

基于认知模拟的自适应机器学习算法研究*

王继成

(同济大学 计算机科学与工程系, 上海 200092)

E-mail: jcwangk@online.sh.cn

http://www.tongji.edu.cn

摘要: 根据认知科学的理论,同时以神经科学、脑科学等学科的最新研究成果为指导,利用计算机科学、数学及工程科学的方法,提出了一种自适应机器学习方法。该方法从微观神经元层次、中间层次和宏观层次,根据人脑的学习机理模拟人脑的学习过程。根据该方法开发的心电图分类知识自动获取系统,可以较好地模拟心电图专家的感知、认知和思维等功能。

关键词: 认知科学;机器学习;神经网络;心电图分析

中图分类号: TP181 **文献标识码:** A

学习是人类智能的主要标志和获得智慧的基本手段。机器学习的研究就是希望计算机能像人类那样具有从现实世界获取知识的能力,同时进一步发现人类学习的机理和揭示人脑的奥秘。在20世纪50年代末和60年代初,人们就开展了基于神经网络的联接学习机制和学习算法的研究,但由于线性感知器的局限性,同时,传统的 Von Neuman 型数字计算机此时正处在发展的全盛时期,而且基于符号的人工智能得到迅速的发展,使得早期的基于联接机制的非符号学习被符号学习的研究所替代。20世纪80年代以来,基于符号机制的机器学习的研究,发展甚为迅速,其间研究了各种学习算法,如讲授学习、解释学习、类比学习、示例学习、概念聚类等。但是,随着研究的不断深入,不难发现,无论是基于联接机制的机器学习研究,还是基于符号机制的机器学习研究,其方法和研究手段都带有一定的局限性和片面性,不利于揭示人脑的思维机制和学习机理,它们都只从单一的层次来描述学习过程。对于人脑如何接受外界信息,通过其内部联接机制的变化,从而反映在系统行为的变化这样的问题,现有的理论不能给出一个大家都能接受的、合理的解释^[1]。

事实上,学习是受一定的意志支配的(即有特定的学习目的),其内部表现为一定的结构(即基于联接机制),其外部表现为一定的行为变化(即基于符号学习)的复杂过程。它涉及到联接理论、认知科学、行为科学、神经科学等多门科学。因此,对于机器学习的研究,只能采用计算机科学、控制论、人工智能、认知科学、神经科学、心理学等多学科交叉的方法,才可望取得机器学习研究的更大进展^[2]。

本文根据认知科学的理论,同时,以神经科学、脑科学等学科的最新研究成果为指导,利用计算机科学、数学及工程科学的方法,提出了一种自适应机器学习方法。该方法从微观神经元层次、中间层次、宏观层次,根据人脑的学习机理模拟人脑的学习过程和学习方法。本文首先提出人脑自适应学习理论,然后研究基于认知模拟的自适应机器学习算法,最后通过根据该算法开发的心电图分类知识自动获取系统的实验和讨论,进一步证实该算法的可行性和先进性。

* 收稿日期: 1999-09-10; 修改日期: 2000-03-28

基金项目: 国家自然科学基金资助项目(69285005)

作者简介: 王继成(1958—),男,江苏盐城人,博士,教授,博士生导师,主要研究领域为人工智能,神经网络,认识科学。

1 人脑自适应学习理论

根据神经科学的研究,人脑是由大约 10^{11} 个神经元组成的巨系统,并且各神经元之间约有 $10^3 \sim 10^4$ 个突触相互联结.人自出生之后,神经元的数目略有增加,但随着年龄的增长而逐步死亡.与此相反的是,神经元之间的联结随着人们不断学习新知识而变得越来越复杂,同时,脑中逐步清晰地形成特定的功能区.从认知的角度来看,脑中形成了大量的现实模板.在此过程中,人脑在微观(神经元水平)、中间(功能核团、功能区水平)以及宏观(皮质水平)层次上均有相应的控制机制来制导整个自适应学习过程,其中包括神经元的状态及其转换、神经网络的结构优化以及神经网络学习函数的选择等^[3].

在神经元级,神经元的自适应学习功能是由神经元的状态及其自适应转移规则实现的.神经元的状态有:睡眠状态(或超抑制状态)、正常状态、临界状态、病态、死亡状态.其状态转换如图1所示.

