

文章编号:1001-9081(2006)09-2044-03

## 动态模糊机器学习模型及验证

张 静,李凡长

(苏州大学 计算机科学与技术学院,江苏 苏州 215006)

(210413066@suda.edu.cn)

**摘 要:**根据学习系统中存在的动态模糊性,提出了动态模糊机器学习模型,给出了动态模糊机器学习算法和它的几何模型描述,并进行了算法的稳定性分析,最后给出了实例验证。实例结果与 BP 算法产生结果相比较,优于 BP 算法的结果。

**关键词:**动态模糊集;动态模糊机器学习系统;动态模糊机器学习模型

**中图分类号:**TP181 **文献标识码:**A

## Dynamic fuzzy machine learning model and its validation

ZHANG Jing, LI Fan-zhang

(College of Computer Science and Technology, Soochow University, Suzhou Jiangsu 215006, China)

**Abstract:** According to the dynamic fuzzy characteristic of the machine learning system, a dynamic fuzzy machine learning model was proposed, put forward the dynamic fuzzy machine learning algorithm, described the description of its geometry model, and finally, the theory was verified by giving an example. The result shows that the algorithm of this paper is superior to that of Back Propagation, so it is verified that the proposed algorithm of this paper is more effective.

**Key words:** dynamic fuzzy sets; dynamic fuzzy machine learning system; dynamic fuzzy machine learning model

### 0 引言

机器学习是人工智能和计算机科学领域的一个永恒的课题。目前,计算机科学,特别是人工智能的研究者一般公认 Simon 对学习的论述:“如果一个系统能够通过执行某种过程而改进它的性能,这就是学习。”这个说法的要点是:其一,学习是一个过程;其二,学习是对一个系统而言的;其三,学习能够改变系统的性能。“过程”、“系统”与“改变性能”是学习的三个要点。从这三个要点可以看出机器学习系统是一个普遍存在动态模糊性的系统:如学习过程,本质上具有动态模糊性;系统的变化,系统的好坏等本质上也是具有动态模糊性的;改变系统性能,其改变的结果如何,改变的性能如何等均具有动态模糊性。因此,我们认为要是建立有实质性进展且能满足 Simon 说法要点的机器学习方法,其关键问题是要能对学习活动中产生的动态模糊性进行有效描述。文献[1,4]给出了动态模糊数据的理论基础,文献[2,3]给出了协调机器学习的模型及相关算法,本文在此基础上提出了动态模糊机器学习模型(Dynamic Fuzzy Machine Learning Model, DFMLM)。

### 1 动态模糊机器学习系统的基本概念

一个系统的学习可以认为是它在运行过程中对自身的调整,体现为系统的结构或参数的一系列的变化,如果用数学语言来描述,学习就可以定义为一个集合到另一个集合的映射。

**定义 1** 动态模糊学习空间

由一切动态模糊学习要素构成的用于描述学习过程的空

间称为动态模糊学习空间。它由{学习样例、学习算法、输入数据、输出数据、表示理论}五个元素组成,可表示为  $(\tilde{S}, \tilde{S}) = \{(\vec{E}x, \vec{E}x), ER, (\tilde{X}, \tilde{X}), (\tilde{Y}, \tilde{Y}), ET\}$

**定义 2** 动态模糊学习

动态模糊学习  $(\tilde{I}, \tilde{I})$  是指在动态模糊空间  $(\tilde{S}, \tilde{S})$  中的一个输入数据集  $(\tilde{X}, \tilde{X})$  到一个输出数据集  $(\tilde{Y}, \tilde{Y})$  的映射:  $(\tilde{I}, \tilde{I}): (\tilde{X}, \tilde{X}) \rightarrow (\tilde{Y}, \tilde{Y})$

**定义 3** 动态模糊机器学习系统

动态模糊学习空间  $(\tilde{S}, \tilde{S})$  中的五要素按照一定的学习机制形成的具有学习能力的计算机系统,称为动态模糊机器学习系统(Dynamic Fuzzy Machine Learning System, DFMLS)。

