

基于强化学习的 LVQ 聚类方法

Reinforcement Learning Based LVQ Clustering Approach

程小平 邱玉辉

(西南师范大学计算机与信息科学学院 重庆400715)

Abstract A reinforcement clustering framework which constitutes Bernoulli stochastic neural units is proposed in this paper. Reinforcement learning mechanism is introduced to LVQ clustering problems. Related algorithm LVQ-R is developed and its property is analyzed in detail. The authors conclude that reinforcement learning can be also introduced to other on-line competitive clustering methods. Experiments show that LVQ-R has better performance than original LVQ approach.

Keywords Reinforcement learning, Bernoulli stochastic unit, LVQ

1 引言

在数据挖掘、智能控制、模式识别中,数据聚类始终是一种基本需要。在缺乏对数据类属的先验知识的情况下,多采用无监督学习方法来训练分类器。

强化学习(reinforcement learning,也称增强学习,自励学习)是一种介于有监督学习和无监督学习之间的一种机器学习方法,它利用少量的评价信息改善系统行为。强化学习与动物学习理论,认知科学,学习自动机理论有密切的关系。近年来得到大量的研究并应用到包括数据挖掘在内的多个领域^[1]。

下文中首先提出一个通用的强化学习聚类方法框架,然后在此框架的基础上将强化学习机制引入常用的学习向量量化(learning vector quantization LVQ)算法,并分析强化 LVQ 方法的特点及本文基本方法的可推广性。

2 强化学习聚类方法的一般结构

数据聚类问题描述如下:

N 个输入数据 $x = (x_1, x_2, \dots, x_N)$, 每个数据是 m 维向量 $x_j = (x_{j1}, x_{j2}, \dots, x_{jm})^T \in R^m$ 需要将其指定为 L 个分类 $W = (w_1, w_2, \dots, w_L)$ 。每个分类的聚类中心由向量 $w_i = (w_{i1}, w_{i2}, \dots, w_{im}) (i=1, 2, \dots, L)$ 表示。

为了很好地体现强化学习的随机性质,我们采用最终输出 y_i 为 1 或 0 二值 Bernoulli 神经元^[2,3]构成的离散随机神经网络来构造强化学习结构,如图1所示。

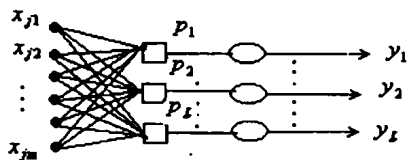


图1 二值 Bernoulli 神经元离散随机神经网络

图1中第一层表示 m 维输入数据;第二层是 s 函数神经元结点,但输出值被归算到 $(0, 1)$ 区间,其意义是概率值。第三层依概率决定 L 个分类的最终输出,即熟知的二值 Bernoulli 公式:

$$p(y=k) = p_i^k (1-p_i)^{1-k}; k=0, 1 \quad (1)$$

即对于 $i=1, 2, \dots, L$, 都有:

$$y_i = \begin{cases} 1; & \text{按概率 } p_i \\ 0; & \text{按概率 } (1-p_i) \end{cases} \quad (2)$$

强化学习聚类方法如下:

(1) 对随机选取的 $x \in X$ 计算到聚类中心的距离

$$s_i = d(x_j, w_i) \quad i=1, 2, \dots, L$$

(2) 计算第二层输出的概率值

$$p_i = h(s_i) = 2(1 - f(s_i)) \quad i=1, 2, \dots, L \quad (3)$$

其中 $f(x)$ 是 s 函数 $f(x) = \frac{1}{1+e^{-x}}$ 由于 $s_i \geq 0$, h 值被归算到 $(0, 1)$ 区间, p_i 随 s_i 单调减小, 即 x 与 w_i 的距离 s_i 越小, 第 i 类输出 $y_i=1$ 的概率越大。

(3) 按通用的随机梯度强化法^[3]更新权值

$$\Delta w_{ij} = a(r - b_{ij}) \frac{\partial g_i(y_i, p_i)}{\partial w_{ij}} \quad (4)$$