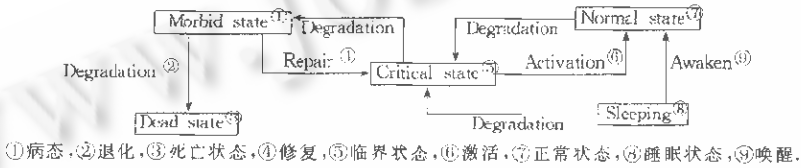


Fig.1 The transformations of neuron states

图1 神经元状态转换图

在中间级(功能核团、功能区),神经网络的自适应学习过程表现在:神经网络学习函数的选择与替换以及多目标神经网络结构优化^[4].神经网络在学习过程中,根据学习的样本和学习的不同阶段选择合适的学习函数,以提高学习速度.神经网络的学习函数可表示为

$$\delta_{ik} = g(y_{ik}) * (t_{ik} - y_{ik}), \quad (1)$$

其中: $g(y_{ik})$ 在神经网络学习的不同阶段,可以根据需要取 $\cos(y)$, $\cosh(y)$, $1 - y^2$, $\tanh(y)$, $\sin(y)$, e^{-y} 等(这里 $y = 1/(1 + e^{-x})$),以提高神经网络的学习速度; t_{ik} 为神经元 i 在输入模式 k 时的期望输出值; y_{ik} 为神经元 i 在输入模式 k 时的实际输出值.

通过研究发现,当神经元 i 在输入模式 k 下的期望输出值为 1、实际输出值为 0 时,选择 $\cos(y)$, $1 - y^2$, e^{-y} , $1 - \tanh(y)$ 作为学习函数可以获得高的收敛速度.同样,当神经元 i 在输入模式 k 下的期望输出值为 0、实际输出值为 1 时,选择 $\cosh(y)$, $\tanh(y)$, $\sin(y)$, $1/\cos(y)$, $\tanh(y^2)$, $\sinh(y)$ 作为学习函数可以获得高收敛速度.多目标神经网络结构优化就是从多个目标函数出发(或许多个目标间是相互制约的),通过神经网络的学习优化神经网络的结构,以形成若干个特定的功能区(功能核团)^[5].具体地说,就是最小化多目标误差能量函数 E :

$$\begin{aligned} E &= W_1 \cdot E_f + W_2 \cdot W_n + W_3 \cdot E_v \\ &= W_1 \left(0.5 \sum_p \sum_i |O_{pi} - d_{pi}|^2 \right) + W_2 \left(0.5 \sum_j \left(\frac{\partial E_f}{\partial I_j} \right)^2 \right) + W_3 \left(\sum_l \sum_k |W_{lk}| \right), \end{aligned} \quad (2)$$

其中 O_{pi} 为输出层第 i 节点在第 p 输入模式下的实际输出, d_{pi} 为输出层第 i 节点在第 p 输入模式下的期望输出, $\frac{\partial E_f}{\partial I_j}$ 是 E_f 关于输入层第 j 节点的导数, W_{lk} 是第 l 节点与第 k 节点间的联接权.

在宏观上,神经系统的自适应学习能力主要表现在:根据不同性质的问题自动选择学习方法和学习方式,以获得最佳的学习效果.人类在学习时,首先从实践中获取大量的感性知识,然后经过去

粗取精、去伪存真、由此及彼、由表及里的理性思维过程, 把感性认识的东西上升为理性知识, 这些理论知识一方面用于解决以后遇到的实际问题, 同时也用于指导以后的学习, 这些知识对以后获取新的知识至关重要. 因此, 神经系统自适应学习问题在宏观上主要是从隐含于神经网络结构的分布知识中提取出符号知识, 并用这些符号知识指导神经网络的学习.^[6]

2 自适应机器学习算法

机器学习研究的目的是希望计算机能像人类那样具有从现实世界中获取知识的能力, 因此, 机器学习算法必须以人脑自适应学习理论为指导, 其中包括: 学习函数的自适应选择、多目标优化学习准则、符号知识的自适应提取等. 当然, 机器学习不可能也没有必要走完全模拟人类学习的道路, 而应该充分吸取人类学习的精华为机器学习所用. 所以, 本节研究的自适应机器学习算法, 首先采用联接机制的方法(其中, 包括上节讨论的神经元级和中间级的自适应学习方法). 从实例中学习, 然后提取出隐含于联接结构中的符号知识、概念, 最后用这些符号知识指导联接机制的学习, 进一步优化神经网络的结构.^[7~10]

为了叙述方便, 本节只讨论单层和多层神经网络的结构. 对于单层神经网络, Hebb 规则可以作为网络的学习规则, 即

$$\Delta W_{ij} = \epsilon * a_i * a_j, \quad (3)$$

其中 ϵ 是学习速率, a_i 是神经元 i 的激活值, a_j 是神经元 j 的激活值, ΔW_{ij} 是神经元 i 与神经元 j 之间联接权的改变量.

