

文章编号: 1001-0920(2005)11-1229-06

分裂-合并竞争学习的研究

安成万^{1,2}, 张永谦¹, 谭民¹

(1. 中国科学院自动化研究所, 北京 100080; 2. 国家电网公司山西分公司, 太原 030001)

摘要: 针对竞争学习在给定的输出节点数目少于实际类数目时的学习结果会在几类数据之间振荡的问题, 提出了MPTOC策略以及基于此策略的分裂-合并竞争学习算法。在假设数据集中的数据对其相应节点产生大小等于二者距离“吸引力”的基础上, 算法通过计算网络中获胜节点在不同方向的“吸引力合力”分布, 间接描述该节点附近数据的分布情况; 采用高维空间模糊熵的方法确定该节点主要的“合力”方向, 并将该节点在这几个方向上进行分裂-合并学习, 从而实现MPTOC策略。通过对二维随机分布数据的实验结果验证了所提出算法的正确性和有效性。

关键词: 竞争学习; 分裂-合并竞争学习; MPTOC; 模糊熵

中图分类号: TP183 **文献标识码:** A

Research on Splitting-merging Competitive Learning

AN Cheng-wan^{1,2}, ZHANG Yong-qian¹, TAN Min¹

(1. Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China; 2. Shanxi Electrical Power Company, SG, Taiyuan 030001, China. Correspondent: AN Cheng-wan, E-mail: acw74s@sohu.com)

Abstract: An MPTOC strategy is presented, which guarantees each cluster has an output unit at least during learning. The concept of “attractive force” in mechanics is adopted to describe the relation between a unit and its corresponding data, which equal to their Euclid distance. Based on MPTOC splitting-merging competitive learning, the distribution of data around their winning unit is estimated indirectly through computing the unit attractive forces. And the unit splitting directions are assigned through method of fuzzy entropy in high dimension space. Then the unit is split and learned along these directions. To avoid over-segmentation of input dataset, the results of splitting-learning are merged with the help of their means and variances. Experiments in 2D space validate the proposed algorithm.

Key words: Competitive learning; Splitting-merging competitive learning (SMCL); MPTOC; Fuzzy entropy

1 引言

数据聚类分析的目的在于确定样本数据集内部的结构关系, 主要包括确定数据集的类数目, 各类数据的类中心以及它们在样本空间内的分布情况³。部分内容: 通常, 聚类分析是一个无监督学习过程, 传统的聚类算法大致分为以下2类: 1) 基于参数统计特性和概率分布的模型识别方法; 2) 基于向量量化的方法, 如自组织映射(SOM), 自适应共振理论(ART)系列和模糊逻辑等^[1-3]。作为单层自组织映射网络, 在已知类数目的前提下, 竞争学习能够较好

地解决上述第2个问题。从神经网络的观点出发, 数据集的类数目对应于竞争学习网络的输出节点数目, 类中心为各节点的权值。对于类数目已知的数据集, 竞争学习聚类就是通过迭代使输出节点的权值收敛到数据集的类中心^[4,5]。传统竞争学习网络的输出节点数目在学习过程中是固定的。若给定的节点数目少于实际类数目时, 学习结果会在几类数据之间振荡, 不能获得真正的类中心^[2,3,6]。因此, 对于类数目未知的数据集, 传统竞争学习仍有较大的局限。

针对此问题, 本文在传统竞争学习的基础上, 提

收稿日期: 2004-10-27; 修回日期: 2005-03-10

基金项目: 国家863计划项目(2F03H03, 2F03H06)

作者简介: 安成万(1974—), 男, 山西阳泉人, 博士生, 从事移动机器人目标识别、视觉导航等研究; 谭民(1962—), 男, 山西运城人, 研究员, 博士生导师, 从事机器人控制、多机器人系统、复杂系统理论与方法等研究

出了MPTOC (Multiprototypes-take-one-cluster) 策略, 保证每类数据在学习过程中至少有一个节点与之对应。通过借用力学中“吸引力”的概念描述节点与其对应数据在样本空间内的相对分布情况, 提出了分裂-合并竞争学习算法。该算法较好地解决了聚类分析的前两个问题, 实现了数据集类数目及类中心的自动确定, 对类数目未知的数据集有较好的聚类结果。