其中 a 为学习率, r 是按某种评价方法给出的强化量值, b_{ij} 为一常量, 称为强化基准, 对于我们图1这个神经网络, $g_i(y_i, p_i)$ 即式(1)的 Bernoulli 分布函数:

$$g_i(y_i, p_i) = p_i^{y_i} (1-p_i)^{1-y_i} \quad (5)$$

3 基于强化信号的 LVQ 聚类方法

对于基本 LVQ 方法, 权值更新仅在距离度量值最小的单元 i^* 进行。

$$\Delta w_{ij} = \begin{cases} a_i(x_j - w_{ij}); & \text{若 } i \text{ 为获胜单元 } i^* \\ 0; & i \neq i^* \end{cases} \quad (6)$$

为了使强化 LVQ 方法有基本 LVQ 竞争学习的特点, LVQ-R 的权值更新也仅在获胜单元(第二层输出 p_i 最大的单元)进行。

由于获胜单元最终输出是由 Bernoulli 分布按概率取 1 或 0, 因此, 强化信号与之一致。获胜单元(按概率)被激活 ($y_i=1$) 时, 给出奖励强化 $r_i > 0$, 而获胜单元(按概率)未被激活 ($y_i=0$) 时, 给出惩罚强化信号 $r_i < 0$, 而对其它单元给强化信号为 0, 使取权值不更新。

不失一般性, 令 $b_{ij}=0$, 则(4)式

$$\begin{aligned} \Delta w_{ij} &= ar \frac{\partial g_i(y_i, p_i)}{\partial w_{ij}} \\ &= ar \frac{\partial g_i(y_i, p_i)}{\partial p_i} \cdot \frac{\partial p_i}{\partial s_i} \cdot \frac{\partial s_i}{\partial w_{ij}} \end{aligned} \quad (7)$$

由(5)式可得:

$$\frac{\partial g_i(y_i, p_i)}{\partial p_i} = \frac{y_i - p_i}{p_i(1-p_i)} \quad (8)$$

由(3)式得

$$\frac{\partial p_i}{\partial s_i} = p_i \left(\frac{p_i}{2} - 1 \right) \quad (9)$$

在 s_i 为欧氏距离度量的情形,即

$$s_i = d(x, w_i) = \sum_{j=1}^L (x_j - w_{ij})^2 \quad (10)$$

$$\frac{\partial s_i}{\partial w_{ij}} = -2(x_j - w_{ij}) \quad (11)$$

将式(8)、(9)、(11)代入(7),整理,得出权值更新表达式为:

$$\Delta w_{ij} = a r_i (y_i - p_i) (x_j - w_{ij}) \cdot \frac{2 - p_i}{1 - p_i} \quad (12)$$

按强化信号奖励策略,选取了强化量

$$r_i = \begin{cases} \frac{1 - p_i}{2 - p_i} > 0, & i = i^* \text{ 且 } y_i = 1 \\ -\frac{1 - p_i}{2 - p_i} < 0, & i = i^* \text{ 且 } y_i = 0 \\ 0, & i \neq i^* \end{cases} \quad (13)$$

将 r_i 代入(12),容易得出:

$$\Delta w_{ij} = \begin{cases} a |y_i - p_i| (x_j - w_{ij}), & i = i^* \\ 0, & i \neq i^* \end{cases} \quad (14)$$

这样,我们就得到在强化学习框架下,LVQ 聚类算法 LVQ-R.

(1) 随机选取 $x \in X$, 并计算距离度量 $s_i = d(x, w_i)$.

(2) 按式(3)计算 p_i , 并选中获胜单元 i^* , $p_i^* = \text{Max}(p_i)$, $i = 1, 2, \dots, L$.

(3) 按 Bernoulli 分布,依概率确定获胜单元的 y_i 并按式(13)计算强化量.

(4) 按式(14)更新权向量.

4 强化 LVQ 方法的分析及强化学习方法的推广

强化学习系统显示出随机爬山法的特性^[3],性能度量是使强化量的数学期望极大化.

即 $\hat{R}(w, r_i) = E(r/w, x) = \text{Max}$, 权值更新的平均值指向 $E(r/w, x)$ 增加的方向,容易得出,在 Bernoulli 随机神经元的情形

$$R(x, w) = \sum_{j=1}^N (p_i^*(x_j) - (1 - p_i^*(x_j))) \quad (15)$$

其中 $p_i^*(x_j)$ 是对应于输入 x_j 的最大概率.

