 话：13255065790 Email:1348284274@qq.com