**定义 4** 动态模糊机器学习模型<sup>[2]</sup> DFMLM

$DFMLM = \{(\tilde{S}, \tilde{S}), (\tilde{L}, \tilde{L}), (\tilde{u}, \tilde{u}), (\tilde{y}, \tilde{y}), (\tilde{p}, \tilde{p}), (\tilde{I}, \tilde{I}), (\tilde{O}, \tilde{O})\}$ , 其中  $(\tilde{S}, \tilde{S})$  是被学习部分(动态环境/动态模糊学习空间),  $(\tilde{L}, \tilde{L})$  是动态学习部分,  $(\tilde{u}, \tilde{u})$  是  $(\tilde{S}, \tilde{S})$  到  $(\tilde{L}, \tilde{L})$  的输出,  $(\tilde{y}, \tilde{y})$  是  $(\tilde{L}, \tilde{L})$  到  $(\tilde{S}, \tilde{S})$  的动态反馈,  $(\tilde{p}, \tilde{p})$  是系统学习性能指标,  $(\tilde{I}, \tilde{I})$  是外界环境对动态模糊学习系统的输入数据集,  $(\tilde{O}, \tilde{O})$  是本系统对外界的输出。

现在对系统采用离散化处理,  $(\tilde{S}, \tilde{S}), (\tilde{L}, \tilde{L})$  用状态空间表示,于是有:

**定义 5** 动态模糊学习模型 DFMLM 描述为:

$$(\hat{x}, \hat{x})(k+1) = G_1((\hat{x}, \hat{x})(k), (\tilde{u}, \tilde{u})(k), \zeta_1(k)) \quad (1)$$

$$(\hat{y}, \hat{y})(k) = G_2((\hat{x}, \hat{x})(k), (\tilde{u}, \tilde{u})(k), \zeta_2(k)) \quad (2)$$

$$(\hat{p}, \hat{p})(k) = \sum_{i=1}^k P((\hat{y}, \hat{y})(i)) \quad (3)$$

其中:  $(\hat{x}, \hat{x})(k)$  是  $(\tilde{S}, \tilde{S})$  在时刻  $k$  的状态变量,  $(\tilde{u}, \tilde{u})(k)$  是

$(\tilde{S}, \tilde{S})$  的动态输出,  $\zeta_1(k)$  是状态方程中的随机干扰,  $(\hat{y}, \tilde{y})(k)$  是  $(\tilde{L}, \tilde{L})$  的动态反馈输出,  $\zeta_2(k)$  是观察随机误差。  $k$  表示时刻, 只取整数。假定其中的向量全部为有限维状态变量。 $(\hat{p}, \tilde{p})$  表示系统学习性能指标,  $P(i)$  是一个标量函数, 表示时刻  $i$  时的系统学习性能。

根据定义可得如下几个命题:

命题 1 DFMLS 是随机系统;

命题 2 DFMLS 是开放系统;

命题 3<sup>[2]</sup> DFMLS 是非线性系统。

## 2 动态模糊机器学习算法

考虑动态模糊机器学习系统由如下动态模糊规则组建:

IF  $(\hat{x}, \tilde{x})(k)$  is  $(\tilde{A}_1^p, \tilde{A}_1^p)$  and  $(\hat{x}, \tilde{x})(k-1)$  is  $(\tilde{A}_2^p, \tilde{A}_2^p)$  and ... and  $(\hat{x}, \tilde{x})(k-m+1)$  is  $(\tilde{A}_m^p, \tilde{A}_m^p)$  and  $(\tilde{u}, \tilde{u})(k)$  is  $(\tilde{B}^p, \tilde{B}^p)$

THEN

$$(\hat{x}, \tilde{x})^p(k+1) = (\tilde{a}_1^p, \tilde{a}_1^p)(\hat{x}, \tilde{x})(k) + \dots + (\tilde{a}_m^p, \tilde{a}_m^p)(\hat{x}, \tilde{x})(k-m+1) + (\tilde{b}^p, \tilde{b}^p)(\tilde{u}, \tilde{u})(k)$$