上式可写为:

$$R(x, w) = \sum_{j=1}^N p_i^*(x_j) - N \quad (16)$$

由于 N 为常数,上式的最大化即是

$$\sum_{j=1}^N p_i^*(x_j) \text{ 的最大化}$$

而 p_i 与距离度量成反比,因此,强化量的最大化对应于距离度量的极小化,这一点与基本 LVQ 方法是一致的.

事实上,可以将基本 LVQ 考虑成 LVQ-R 的特例.令 $\rho = \frac{1 - p_i}{2 - p_i} > 0$, 容易证明,只要选取

$$r_i = \begin{cases} \frac{1}{y_i - p_i} \rho, & i = i^* \text{ 且 } y_i = 1 \\ -\frac{1}{y_i - p_i} \rho, & i = i^* \text{ 且 } y_i = 0 \\ 0, & i \neq i^* \end{cases} \quad (17)$$

则按 LVQ-R 算法得出的权值更新与基本 LVQ 算法相同.

比较 LVQ-R 与 LVQ 的参数更新方程,差别仅在于 LVQ-R 中多一项 $|y_i - p_i|$. 由于 y_i 依概率取 1 或 0, 因此 $|y_i - p_i|$ 依输出值 y_i 不同而不同,在相同的条件下, LVQ-R 与基本 LVQ 的权值更新的强度就可能不同,这相当于在 LVQ 权值更新过程中引入随机噪声,使得 LVQ-R 比较不易陷入局部极小.

另外,现有的许多在线竞争学习聚类算法,大多数情况下距离度量可求导,且参数更新算法与 LVQ 有相似的形式.因此,可以适当地选取强化量 r_i , 使得这些算法可以在强化学习的框架下描述.这种描述,一方面为分析这些算法提供了一个新的视角;另一方面,由于引入了随机性及强化学习机理,将会减少陷入局部极小的机会,并有望改变这些算法的性能.下面列举两个常用的算法例子:

(1) 自组织特征映射^[5] (Self Organizing Feature Mapping) (SOFM)

距离度量为 $d(x, w_i) = (x - w_i)^2$, 取

$$r_i = \begin{cases} \rho h_{\sigma}(i, i^*); & y_i = 1 \\ -\rho h_{\sigma}(i, i^*); & y_i = 0 \end{cases}$$

$h_{\sigma}(i, i^*)$ 即是 SOFM 中的高斯函数,方差为 σ , 均值为 i^* .

(2) 按频度加权竞争学习^[4] (Frequency Sensitive Competitive Learning, FSCL)

距离度量空间 $d(x, w_i) = r_i |x - w_i|^2$

$$r_i = \frac{n_i}{\sum_j n_j}; n_i \text{ 表示获胜单元 } i \text{ 的计次, 取}$$

$$r_i = \begin{cases} \rho, & i = i^* \text{ 及 } y_i = 1 \\ -\rho, & i = i^* \text{ 及 } y_i = 0 \\ 0, & i \neq i^* \end{cases}$$

很明显,常用的 LVQ2、LVQ3 可以容易地在强化学习的框架下加以描述.

5 实验比较

对 LVQ-R 及基本 LVQ 作了性能对比实验,数据源采用熟知的 IRIS 数据集,其有 150 个 4 维点,对应三个分类,每类 50 个点,比较所采用的目标函数 J 为欧氏空间的距离度量,将分类数分别设为 3, 4, 5, 对两种方法进行了性能对比.两种方法采用相同的、随机设定的初始权向量,下表中列出运行 10 次后目标函数的平均值.

表1 LVQ 与 LVQ-R 性能比较

聚类数	LVQ	LVQ-R
3	105.3	98.5
4	95.7	76.4
5	68.5	65.2

可见, LVQ-R 有较好的性能.

结束语 本文在一个采用离散二值 bernoulli 单元的神经网络中,建立强化学习聚类方法框架,并将 LVQ 算法在此框架下加以描述,实现了 LVQ-R 算法, Bernoulli 单元的随机选择性与强化学习的奖励规则结合,很好地模拟了生物界的学习过程.实验结果验证了 LVQ-R 的性能优于基本 LVQ. 本文还提出在强化学习框架下描述其它在线竞争学习聚类算法的设想.需要进一步的研究纳入强化学习机制后,其它竞争聚

(下转第 132 页)

于重视弃权的不同考虑给出两种非概率 Vague 熵准则及相应的熵表达式。首先令 A 和 B 是论域 $U = \{\mu_1, \mu_2, \dots, \mu_n\}$ 上两个 Vague 集, 定义同本文 5.1 节。