通过学习, 神经网络从大量实例中获取知识, 并以分布的形式存在于神经网络的结构中. 对于多层神经网络, 采用反向传播(BP)算法训练神经网络, 即

$$\Delta W_{ij} = \epsilon * \delta_{pj} * a_{pj}. \quad (4)$$

当神经元 j 为输出神经元时,

$$\delta_{pj} = (t_{pj} - a_{pj}) * f'_{pj}(\text{net}_{pj});$$

当神经元 j 为隐含神经元时,

$$\begin{aligned} \delta_{pj} &= f'_{pj}(\text{net}_{pj}) * \delta_{pk} * W_{kj}, \\ f'_{pj}(\text{net}_{pj}) &= a_{pj} * (1 - a_{pj}), \end{aligned}$$

其中 ΔW_{ij} 是神经元 i 与神经元 j 之间联接权的改变量, ϵ 是学习速率, a_{pj} 是神经元 j 关于输入模式 p 的激活值, δ_{pj} 是联接到神经元 i 的联接权关于输入模式 p 的改变量, t_{pj} 是神经元 j 关于输入模式 p 的期望激活值.

神经网络通过学习, 提取出隐含于样本数据中的知识, 这些知识分布存放于神经网络的结构中. 本节讨论的基于认知模拟的自适应机器学习方法, 首先采用联接学习方法(其中包含了神经元的状态及其转换、学习函数自适应选择、多目标优化神经网络等)获取样本数据中的知识, 然后采用下面讨论的方法从神经网络结构中提取出符号知识.

对于单层神经网络的学习, 学习结束后首先找出 W_{ik} , 使之满足:

$$W_{ik} = \max(W_{1k}, \dots, W_{mk}), \quad (5)$$

其中 W_{ik} 是输入神经元 i 与输出神经元 k 之间的联接权. 选择

$$W_{jk} \geq \epsilon * W_{ik},$$

其中 ϵ 是比例系数.

根据上式, 构造以下规则:

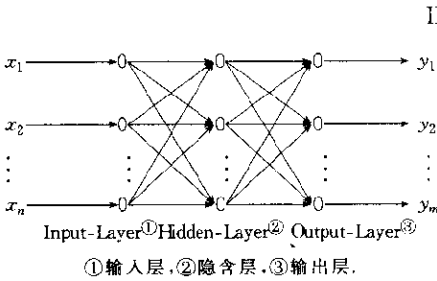


Fig. 2 The multi-layer neural network structure
图2 多层神经网络结构

$$\text{IF } a_j[\text{and } a_j]^* \rightarrow b_k, \tag{6}$$

其中 a_j 为输入神经元 j 的激活值, b_k 为输出神经元 k 的输出。

重复上述过程, 可以进一步提取出隐含在样本数据中的概念。

对于多层神经网络的学习, 为了叙述方便, 不妨设神经网络的结构如图 2 所示。

为了研究其输入-输出关系, 把输出神经元的输出作为神经网络的附加输入信号, 经过此变换后, 新的神经网络结构如图 3 所示。

神经网络结构如图 3 所示。

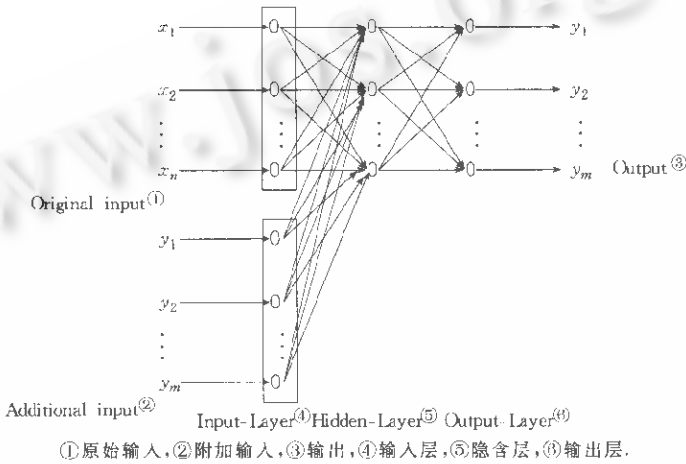


Fig. 3 The transferred multi-neural network structure
图3 变换后的多层神经网络结构

使用 BP 学习算法训练神经网络, 学习结束后, 如果从某个输入(原始)到隐含神经元的联接权与输入层中附加输入到该隐含神经元的联接权相等或相近, 则假设该输入与对应的输出(附加的输入)有密切的联系。为了描述这种联系, 设

