机器学习的主要策略综述

闫友彪, 陈元琰
(广西师范大学 数学与计算机学院, 广西 桂林 541004)

摘 要: 当前人工智能研究的主要障碍和发展方向之一就是机器学习。机器学习与计算机科学、心理学、认知科学等各种学科都有着密切的联系,牵涉的面比较广,许多理论及技术上的问题尚处于研究之中。对机器学习的一些主要策略的基本思想进行了较全面的介绍,同时介绍了一些最新的进展和研究热点。

关键词: 机器学习; 学习策略; 支持向量机; 强化学习; 遗传算法
中图法分类号: TP181 文献标识码: A 文章编号: 1001-3695(2004)07-0004-07

A Survey on Machine Learning and Its Main Strategy

YAN You-biao, CHEN Yuan-yan
(*College of Mathematics & Computer Science, Guangxi Normal University, Guilin Guangxi* 541004, *China*)

Abstract: The main obstacle of the current Artificial Intelligence(AI) study is machine learning, which is one of the developing directions of AI as well. Machine learning has much to do with various disciplines, such as computer science, psychology, and cognitive science etc. It involves too much different fields and many theories and technical problems are still under study. The main strategy of machine learning, as well as other new developments and hot points of the study are discussed.

Key words: Machine Learning; Learning Strategy; Support Vector Machine; Reinforcement Learning; Genetic Algorithms

1 机器学习的基本概念与学习系统

1.1 基本概念

机器学习的核心是学习。关于学习,至今却没有一个精确的、能被公认的定义。这是因为进行这一研究的人们分别来自不同的学科,更重要的是学习是一种多侧面、综合性的心理活动,它与记忆、思维、知觉、感觉等多种心理行为都有着密切的联系,使得人们难以把握学习的机理与实现。

目前在机器学习研究领域影响较大的是 H. Simon 的观点:学习是系统中的任何改进,这种改进使得系统在重复同样的工作或进行类似的工作时,能完成得更好[1]。学习的基本模型就是基于这一观点建立起来的。

机器学习就是要使计算机能模拟人的学习行为,自动地通过学习获取知识和技能,不断改善性能,实现自我完善。机器学习研究的就是如何使机器通过识别和利用现有知识来获取新知识和新技能。作为人工智能的一个重要的研究领域,机器学习的研究工作主要围绕学习机理、学习方法、面向任务这三个基本方面的研究。

1.2 学习系统

为了使计算机系统具有某种程度的学习能力,使它能通过学习增长知识,改善性能,提高智能水平,需要为它建立相应的学习系统。一个学习系统必须具有适当的学习环境,一定的学习能力,并且能应用学到的知识求解问题,其目的是能提高系统的性能。一个学习系统一般应该由环境、学习、知识库、执行与评价四个基本部分组成。各部分的关系如图 1 所示。

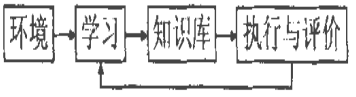

图 1 学习系统的基本结构

在图 1 中,箭头表示信息的流向;环境指外部信息的来源,它将为系统的学习提供有关信息;学习指系统的学习机构,它通过对环境的搜索取得外部信息,然后经过分析、综合、类比、归纳等思维过程获得知识,并将这些知识存入知识库中;知识库用于存储由学习得到的知识,在存储时要进行适当的组织,使它既便于应用又便于维护;执行与评价由执行和评价两个环节组成,执行环节用于处理系统面临的现实问题,即应用学习到的知识求解问题,如定理证明、智能控制、自然语言处理、机器人行动规划等;评价环节用于验证、评价执行环节的效果,如结论的正确性等。另外,从执行到学习必须有反馈信息,学习将根据反馈信息决定是否要从环境中索取进一步的信息进行学习,以修改、完善知识库中的知识。这是学习系统的一个重要特征。

2 机器学习的主要策略

机器学习的发展极为迅速,应用亦日益广泛,有很多优秀的学习算法,基本上可以分为基于符号学习和基于非符号学习(连接学习)。其中符号学习比较好的有机械式学习、指导式学习、示例学习、类比学习、基于解释的学习。