6.1 Vague 熵准则 I 及相应的一种 Vague 熵

Vague 熵准则 I 如下:

- (1) $\hat{E}(A) = 0$, 当且仅当对所有 $\mu_i, |t_A(\mu_i) - f_A(\mu_i)| = 1$
- (2) $\hat{E}(A) = 1$, 当且仅当对所有 $\mu_i, t_A(\mu_i) - f_A(\mu_i) = 0$
- (3) $\hat{E}(A) \leq \hat{E}(B)$ 成立条件
 $|t_A(\mu) - f_A(\mu)| \geq |t_B(\mu) - f_B(\mu)|$
- (4) $\hat{E}(A) = \hat{E}(\bar{A})$

当 $t_A(\mu_i) = f_A(\mu_i)$ 时, $\hat{E}(A)$ 为最大值 1, 类似于文 [10] 确定模糊熵的思路, 我们从 $[t_A(\mu_i), f_A(\mu_i)]$ 与 $\hat{E}(A)$ 最大值时的 $[t_A(\mu_i), t_A(\mu_i)]$ 贴近程度的角度得出响应 Vague 熵准则 I 的相应一种 Vague 熵表达式如下:

定义 7 $\hat{E}(A) = 1 - \frac{1}{n} \left(\sum_{i=1}^n |t_A(\mu_i) - f_A(\mu_i)|^p \right)^{1/p}$ 是 Vague 熵准则 I 的熵。(易于验证, 证略)

Vague 熵准则 I 是从不细化分析弃权或以概率的形式细化分析弃权的角度给出的熵准则。可以这样理解: 为细化分析弃权蕴含的意义, 引入两个参数 ω_i, ω_f , ω_i 代表弃权人群中有赞同倾向而因压力或私利等原因作弃权决定的人的百分比, ω_f 代表弃权人群中有反对倾向的人的百分比, $\omega_i + \omega_f \leq 1$ 。此时新的真、假、弃权部分隶属度值 $t'_A(\mu_i), f'_A(\mu_i), 1 - t'_A(\mu_i) - f'_A(\mu_i)$ 分别为: $t'_A(\mu_i) = t_A(\mu_i) + \omega_i(1 - t_A(\mu_i) - f_A(\mu_i))$, $f'_A(\mu_i) = f_A(\mu_i) + \omega_f(1 - t_A(\mu_i) - f_A(\mu_i))$, $1 - t'_A(\mu_i) - f'_A(\mu_i) = (1 - \omega_i - \omega_f)(1 - t_A(\mu_i) - f_A(\mu_i))$ 。细化分析弃权后的结果 $t'_A(\mu_i), f'_A(\mu_i), 1 - t'_A(\mu_i) - f'_A(\mu_i)$ 仍适用 Vague 熵准则 I 及相应的 Vague 熵定义 7, 不细化分析弃权情况即 $\omega_i = \omega_f = 0$ 。另外, 当 $t_A(\mu_i) + f_A(\mu_i) = 1$ 或 $\omega_i + \omega_f = 1$ (即 $t'_A(\mu_i) + f'_A(\mu_i) = 1$) 时, Vague 熵准则 I 演化为 De Luca-Termini 准则 (即 De Luca-Termini 准则是 Vague 熵准则 I 的特例), 定义 7 演化为文 [10] 提出的模糊熵 R_p 。另外值得一提的是, 从熵的角度来看, 前文所研究的模糊相似度量 M_C 和 M_L 融入了对 Vague 值所蕴含的平均信息量的重视。

6.2 Vague 熵准则 II 及相应一种 Vague 熵

Vague 熵准则 II 如下:

- (1) $\hat{E}(A) = 0$, 当且仅当对所有 $\mu_i, |t_A(\mu_i) - f_A(\mu_i)| = 1$
- (2) $\hat{E}(A) = 1$, 当且仅当对所有 $\mu_i, t_A(\mu_i) - f_A(\mu_i) = 0$ 且 $1 - t_A(\mu_i) - f_A(\mu_i) = 1$
- (3) 当 $|t_A(\mu) - f_A(\mu)| - (1 - t_A(\mu) - f_A(\mu)) \geq |t_B(\mu) - f_B(\mu)| - (1 - t_B(\mu) - f_B(\mu))$ 时, $\hat{E}(A) < \hat{E}(B)$
- (4) $\hat{E}(A) + \hat{E}(\bar{A})$