$$SSE_{ab} = \sum_{j=1}^k (W_{bj} - W_{aj})^2, \tag{7}$$

其中 k 是隐含神经元数, W_{bj} 是附加输入神经元 b 与隐含神经元 j 之间的联接权, W_{aj} 是原始输入神经元 a 与隐含神经元 j 之间的联接权。

从上式可以看出, SSE_{ab} 的值越小, 表示输入神经元 a 与输出神经元 b 联系越紧密。因此, 在此基础上的分析, 可以在中间层次上描述神经网络的功能核团或功能区, 提取出隐含于神经网络的符号知识。这些符号知识一方面用于高层推理, 进一步获取深层知识; 另一方面, 用于优化神经网络结构, 删除神经网络中冗余的联接权和冗余神经元, 从而提高神经网络的泛化能力和学习速度。

3 实验与讨论

为了研究基于认知模拟的自适应机器学习算法的性能及其特点, 我们实现了心电图分类知识自动获取系统。用美国 MIT BME 中心提供的 MIT/BIH 心电图数据库中的部分心电图波形 T_{10c} , T_{105} , T_{106} , T_{108} , T_{111} , T_{112} , T_{217} , T_{219} , T_{220} 和 T_{221} 进行心电图分类知识自动获取的实验研究。图 4 示出