2 传统竞争学习

2.1 竞争学习的模型

设数据集 $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in R^d$ 为 d 维向量。对于包含 k 个输出节点的单层自组织映射网络, 每个输出节点记为 μ_i , 权值记为 $\omega_i \in R^d, i \in [1, k]$ 。那么标准竞争学习算法包含以下 2 步^[4]:

1) 从数据集 X 中任意选取一个输入数据 x , 对于每个输出节点 μ_i 有

$$\mu_i = \begin{cases} 1, & \text{如果 } i = c, \text{ 使得} \\ & x - \omega_i = \min_j \|x - \omega_j\|^2; \\ 0, & \text{其他} \end{cases} \quad (1)$$

2) 更新输出节点的权值 $\Delta\omega_i = \alpha\mu_i(x - \omega_i)$, 即

$$\Delta\omega_i = \begin{cases} \alpha(x - \omega_i), & \mu_i = 1; \\ 0, & \text{其他} \end{cases} \quad (2)$$

式中: $\mu_i = 1$ 的节点称为获胜节点, α 为学习速率。为保证算法收敛, 通常 α 取比较小的值。目前, 研究人员已提出许多基于标准竞争学习的学习算法, 主要区别在于学习规则和类数目的选取方式两方面。

2.2 竞争学习的学习规则

竞争学习的学习规则分为WTA (Winner-take-all) 与WTM (Winner-take-more) 两种, 分别称为“硬”竞争学习 (Hard competitive learning) 与“软”竞争学习 (Soft competitive learning)^[2,3]。在WTA规则下, 数据集中的任意数据的获胜节点只有一个, 如标准竞争学习和 k 均值算法等。WTA 规则存在一个重要问题: 对于不同初始权值的节点, 学习过程中可能存在死节点 (Dead-nodes) 或不充分利用 (Under-utilization) 的现象。为此, 研究人员提出了WTM 规则, 它放宽了WTA 规则的限制, 数据的获胜节点数目从一个变为多个, 如SCS (Soft competition scheme)^[7], NGN (Neural gas network)^[8,9] 和 FCL (Fuzzy competitive learning)^[10,11] 等算法。WTM 规则削弱了学习对节点初始权值的依赖, 但造成数据的获胜节点不确定, 使得部分节点可能偏离其实际的类数据。频率敏感竞争学习 (FSCL) 与模糊频率敏感竞争学习 (FFSCL) 引入了节点的获胜频率, 使获胜频率较低

的节点增加了获胜的机会, 在某些情况下可以显著地改善这个问题。但是, FSCL 和 FFSCL 采用非欧氏距离确定获胜节点, 可能导致一些节点收敛到同一类数据中。针对该问题, 文献[5] 提出了惩罚对手的竞争学习 (RPCL): 对于数据集中的每个数据, 不仅其获胜节点以一定的学习速率向该数据“靠近”, 而且其竞争对手 (第 2 获胜者) 以更小的学习速率 (惩罚速率) 被“推离”该数据。通常, 获胜节点的竞争对手会被“推”到远离该类数据的位置, 减小了对学习过程的干扰。然而, RPCL 对惩罚速率比较敏感, 如果惩罚速率选择不当, 则对某些节点可能造成过度惩罚或惩罚不充分^[7]。

2.3 类数目的选取

竞争学习的另外一个关键问题是数据集类数目的选取, 也就是网络输出节点个数的选取。对于上述竞争学习算法, 类数目需要在学习前正确指定, 否则聚类效果会很差。尽管 RPCL 对于较大的初始类数目是有效的, 但它不能解决初始类数目小于实际类数目的情形。