其中:  $(\tilde{A}_i^p, \tilde{A}_i^p)$ 、 $(\tilde{B}^p, \tilde{B}^p)$  为动态模糊集,  $(\tilde{a}_i^p, \tilde{a}_i^p)$ 、 $(\tilde{b}^p, \tilde{b}^p)$  为常动态模糊数,  $(p=1, 2, \dots, m)$ ,  $(\tilde{u}, \tilde{u})(k)$  见定义 4;

$(\tilde{X}, \tilde{X})(k) = ((\hat{x}, \tilde{x})(k), (\hat{x}, \tilde{x})(k-1), \dots, (\hat{x}, \tilde{x})(k-m+1))^T \in (\tilde{U}, \tilde{U})$  为系统状态向量。

则由定义 5 给出的动态模糊机器学习系统的学习算法如下:

算法:

procedure ER

/\*  $(\tilde{u}_{Ap}, \tilde{u}_{Ap})$ ,  $(\tilde{u}_{Bp}, \tilde{u}_{Bp})$  为相应的动态模糊隶属度,  $(\hat{y}_d, \tilde{y}_d)$  为期望输出 \*/

$$(\hat{x}, \tilde{x})(k+1) = \frac{\sum_{p=1}^m (\hat{x}, \tilde{x})^p(k+1)(\tilde{v}, \tilde{v})^p}{\sum_{p=1}^m (\tilde{v}, \tilde{v})^p} \quad (4)$$

$$(\tilde{v}, \tilde{v})^p = \prod_{i=1}^m (\tilde{u}_{Ap}, \tilde{u}_{Ap}) [(\hat{x}, \tilde{x})(k-i+1)] \times (\tilde{u}_{Bp}, \tilde{u}_{Bp}) [(\tilde{u}, \tilde{u})(k)] \quad (5)$$

$$(\hat{y}_k, \tilde{y}_k) = \frac{\sum_{j=1}^k (\tilde{h}_j, \tilde{h}_j)(\hat{x}, \tilde{x})(j)}{\sum_{j=1}^k (\tilde{h}_j, \tilde{h}_j)} \quad (6)$$

$$E(\hat{y}_k, \tilde{y}_k) = (\hat{y}_d, \tilde{y}_d) - (\hat{y}_k, \tilde{y}_k) \quad (7)$$

$$(\hat{p}, \tilde{p})(k) = \frac{1}{2} \sum_{i=1}^k [E(\hat{y}_i, \tilde{y}_i)]^2 \quad (8)$$

If  $|E(\hat{y}_k, \tilde{y}_k)| > (\tilde{\varepsilon}, \tilde{\varepsilon})$

then 舍弃  $(\hat{y}_k, \tilde{y}_k)$ , 反馈给规则库, 通知其修改规则中的常动态模糊系数

//  $(\tilde{\varepsilon}, \tilde{\varepsilon})$  是一个很小的值, 它可根据系统要求或经验给定

$$(\tilde{u}, \tilde{u})(k) = (\tilde{u}, \tilde{u})(k) + \alpha E(\hat{y}_k, \tilde{y}_k)$$

继续从头执行

else if  $|E(\hat{y}_k, \tilde{y}_k)| \leq (\tilde{\delta}, \tilde{\delta})$

/\*  $(\tilde{\delta}, \tilde{\delta})$  是一个比  $(\tilde{\varepsilon}, \tilde{\varepsilon})$  小的多的值, 它可根据系统要求或经验给定 \*/

then 本次学习结束, 输出  $(\hat{y}_k, \tilde{y}_k)$

else

$$(\tilde{u}, \tilde{u})(k) = (\tilde{u}, \tilde{u})(k) + \alpha E(\hat{y}_k, \tilde{y}_k)$$

继续从头执行

endif

endif

end ER.