了各种心电图的部分波形。

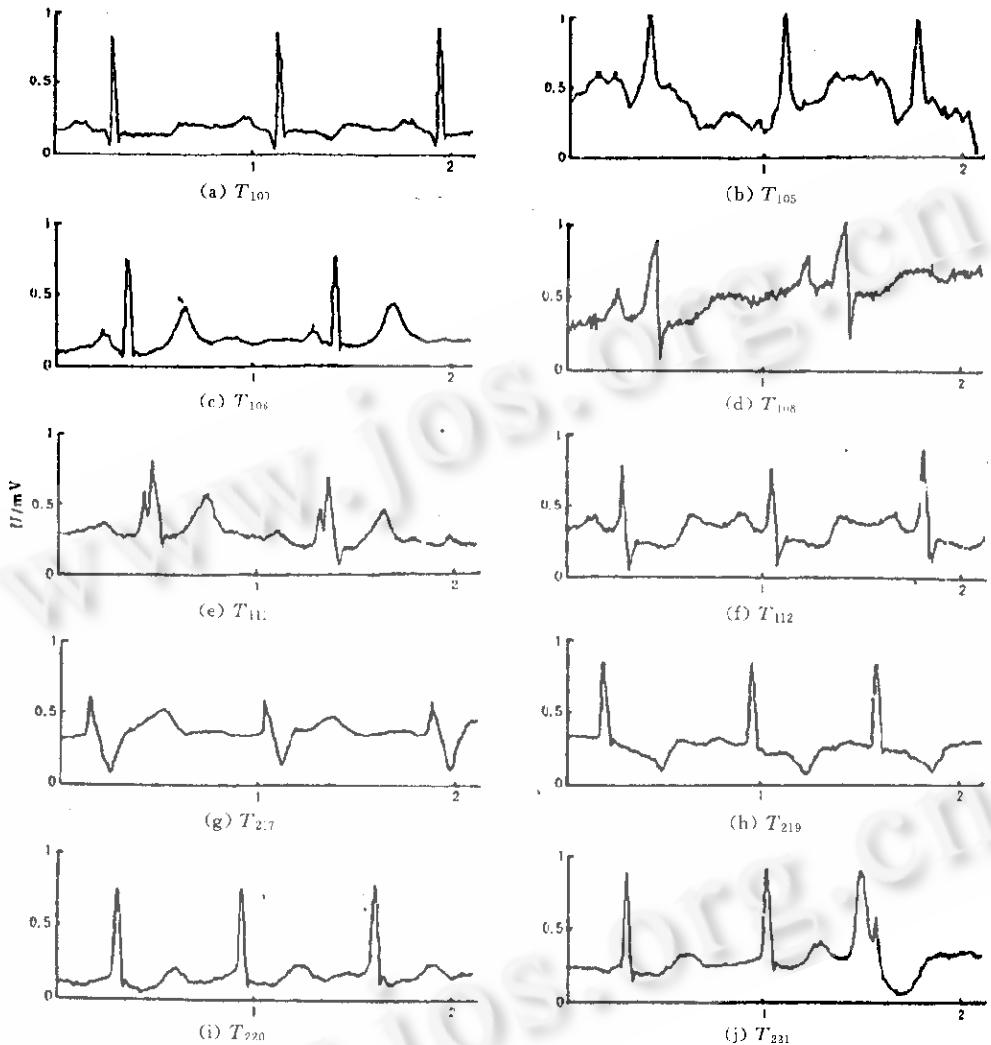


Fig. 4 The waveforms of electro-cardiogram $T_{100} \sim T_{221}$

图 4 心电图 $T_{100} \sim T_{221}$ 部分波形

实验中所有的参数设置如下:心电图的各心搏取 R 波顶点之前 105 点, R 波顶点之后 165 点。对这 270 点数据进行抽用,即在 R 波前后 15 点全部使用,其他部分每隔 6 点采样,其结果为每一心搏取 70 个数据。

心电图分类知识自动获取系统首先由神经网络(其结构为输入层神经元数为 70,隐含层神经元数为 20,输出层神经元数为 10。学习算法为 BP,且引入神经元的自适应学习规则,即神经元的 5 种状态:正常、病态、临态、死亡、睡眠及其状态之间的转换)^[2]从心电图波形 $T_{100} \sim T_{221}$ 中获取心电图分类知识。这些知识以分布的形式存放于神经网络结构中。然后,采用第 2 节中讨论的规则提取算法,从神经网络中提取出符号知识,从中发现,神经网络经过学习后自动形成心电图中的 P 波、QRS 复合波和 T 波分类的功能区,表示了心电图分类的深层知识。最后,根据从神经网络中提取出的符号知识以及由符号机器学习算法从专家和书本上获取的心电图分类知识来优化神经网络的结

构,以获得更好的学习效果。

通过实验发现,该系统对于已学习过的和与已学习过的心电图波形相似的心电图波形,可以由神经网络直接给出心电图波形的正确的分类结果;对于没有学习过的且与已学习过的心电图波形相差较大的波形,该系统可以根据从神经网络提取出的符号知识和符号机器学习算法获取的知识,对心电图波形进行分析、推理而获得正确的结果。表1~表3分别列出了基于认知模拟的机器学习算法用于心电图分类知识获取、神经网络学习算法用于心电图分类知识获取、符号机器学习算法用于心电图分类知识获取系统对心电图进行分类的结果。

Table 1 The experimental results of the adaptive machine learning algorithm based on cognition simulation
表 1 基于认知模拟的自适应机器学习算法的实验结果

Learning pattern data ^①	Knowledge expression form ^②	Learning time ^③	Correct classification rate ^④ (%)		Classification time ^⑤ (μ s)	
			Learned patterns ^⑥	Non-Learned patterns ^⑦	Learned patterns	Non-Learned patterns
Electro-Cardiogram ^⑧ $T_{100} \sim T_{221}$	Neural network structure, connecting weights, rules ^⑨	15 minutes ^⑩ (Pentium 100)	100	98.3	5	20

①学习样本数据,②知识表示形式,③学习时间,④正确识别率,⑤分类时间,⑥学习过的,⑦未学习过的,⑧心电图,⑨神经网络的拓扑结构、联接权矩阵、规则,⑩15分钟。

Table 2 The experimental results of the machine learning algorithm based on BP
表 2 基于BP神经网络的机器学习算法的实验结果

Learning pattern data ^①	Knowledge expression form ^②	Learning time ^③	Correct classification rate ^④ (%)		Classification time ^⑤ (μ s)	
			Learned patterns ^⑥	Non-Learned patterns ^⑦	Learned patterns	Non-Learned patterns
Electro-Cardiogram ^⑧ $T_{100} \sim T_{221}$	Neural network structure, connecting weights ^⑨	15 minutes ^⑩ (Pentium 100)	100	90.1	5	20

①学习样本数据,②知识表示形式,③学习时间,④正确识别率,⑤分类时间,⑥学习过的,⑦未学习过的,⑧心电图,⑨神经网络的拓扑结构、联接权矩阵,⑩15分钟。

Table 3 The experimental results of the machine learning algorithm based on IQ3
表 3 基于IQ3机器学习算法的实验结果