对于数据集类数目的选取, 一个可行的办法是在竞争学习中引入网络节点自生成机制 (Self-creating mechanism), 即在学习过程中网络可以根据数据集自动地增加新的节点^[6]。如 GCS (Growing cell structure), GNG (Growing neural gas), SSCL (Self-splitting competitive learning), SSMCL (Self-splitting and merging competitive learning) 和 BCL (Branching competitive learning) 等算法^[1-3,6-9]。GCS 与 GNG 缺乏节点增加与停止增加的有效准则, 而且要求初始类数目至少为 2, 不适合只包含一类数据的数据集。BCL 认为, 当某个节点为几类不同数据的获胜者时, 其学习轨迹会在这几类数据方向之间振荡。通过判断节点学习轨迹在样本空间中的变化, BCL 对包含几类数据的节点进行分裂来自动调整网络的规模, 对各类数据的数目基本相近的数据集 BCL 是有效的。但当类间数据的数目相差较大时, 节点的学习轨迹不足以充分反映节点附近数据的分布。SSCL 和 SSMCL 每次仅采用一个初始类中心来确定数据集中的一类数据, 通过多次迭代实现多类数据的聚类。为防止对数据集的过度分割, SSMCL 根据 SSCL 分类结果的均值和方差合并相应的类。SSCL 和 SSMCL 需要对数据集进行多次聚类, 使得运算量显著增加, 而且 SSCL 算法对控制迭代聚类过程的阈值比较敏感。

3 分裂-合并竞争学习

3.1 OPTMC, OPTOC 与 MPTOC

在数据聚类分析中, 当指定的节点 (类中心) 数

目少于实际数目时, 至少会有两类数据拥有同一节点, 该现象称为 OPTMC (One - prototype - take - multiclusters)^[1]. 这种情况会使得该节点在两类或几类数据之间振荡. 图 1(a) 给出了一个初始节点 p 对数据集 $X = \{X_1, X_2, X_3\}$ 的学习轨迹. 当指定的节点数目与实际类数目相等时, 节点的学习轨迹还受其初始位置的影响, 仍然可能出现 OPTMC 现象. 对于 X , 假定 3 个节点分别为 p_1, p_2 和 p_3 , 图 1(b) 和图 1(c) 分别给出了两种不同初始位置情况下 3 节点的学习轨迹. 可以看出, 图 1(b) 中 p_3 正确地收敛到类 X_3 中, 而 p_1 没有参与学习, p_2 则仍出现了 OPTMC 现象. 在图 1(c) 中实现了每个节点对应一类数据, 这是聚类分析最理想的结果, 称为 OPTOC (One-prototype-take-one-cluster)^[1]. 但是, 在不知道数据集实际类数目和数据分布情况的条件下, OPTOC 的实现有较大的难度. 基于此, 本文提出了 MPTOC (Multi prototypes-take-one-cluster) 策略: 在学习过程中, 一类数据可能有多个节点与之对应; 但保证每类数据至少要有一个节点与之对应. 图 1(d) 给出了对数据集 X 采用 MPTOC 策略的节点学习轨迹. 然而, 采用 MPTOC 策略可能会使数据集被过度分割, 使原本属于一类的数据分为两类或几类. 因此, 应该对采用 MPTOC 策略获得的聚类结果进行合并, 从而实现 OPTOC. 下面给出基于此策略的分裂 - 合并竞争学习算法.

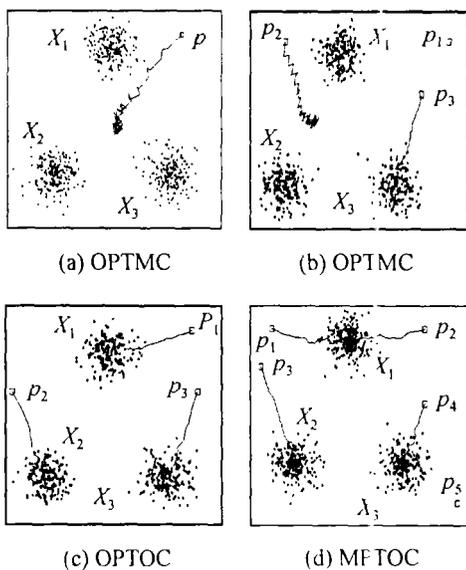


图 1 二维数据聚类中节点的几种学习轨迹示意图

3.2 样本空间内节点的“合力”分布

分裂 - 合并竞争学习通过计算获胜节点与其附近数据的相互位置关系, 自动调整网络的输出节点数目. 为了方便地进行上述位置关系的分析、计算, 本文引入力学中“吸引力”的概念.