注:

1) 根据本系统的特点, 我们选取动态模糊隶属函数为正态函数:

$$(\tilde{u}_{A_i^p}(x_i), \tilde{u}_{A_i^p}(x_i)) = \left( \exp \left[ - \left( \frac{(\hat{x}_i, \tilde{x}_i) - (\tilde{m}_i^j, \tilde{m}_i^j)}{\sqrt{2}(\tilde{\sigma}_i^j, \tilde{\sigma}_i^j)} \right)^2 \right], \exp \left[ - \left( \frac{(\hat{x}_i, \tilde{x}_i) - (\tilde{m}_i^j, \tilde{m}_i^j)}{\sqrt{2}(\tilde{\sigma}_i^j, \tilde{\sigma}_i^j)} \right)^2 \right] \right)$$

2)  $(\tilde{h}_j, \tilde{h}_j)$  为相应的权系数, 一般地, 取隶属函数作为权系数。

## 3 动态模糊机器学习几何模型描述

### 3.1 动态模糊机器学习系统几何模型描述

如图 1 所示, 我们把论域定义为一个动态模糊球空间 (大球体), 在这个球空间中, 用一些小的球体来表示论域中的动态模糊集, 每个动态模糊数定义为动态模糊集 (小球体) 中的一点, 每个动态模糊数的隶属度由它所在的动态模糊集 (小球体) 在论域 (大球体) 中的位置和半径及该数在该动态模糊集的位置决定。

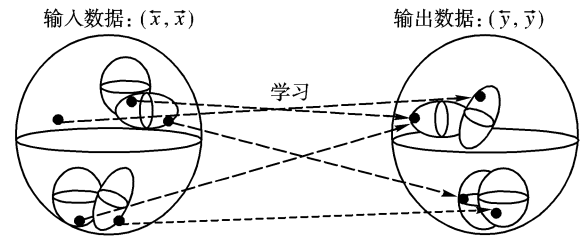


图 1 动态模糊机器学习系统几何模型

### 3.2 动态模糊机器学习算法的几何模型

两球的中心点为期望值  $(\hat{y}_d, \tilde{y}_d)$ , 大球的半径为  $(\tilde{\varepsilon}, \tilde{\varepsilon})$ , 小球的半径为  $(\tilde{\delta}, \tilde{\delta})$  ( $(\tilde{\varepsilon}, \tilde{\varepsilon})$ 、 $(\tilde{\delta}, \tilde{\delta})$  的意义同第 2 节动态模糊学习算法)。

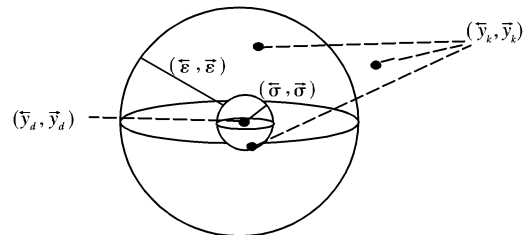


图 2 动态模糊机器学习算法的几何模型

该几何模型描述:

1) 学习算法所得的值  $(\hat{y}_k, \tilde{y}_k)$  落在大球外, 则视本次学习无效, 舍弃  $(\hat{y}_k, \tilde{y}_k)$ ;

2) 若  $(\hat{y}_k, \tilde{y}_k)$  落在大球与小球之间, 则修正  $(\hat{y}_k, \tilde{y}_k) = (\hat{y}_k, \tilde{y}_k) + \beta E(\hat{y}_k, \tilde{y}_k)$ ;

3) 若  $(\hat{y}_k, \tilde{y}_k)$  落在小球内, 则认为已达到精度要求, 结束学习, 输出  $(\hat{y}_k, \tilde{y}_k)$ 。