定义 8 $\hat{E}(A) = 1 - \frac{1}{2n} \left(\sum_{i=1}^n [t_A(\mu_i) - f_A(\mu_i)] + (t_A(\mu_i) + f_A(\mu_i)) \right)$ 是 Vague 熵准则 II 下的熵。(证略)

此处是发源于对弃权对象绝对重视的观点, 实质上是基于真假差距削弱弃权的考虑。

本节给出了两种 Vague 熵准则和相应的 Vague 熵, 具体应用时应根据需要进行选取。

结论和展望 Vague 集考虑了弃权的因素, 但现有的 Vague 相似度量方法中并未深入彻底地考虑弃权的发展趋向或动向。为了使结果与真实情况更加接近, 应重视对弃权部分的细致分析, 深入研究弃权者的心理态势, 这是改进现有 Vague 相似度量方法的渠道之一。另外当前方法都建立在距离基础上, 只能表示 Vague 集之间和元素之间总的距离, 无法表示 Vague 集之间最小的相似性和元素之间最大的相似性, 可以尝试把模糊集的一些研究成果灵活地运用于 Vague 集, 进一步推进对 Vague 相似度量的研究。

总之, 本文讨论了 Vague 值 (集或元素) 之间的相似度量问题, 假定有 n 个元素: $x_1, x_2, \dots, x_n$, 且分为 n 组: $A_1, A_2, \dots, A_n$, 我们不清楚 x_i 和 A_i 的情况, 仅知道每个元素在多大程度上包含在每个组中, 那么本文讨论的方法可以较好地回答: 1. 任意两个组 A_i 和 $A_j (i \neq j)$ 在多大程度上可以组合在一起? 2. x_i 和 $x_j (i \neq j)$ 有多大可能属同一类? 这种解决问题的方法可以广泛应用在协商、投票、市场定位、管理、方案评估模型中。另外, 基于对弃权对象的不同考虑, 本文还给出了两种 Vague 熵准则和相应的 Vague 熵, 它们对 Vague 集的研究和应用有实际意义, 在此希望抛砖引玉, 引发对它的深入探讨。

参考文献

- 1 Zadeh L A. Fuzzy sets. Inform. and Control, 1965, 8: 338~356
- 2 Gau W L, Buehrer D J. Vague sets. IEEE Trans, Syst. Man, Cybern, 1993, 23(2): 610~614
- 3 Atanassov K. Intuitionistic fuzzy sets. Fuzzy Sets Syst, 1986, 20(1): 87~96
- 4 Chen S M. Measures of similarity between vague sets. Fuzzy Sets Syst, 1995, 74(2): 217~223
- 5 Chen S M. Similarity measures between vague sets and between elements. IEEE Trans, Syst. Man, Cybern, 1997, 27(1): 153~158
- 6 Hong D H, Kim C. A note on similarity measures between vague sets and between elements. Information Science, 1999, 115: 83~96
- 7 李凡, 徐章艳. Vague 集之间的相似度量. 软件学报, 2001, 12(6): 922~927
- 8 李凡. 模糊信息处理系统. 北京: 北京大学出版社, 1998
- 9 Luca A D, Termini S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. Inform. and Control, 1972, 20: 301~312
- 10 Deqin Y. A new approach on information measures for fuzzy sets. Pattern Recognition and Artificial Intelligence, 2001, 14(1): 23~27

(上接第 134 页)

类算法的性能, 并进一步探讨将本文的方法在数据挖掘及基于强化学习的智能控制中应用的可能性。

参考文献

- 1 Kaelbling L, Littman M, Moore A. Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research, 1996, 4: 237~285
- 2 Barto A G, Anandan P. Pattern Recognizing Stochastic Learning

Automata. IEEE Transactions on systems, Man and Cybernetics, 1985, 15: 360~375

- 3 Williams R J. Simple statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning. Machine Learning, 1992, 8: 229~256
- 4 Ahalt S C, et al. Competitive Learning Algorithms for Vector Quantization. Neural Networks, 1990, 3: 277~291
- 5 张智星, 孙春在, 水谷英二. 神经-模糊和软计算. 张平等译, 西安交通大学出版社, 2002. 2