定义 1 令 $\vec{f}_i = (f_{i1}, f_{i2}, \dots, f_{id}), \vec{f}_j = (f_{j1}, f_{j2}, \dots, f_{jd})$ 分别为两个 d 维向量, f_i 与 f_j 为二者在 d 维空间中对应的点. 假设二者之间存在“吸引力”, 称向量 \vec{f}_{ij} 为点 f_i 对点 f_j 的“吸引力”, 有

$$\vec{f}_{ij} = \vec{f}_i - \vec{f}_j = (f_{i1} - f_{j1}, f_{i2} - f_{j2}, \dots, f_{id} - f_{jd}). \quad (3)$$

定义 2 对于 d 维向量 $\vec{f} = (f_1, f_2, \dots, f_d)$, 称 $d + 1$ 维向量 $\vec{p} = (f, \theta_1, \theta_2, \dots, \theta_d)$ 为 \vec{f} 的极坐标形式向量, 有

$$f = \vec{f} \cdot \vec{e}_1, \theta = \cos^{-1}(f_i/f), \theta \in [0, \pi] \quad (4)$$

其中: 分量 f 为 \vec{f} 的 2-范数, 即 \vec{f} 的模; 其余 d 个分量 θ 是 \vec{f} 与 d 维空间各坐标轴的夹角. 此处, \vec{f} 的极坐标形式向量 \vec{p} 是 2 维空间极坐标在 d 维空间的推广.

设数据集 $X = \{x_1, x_2, \dots, x_n\}, x_i \in R^d$, 竞争学习网络包含 k 个输出节点, 每个输出节点记为 μ_i , 权值记为 $\omega \in R^d, i \in [1, k]$. 在每次迭代学习过程中, 根据 WTA 规则, 数据集 X 分为若干个数据子集, 每个数据子集对应不同的输出节点. 设数据子集 $Z_i = \{\vec{z}_{i1}, \vec{z}_{i2}, \dots, \vec{z}_{im}\}$ 包含 m 个数据, 其对应的输出节点为 μ_i , 即

$$\vec{z}_{il} - \vec{\omega}_i = \min_{j \in [1, k]} \|\vec{z}_{il} - \vec{\omega}_j\|, \vec{z}_{il} \in X, l \in [1, m], j \in [1, k] \quad (5)$$

设子集 Z_i 中的每个点对节点 μ_i 存在“吸引力”. 根据定义 1, Z_i 中所有点对节点 μ_i 的“吸引力”子集为

$$F_i = \{\vec{f}_{i1}, \vec{f}_{i2}, \dots, \vec{f}_{im}\}; f_{il} = \vec{z}_{il} - \vec{\omega}_i, l \in [1, m] \quad (6)$$

可以看出, 子集 F_i 为节点 μ_i 所受到的所有“吸引力”的集合. 为了计算节点 μ_i 在 d 维空间中各方向上的“吸引力合力”, 将“吸引力”子集 F_i 中的各元素转换为其极坐标形式 $J_i = \{\vec{p}_{i1}, \vec{p}_{i2}, \dots, \vec{p}_{im}\}$. 根据力学中关于质点受力分析的方法, 计算节点 μ_i 在 d 维空间中不同方向上的“合力”子集 $H_i = \{\vec{h}_{i1}, \vec{h}_{i2}, \dots, \vec{h}_{im}\}$, 其中 $M = m$ 为节点 μ_i 所有的受力方向. 此处, 令向量 $\vec{p}_{i1}, \vec{p}_{i2}, \dots, \vec{p}_{im}$ 的各分量角 θ 取整数, 则

$$\vec{h}_{ij} = \{f, \theta_1, \theta_2, \dots, \theta_d\}, f = \vec{p}_{il}(1),$$

如果

$$\vec{p}_{il}(w) = \vec{h}_{ij}(w), w \in [2, d + 1], l \in [1, m], j \in [1, M] \quad (7)$$

其中 $\vec{p}_{il}(w)$ 和 $\vec{h}_{ij}(w)$ 分别表示向量 \vec{p}_{il} 和 \vec{h}_{ij} 的第 w 个分量.