## 4 稳定性分析

定义状态向量  $(\tilde{X}, \tilde{X})(k) = ((\hat{x}, \tilde{x})(k), (\hat{x}, \tilde{x})(k-1), \dots, (\hat{x}, \tilde{x})(k-m+1))^T$

$$(\vec{A}_p, \vec{A}_p) = \begin{bmatrix} (\vec{a}_1^p, \vec{a}_1^p) & (\vec{a}_2^p, \vec{a}_2^p) & \cdots & (\vec{a}_{m-1}^p, \vec{a}_{m-1}^p) & (\vec{a}_m^p, \vec{a}_m^p) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

则式(4)可写为:

$$(\vec{X}, \vec{X})(k+1) = \frac{\sum_{p=1}^m (\vec{A}_p, \vec{A}_p)(\vec{X}, \vec{X})(k)(\vec{v}, \vec{v})^p}{\sum_{p=1}^m (\vec{v}, \vec{v})^p} \quad (9)$$

∴ 当  $(\vec{X}, \vec{X})(k) = (\vec{0}, \vec{0})$  时, (9) 式右边等于  $(\vec{0}, \vec{0})$

$$\begin{aligned} \Delta V[(\vec{X}, \vec{X})(k)] &= (\vec{X}, \vec{X})^T(k+1)(\mathbf{P}, \mathbf{P})(\vec{X}, \vec{X})(k+1) - (\vec{X}, \vec{X})^T(k)(\mathbf{P}, \mathbf{P})(\vec{X}, \vec{X})(k) \\ &= (\vec{X}, \vec{X})^T(k) \left\{ \left[ \frac{\sum_{p=1}^m (\vec{A}_p, \vec{A}_p)^T(\vec{v}, \vec{v})^p}{\sum_{p=1}^m (\vec{v}, \vec{v})^p} \right] \times (\mathbf{P}, \mathbf{P}) \left[ \frac{\sum_{p=1}^m (\vec{A}_p, \vec{A}_p)(\vec{v}, \vec{v})^p}{\sum_{p=1}^m (\vec{v}, \vec{v})^p} \right] - (\mathbf{P}, \mathbf{P}) \right\} (\vec{X}, \vec{X})(k) \\ &= \frac{\sum_{p=1}^m \sum_{q=1}^m \{ (\vec{v}, \vec{v})^p (\vec{v}, \vec{v})^q (\vec{X}, \vec{X})^T(k) \times [ (\vec{A}_p, \vec{A}_p)^T(\mathbf{P}, \mathbf{P})(\vec{A}_q, \vec{A}_q) - (\mathbf{P}, \mathbf{P}) ] (\vec{X}, \vec{X})(k) \}}{\sum_{p=1}^m \sum_{q=1}^m (\vec{v}, \vec{v})^p (\vec{v}, \vec{v})^q} \end{aligned}$$

由式(10)及  $(\vec{v}, \vec{v})^p \geq (\vec{0}, \vec{0})$ , 得到  $V[(\vec{X}, \vec{X})(k)] < (\vec{0}, \vec{0})$ 。根据 Lyapunov 稳定性定理, 本定理得证。