Learning pattern data ^①	Knowledge expression form ^②	Learning time ^③	Correct classification rate ^④ (%)		Classification time ^⑤ (ms)	
			Learned patterns ^⑥	Non-Learned patterns ^⑦	Learned patterns	Non-Learned patterns
Electro-Cardiogram ^⑧ $T_{100} \sim T_{221}$	Decision tree ^⑨	15 minutes ^⑩ (Pentium 100)	100	81.7	100	100

①学习样本数据,②知识表示形式,③学习时间,④正确识别率,⑤分类时间,⑥学习过的,⑦未学习过的,⑧心电图,⑨决策树,⑩15分钟。

从表1~表3的实验结果可以看出,由于基于认知模拟的自适应机器学习算法采用了较为灵活的学习算法(根据问题的特点,选择符号学习算法和联接学习算法)和知识表示方法(联接表示和符号表示融于一体,符号表示由联接表示经规则提取而获取,符号表示的知识又作为联接学习的基础和指导,优化联接学习效果和神经网络结构),所以基于该学习算法实现的心电图分类知识获取系统,无论在学习时间,还是在心电图的正确分类率和分类时间整体性能指标上都优于基于BP神经网络和IQ3的心电图分类知识获取系统。更为重要的是,由此可以进一步地深入探讨人类的思维机制和学习机理。

4 结束语

本文通过研究人脑的自适应学习理论,提出了基于认知模拟的自适应机器学习方法,并通过实

现的心电图分类知识自动获取系统的实验,进一步证实该机器学习方法的先进性和有效性.对于进一步研究人类智能、人工智能和认知科学具有一定的指导意义.

References:

- [1] Wang, Ji-cheng., Lu, Wei-xue. Multiobjective optimization of neural network. *Science in China (Series B)*, 1995,38(8): 971~978.
- [2] Wang, Ji cheng, Cui, Yi-fa, Lu, Wei-xue. A new method on automatically building neural network structure. *Acta Automatica Sinica*, 1996,22(1):19~25 (in Chinese).
- [3] Rumelhart, D. E., McClelland, J. L. *Parallel Distributed Processing (Vol. 1)*. Cambridge, MA: MIT Press. 1986.
- [4] Quinlan, J. R. Induction of decision trees. *Machine Learning*, 1986,40(1):81~103.
- [5] Fahlman, S. E., Hinton, G. E. Connectionist architectures for artificial intelligence. *IEEE Computer*, 1987,33(1):98~195.
- [6] Diederich, J. KRITON: a knowledge acquisition for expert systems. *International Journal of Man-Machine Studies*, 1987, 35(5):26~32.
- [7] Ribaric, S. Knowledge representation scheme based on Petri net theory. *International Journal of Pattern Recognition and Intelligence*, 1988,2(4):691~700.
- [8] Birman, K. P. Rule-Based learning for more accurate ECG analysis. *IEEE Pattern Analysis Intelligence*, 1982,30(4):369~375.
- [9] Xue, Q. Z., Hu, Y. H., Tompkins, W. J. Neural-Network-Based adaptive matched filtering for QRS detection. *IEEE Transactions on Biomedical Engineering*, 1992,39(6):317~329.
- [10] Lin, K. P., Chang, W. H. Classification of QRS pattern by an associative memory model. In: Aderson, J. A., ed. *Proceedings of the Annual International Conference on IEEE Engineering Medical Biological Society*. New York: Academic Press, 1989. 2017~2018.

附中文参考文献:

- [2] 王继成、蔡义发、吕维雪.一种自动生成神经网络结构的新方法. *自动化学报*,1996,22(1):19~25.

Research on the Adaptive Machine Learning Algorithm Based on Cognition Simulation*

WANG Ji-cheng

(Department of Computer Science and Engineering, Tongji University, Shanghai 200092, China)

E mail: jcwangk@online.sh.cn

http://www.tongji.edu.cn

Abstract: In this paper, according to the theory of cognitive science and the research results of neurology and brain science, a kind of adaptive machine learning algorithm is presented by using the methods of computer science, mathematics and engineering science. This algorithm simulates the learning process of man brain from the micro-neuron-level, middle-level and macro-level. The developed automatic electro-cardiogram classification knowledge acquisition system can well simulate the sensation, cognition and thinking of electro-cardiogram experts.

Key words: cognitive science; machine learning; neural network; electro-cardiogram analysis

* Received September 10, 1999; accepted March 28, 2000

Supported by the National Natural Science Foundation of China under Grant No. 69285005