如果数据集 X 的数据分为若干类, 则这些数据在 d 维空间内会形成若干“团”点的集合, 并且各

“团”点在其类中心附近具有较高的分布密度 从分裂 - 合并竞争学习的观点出发, 上述情况可表述为: 对于数据集 X , 根据 WTA 规则可确定其确定数据子集 Z_i 及对应的输出节点 μ_i , 并且节点 μ_i 的“合力”子集 H_i 在数据子集 Z_i 的方向上具有较大的“合力” 因此, 一定程度上节点的“合力”子集可以反映该节点与其对应数据子集在 d 维空间中的相对分布情况 图 2 为 2 维数据的示意图 图 2(a) 中节点 p 附近有 3 类不同的数据 X_1, X_2, X_3 图 2(b) 中的各点表示各类数据对节点的“吸引力”相对于 p 的大小与方向; d_1, d_2, d_3 分别表示节点主要的 3 个“受力”方向 可以看出, 节点 p 的“合力”子集元素主要集中在这 3 类数据的方向上

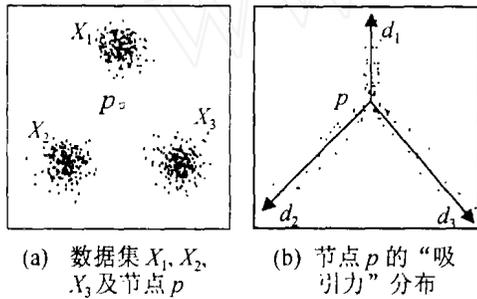


图 2 节点“吸引力”分布示意图

3.3 分裂准则

通过 3.2 节的分析, 可得出如下结论: 如果节点 μ_i 在某些方向上具有较大的“合力”, 则表明在这些方向上有不同类的数据, 或者是具有不同局部分布密度的某类数据与之对应 为了保证每类数据至少有一个节点与之对应, 该节点应该在这几个主要方向上进行分裂, 以生成其数据子集中不同部分数据的获胜节点 因此, 节点的分裂问题就变为其主要“合力”方向的提取

定义 3 对于 $d + 1$ 维向量 \bar{p}_i 与 \bar{p}_j , 称 $\theta(\bar{p}_i, \bar{p}_j)$ 为 \bar{p}_i 与 \bar{p}_j 的夹角, 有

$$\theta(\bar{p}_i, \bar{p}_j) = \cos^{-1} \frac{\bar{p}_i \cdot \bar{p}_j}{\|\bar{p}_i\| \cdot \|\bar{p}_j\|} \quad (8)$$

其中 $\bar{p}_i \cdot \bar{p}_j$ 表示 \bar{p}_i 与 \bar{p}_j 的内积

定义 4 在 $d + 1$ 维空间中, 若满足条件 $|\theta(\bar{y}, \bar{y}')| < \theta_{\text{threshold}}$ (9) 的所有向量 \bar{y}' 有 $\bar{y}'(1) > \bar{y}(1)$, 则称 \bar{y} 为给定阈值 $\theta_{\text{threshold}}$ 条件下的局部极大值点, 简称局部极大值点 其中 $\theta_{\text{threshold}}$ 为给定阈值, 是为了求取“吸引力”局部极大值, 主要针对高维空间中“吸引力”的一些局部“振荡”现象 其主要作用是模糊熵方法确定“吸引力”的模糊区域, 其值越大, 确定的模糊区域越宽, 对“吸引力”子集的平滑作用越大, 但可能会

造成几类样本对应一个节点的情况 反之, 则平滑作用较小, 模糊熵方法对“吸引力”子集造成的“失真”程度较小, 但同时带来计算量增加的问题 通常, $\theta_{\text{threshold}}$ 应采用比较小的值

本文节点 μ_i 主要“合力”方向的提取采用基于高维空间模糊熵的方法 设节点 μ_i 的“合力”子集 $H_i = \{\bar{h}_{i1}, \bar{h}_{i2}, \dots, \bar{h}_{iM}\}$, 则基于高维空间模糊熵的主要“合力”方向提取分为以下 3 步:

Step 1: 最大模糊区域 $\Delta\theta$ 的确定 设 $S_i = \{\bar{s}_1, \bar{s}_2, \dots, \bar{s}_N\}$ 为子集 H_i 在定义 4 下的局部极大值点子集, 根据定义 3, 计算各极大值点与其他极大值点的最小角度差集

$$T = \{t_1, t_2, \dots, t_N\},$$

$$t_i = \min(\theta(\bar{s}_i, \bar{s}_j)),$$

$$i \in [1, N], j \in [1, N], i \neq j, \quad (10)$$

则最大模糊区域 $\Delta\theta = \max(t_i), t_i \in T$.