### 5 实例验证

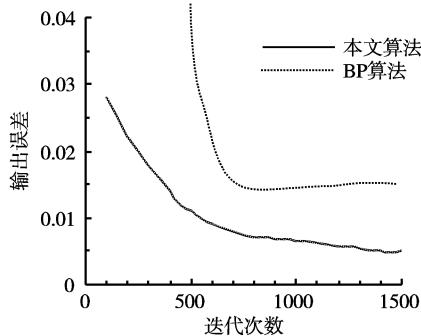


图3 两种算法的误差与迭代步数比较

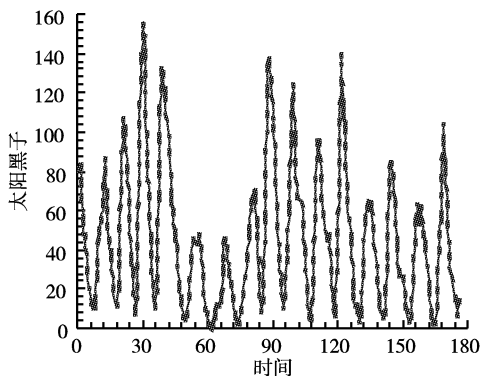


图4 本文算法的预测值与实际值比较

太阳黑子问题: 数据取自文献[8], 从 1749 年到 1924 年, 共有 176 个数据, 其中前 100 个数据作为训练样本数据, 其余的用作检验数据。

图3 是 BP 算法和本文算法的误差与迭代步数比较。

∴ 论域中的原点  $(\vec{0}, \vec{0})$  即为(9) 式的平衡点

定理<sup>[5-7]</sup> 动态模糊学习系统表达式(9) 的平衡点

$(\vec{0}, \vec{0})$  是全局渐近稳定的, 若存在一正定动态模糊矩阵  $(\mathbf{P}, \mathbf{P})$ , 使得:

$$(\vec{A}_p, \vec{A}_p)^T(\mathbf{P}, \mathbf{P})(\vec{A}_p, \vec{A}_p) - (\mathbf{P}, \mathbf{P}) < (\vec{0}, \vec{0}), \quad p = 1, 2, \dots, m \quad (10)$$

证明 考虑如下的 Lyapunov 函数:

$$V[(\vec{X}, \vec{X})(k)] = (\vec{X}, \vec{X})^T(k)(\mathbf{P}, \mathbf{P})(\vec{X}, \vec{X})(k)$$

其中  $(\mathbf{P}, \mathbf{P})$  为一正定动态模糊矩阵, 则有:

BP 算法当  $E = 0.4$  时陷入局部极小, 所需迭代步数为 2700; 本文算法当误差  $E(\vec{y}_k, \vec{y}_k) < (0.4, 0.4)$  时所需的迭代步数为 450。当误差  $E(\vec{y}_k, \vec{y}_k) = (0.34, 0.34)$ , 训练基本稳定, 所需迭代步数为 1100。图4 是置初始值  $(\vec{u}_0, \vec{u}_0)(t) = (\vec{0}, \vec{0})$ , 增益学习系数  $\alpha = 0.3, \beta = 0.2, (\vec{\varepsilon}, \vec{\varepsilon}) = (0.5, 0.5), (\vec{\delta}, \vec{\delta}) = (0.005, 0.005)$ , 当误差  $E(\vec{y}_k, \vec{y}_k) = (0.34, 0.34)$  时实际值与预测值的比较图。

### 6 结语

综上所述, 本文针对机器学习系统中具有动态模糊性, 利用动态模糊集构建了动态模糊机器学习空间的基本概念, 描述了动态模糊机器学习模型并给出了相关的学习算法, 通过和 BP 算法的比较, 取得了比较满意的结果。

#### 参考文献:

- [1] 李凡长, 郑家亮. 动态模糊数据模型研究[J]. 计算机研究与发展, 1998, 35(8): 714-718.
- [2] 李凡长. 基于 DFS 的协调机器学习模型[J]. 计算机工程, 2001, 27(3): 106-110.
- [3] 李凡长. 协调机器学习的稳定性研究[J]. 小型微型计算机系统, 2002, 23(3): 314-317.
- [4] 李凡长, 刘贵全, 余玉梅. 动态模糊逻辑引论[M]. 昆明: 云南科技出版社, 2005.
- [5] 曹建福, 韩崇昭, 方洋旺. 非线性系统理论及应用[M]. 西安: 西安交通大学出版社, 2001.
- [6] 蔡自兴, 肖晓明. 动态模糊系统的鲁棒稳定性分析与控制器设计[J]. 中南工业大学学报, 1999, 30(4): 418-421.
- [7] KHALIL HK. Nonlinear Systems[M]. Third Edition. Publishing House of Electronics Industry, 2005.
- [8] 杨叔子, 吴雅. 时间序列分析的工程应用[M]. 武汉: 华中理工大学出版社, 1994.