Step 2: 以 $\Delta\theta$ 为固定的模糊区域宽度, 在不同渡越点^[12] 情况下计算 S_i 的模糊熵 $E(S_i) = \{E(\bar{s}_1), E(\bar{s}_2), \dots, E(\bar{s}_N)\}$.

Step 3: 根据模糊熵原理, 在 $E(S_i)$ 中确定主要“合力”方向

Step 2 和 Step 3 阶段的计算与二维空间的模糊熵方法相仿, 详见文献[12~14] 设节点 μ_i 经上述阶段确定的分裂向量 $DS = \{\bar{d}_1, \bar{d}_2, \dots, \bar{d}_N\}, \bar{d}_i \in R^d$, 则该节点分裂为 N 个新的节点 $\mu_{i1}, \mu_{i2}, \dots, \mu_{iN}$. 各节点的权值确定如下:

$$\bar{\omega}_j = \bar{\omega} + \alpha^* \bar{d}_j, j \in [1, N], \quad (11)$$

其中 α 为网络的学习速率 从式(11)可以看出, 当节点只有一个分裂向量时, 它仅在其最大的“合力”方向上学习, 这与标准竞争学习算法基本相同

对于任意数据集, 经过充分地学习初始类中心可能分裂为整个数据集中的点, 但此程度的分裂结果对聚类没有实际意义 因此, 需要控制节点的迭代分裂 在不知道数据集的类数目及其在样本空间中分布的情况下, 本文采用节点与其对应的数据子集的特征距离控制迭代分裂 设经过一定的学习, 数据集 X 中对应于节点 μ_i 的数据子集与它的最大、最小距离为 l_1, l_2 , X 中数据之间的最小距离与最大分裂距离阈值分别为 $l_{\text{min}}, l_{\text{threshold}}$ 停止分裂的准则为:

- 1) 若 $l_1 > l_{\text{threshold}}$ 且 $l_2 > l_{\text{min}}$, 则该节点停止分裂;
- 2) 其他情况, 则继续分裂

若所有节点都满足该准则, 则分裂学习结束 $l_{\text{min}}, l_{\text{threshold}}$ 的作用是为了控制分裂学习的终止与否 取值越大, 分裂的节点数目越少; 反之则多 如

果产生的分裂节点少, 则可能产生几类分为一类的情况 但是, 分裂节点多的情形可用后续的合并学习加以校正 因此, $l_{min}, l_{threshold}$ 应取比较大的值 通常, 可以采用输入样本的一些统计值, 如方差、样本与均值的的最小距离等, 或采用分类问题所要求的具体分类精度等

3.4 合并准则

经上述分裂获得的聚类结果, 可能存在过度分割的情况, 因此需要对分裂结果进行合并 本文采用文献[3]中提出的合并准则, 设经过分裂, 数据集 X 分为 K 类, $X = X_1 \ X_2 \ \dots \ X_K$, 类 X_i 数据的均值与方差分别为 \bar{c}_i, δ_i 对于任意给定的两类 X_i 和 X_j , 若满足

$$\bar{c}_i - \bar{c}_j \leq \frac{1}{2} (\delta_i + \delta_j), \quad (12)$$

则二者合并 对于所有的分裂结果, 运用上述条件将所有相互相邻的类合并, 并根据合并结果计算各类的类中心, 获得最后的聚类结果

4 仿真结果

本文采用 3 类 2D 随机分布数据进行分裂-竞争学习算法的仿真实验 输入数据及节点初始位置如图 3(a) 所示, 输入数据集 $X = \{X_1, X_2, X_3\}$, 初始节点为 3 个, 其初始位置任意给定 其中 $l_{threshold}$ 取输入数据的标准方差 图 3(b) 为节点的学习轨迹及分裂结果, 图 3(c) 为分裂节点合并结果, 最后的聚类结果如图 3(d) 所示

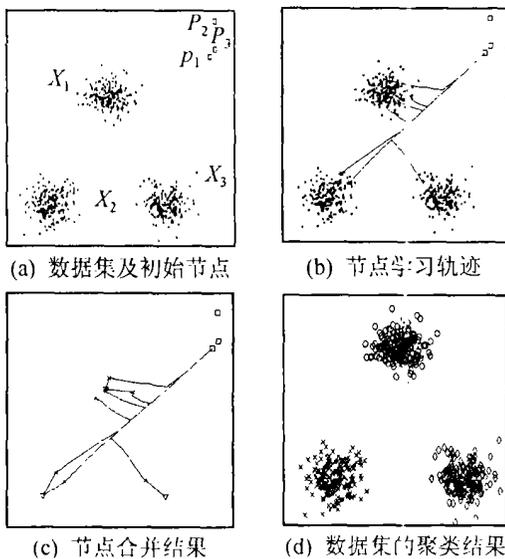


图 3 采用 3 类 2D 随机分布数据的仿真实验

从图 4 可以看出, 在任意给定初始位置的 3 个节点 P_1, P_2, P_3 中, 节点 P_1 与 3 类数据集的距离最小, 因此只有 P_1 成为 3 类数据的获胜节点, 这与只给定一个初始节点情况相同 随着不断地学习, P_1 根据其附近数据的分布情况最终分裂为 7 个节点,

且每类数据至少有一个节点实现了 MPTOC 策略 但是, 对于 X_1 及 X_2 类数据, 对应的节点多于 1 个, 产生了过度分割现象 采用本文给出的合并准则, 这两类数据的 6 个节点分别合并为二者各自的类中心 可以看出, 3 类的类中心基本都保持在各类数据的中心位置附近 图 4 分别给了 3 个节点在不同初始位置的分裂-学习过程及结果

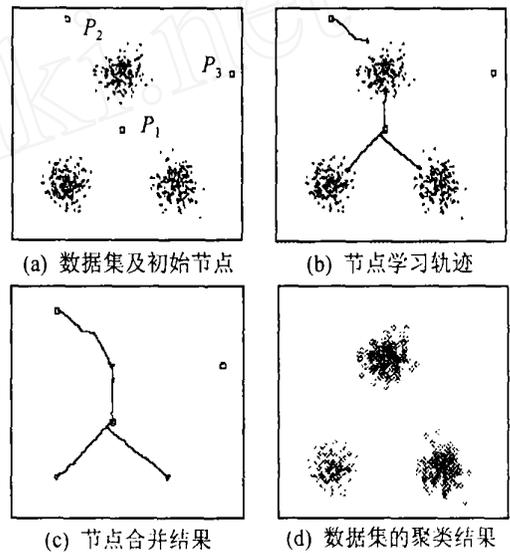


图 4 本文方法的仿真实验

从上述两个仿真实例可以看出, 在节点不同初始数目和位置的情况下, SMCL 算法能对输入样本数据实现较好的分类, 通过节点的分裂使得每类样本数据都能有各自的类中心 特别是在第 2 个示例中, 即使节点 p_1 的初始位置位于 3 类样本数据之间, 但在 3 类数据 3 个不同方向“吸引力”的作用下, 从分裂开始, p_1 就在这 3 个方向分裂为 3 个不同的节点, 并最终形成各类的类中心

对于分裂-竞争学习, 节点的分裂学习受 $l_{threshold}$ 的影响 从图 3(c) 可以看出, 对于 X_2 和 X_3 类数据, 它们的分裂结果并没有进入类数据中, 而只是位于其边缘附近 如果 $l_{threshold}$ 取较小的值, 则分裂结果可以进入对应类数据中, 但会造成分裂节点的数目过多; 如果取得太大, 又容易使类间距离较小的相邻类拥有同一节点, 造成错分 对于 $l_{threshold}$ 的选取, 可以采用输入样本的一些统计值或不同分类问题相应的分类精度等具体要求

5 结 论

本文借用力学中“吸引力”的概念, 通过计算竞争学习网络中获胜节点在不同方向上的“合力”分布, 间接反映该节点附近数据的分布情况 基于本文提出的 MPTOC 策略, 并采用高维空间模糊熵的方法, 提出了分裂-合并竞争学习算法 该算法无需事

先指定网络输出节点的数目,且不受节点初始权值的干扰,能够自适应地确定数据集的类数目和各类类中心。在2维空间内的实验结果表明了该算法的正确性和有效性。下一阶段将从节点对应样本数据的分布密度出发,对停止分裂准则以及分类结果的合并准则作进一步的研究,以进一步提高该算法的鲁棒性。

参考文献(References)

- [1] Raffaella Mattone. The Growing Neural Map: An On-line Competitive Clustering Algorithm [A]. *Proc of the 2002 IEEE Int Conf on Robotics and Automation* [C]. Washington DC: IEEE, 2002: 3888-3893
- [2] Zhang Y J, Liu Z Q. Self-splitting Competitive Learning: A New On-line Clustering Paradigm [J]. *IEEE Trans on Neural Network*, 2002, 13(2): 369-380
- [3] Shuanhu Wu, Alan Wee-Chung Liew, Hong Yan, et al. Cluster Analysis of Gene Expression Data Based on Self-splitting and Merging Competitive Learning [J]. *IEEE Trans on Information Technology in Biomedicine*, 2004, 8(1): 5-15
- [4] Teuvo Kohonen. The Self-organizing Map [A]. *Proc of the IEEE* 78(9) [C]. Tokyo: IEEE, 1990: 1464-1480
- [5] Lei Xu, Adam Krzyzak, Erkki Oja. Rival Penalized Competitive Learning for Clustering Analysis, RBF Net and Curve Detection [J]. *IEEE Trans on Neural Network*, 1993, 4(4): 636-649
- [6] Huilin Xiong, Swamy M N S, Omair Ahmad. Branching Competitive Learning Network: A Novel Self-creating Model [J]. *IEEE Trans on Neural Network*, 2004, 15(2): 417-429
- [7] Liu Z Q, Glickman M, Zhang Y J. Soft-competitive Learning Paradigms [A]. *Soft Computing and Human-centered Machines* [C]. Berlin: Springer-Verlag, 2000: 131-161
- [8] Fritzke B. Growing Cell-structures — A Self-organizing Network for Unsupervised and Supervised Learning [J]. *Neural Networks*, 1994, 7(9): 1441-1460
- [9] Fritzke B. *A Growing Neural Gas Network Learns Topologies*, *Advances in Neural Information Processing Systems* [M]. Cambridge: MIT Press, 1995: 626-632
- [10] Hichem Frigui, Raghu Krishnapuram. A Robust Clustering Algorithm Based on Competitive Agglomeration and Soft Rejection of Outliers [A]. *Proc CVPR '96* [C]. San Francisco CA: IEEE, 1996: 550-555
- [11] Frigui H, Krishnapuram R. A Robust Competitive Clustering Algorithm with Applications in Computer Vision [J]. *IEEE Trans on Pattern Analysis Machine Intelligence*, 1999, 21(5): 450-465
- [12] Cheng H D, Che C H, Chiu H H, et al. Fuzzy Homogeneity Approach to Multilevel Thresholding [J]. *IEEE Trans on Image Processing*, 1998, 7(7): 1084-1088
- [13] Cheng H D, Li J. Fuzzy Homogeneity and Scale-space Approach to Color Image Segmentation [J]. *Pattern Recognition*, 2003, 36(7): 1545-1562
- [14] Cheng H D, Jiang X H, Wang J L. Color Image Segmentation Based on Homogram Thresholding and Region Merging [J]. *Pattern Recognition*, 2002, 35(2): 373-393

(上接第1228页)

- [9] 陶永华. *新型PID控制及其应用* [M]. 北京: 机械工业出版社, 2002: 160-169
(Tao Y H. *New PID Control and Its Application* [M]. Beijing: Mechanical Industry Publisher, 2002: 160-169)
- [10] 谭连生, 尹敏. 计算机高速互连网中一类基于速率的 PD 拥塞控制方法 [J]. *自动化学报*, 2003, 29(1): 54-61
(Tan L S, Yin M. A Rate-based PD Congestion Controller for High-speed Computer Communication Networks [J]. *Acta Automatica Sinica*, 2003, 29(1): 54-61)