随着人工智能研究的进展,人们逐渐发现研究人工智能的最好方法是向人类自身学习,因而引入了一些模拟进化的方法来解决复杂优化的问题,其中富有代表性的是遗传算法。遗传

收稿日期: 2003-07-06; 修返日期: 2003-09-11

算法的生物基础是人类生理的进化及发展,这种方法被称为进化主义;另一方面,神经网络的理论是基于人脑的结构,其目的是揭示一个系统是如何向环境学习的,此方法被称为连接主义。这两种方法与传统方法大相径庭,因而近年来许多科学家致力于这两种方法的研究。

另外由于统计学习理论的发展,提出了支持向量机的学习算法,由于其出色的学习性能尤其是泛化能力,从而引起了人们对这一领域的极大关注。该技术已成为机器学习界的研究热点,并在很多领域都得到了成功的应用。

在20世纪80年代,基于试错方法、动态规划和瞬时误差方法形成了强化学习(Reinforcement Learning)理论。这是一种不同于传统机器学习理论的学习方法。目前,强化学习理论在智能控制、机器人学、导弹制导及分析预测等领域的研究中有着大量的应用。但在国内,强化学习的研究还处于起步阶段。

2.1　机械式学习

机械式学习(Rote Learning)又称死记式学习,这是一种最简单也是最原始、最基本的学习策略。通过记忆和评价外部环境所提供的信息达到学习的目的,学习系统要做的工作就是把经过评价所获取的知识存储到知识库中,求解问题时就从知识库中检索出相应的知识直接用来求解问题。

当机械式学习系统的执行部分解决完一个问题之后,系统就记住这个问题和它的解。可以把执行部分抽象地看成某一函数,这个函数在得到自变量输入值$(x_1, \cdots, x_n)$之后,计算并输出函数值$(y_1, \cdots, y_p)$。实际上它就是简单的存储联合对$[(x_1, \cdots, x_n), (y_1, \cdots, y_p)]$。在以后遇到求自变量输入值为$(x_1, \cdots, x_n)$的问题的解时,就从存储器中把函数值$(y_1, \cdots, y_p)$直接检索出来而不是进行重新计算。机械式学习过程可用模型示意如下:

(1)学习过程

$$(x_1, \cdots, x_n) \xrightarrow{\text{计算}} (y_1, \cdots, y_p) \xrightarrow{\text{存储}} [(x_1, \cdots, x_n), (y_1, \cdots, y_p)]$$

(2)应用过程

$$(x_1, \cdots, x_n) \xrightarrow{\text{检索}} [(x_1, \cdots, x_n), (y_1, \cdots, y_p)] \xrightarrow{\text{输出}} (y_1, \cdots, y_p)$$

机械式学习是基于记忆和检索的方法,学习方法很简单,但学习系统需要几种能力:①能实现有组织的存储信息;②能进行信息结合;③能控制检索方向。对于机械式学习,需要注意三个重要的问题:存储组织信息、环境的稳定性与存储信息的适用性以及存储与计算之间的权衡。机械式学习的学习程序不具有推理能力,只是将所有的信息存入计算机来增加新知识,其实质上是用存储空间换取处理时间,虽然节省了计算时间,但却多占用存储空间。当因学习而积累的知识逐渐增多时,占用的空间就会越来越大,检索的效率也将随之下降。所以,在机械式学习中要全面权衡时间与空间的关系。

2.2　指导式学习

比机械式学习更复杂一点的学习是指导式学习。指导式学习(Leaming by Being Told)又称嘱咐式学习或教授式学习。在这种学习方式下,由外部环境向系统提供一般性的指示或建议,系统把它们具体地转换为细节知识并送入知识库。在学习过程中要反复对形成的知识进行评价,使其不断完善。

对于使用指导式学习策略的系统来说,外界输入知识的表达方式与内部表达方式不完全一致,系统在接收外部知识时需要一点推理、翻译和转换工作。MYCIN, DENDRAL等专家系统在获取知识上都采用这种学习策略。一般地说,指导式学习系统需要通过如下步骤实现其功能:

(1)请求——征询指导者的指示或建议;

(2)解释——消化吸收指导者的建议并把它转换成内部表示;

(3)实用化——把指导者的指示或建议转换成能够使用的形式;

(4)并入——并入到知识库中;

(5)评价——评价执行部分动作的结果,并将结果反馈到第一步。

指导式学习是一种比较实用的学习方法,可用于专家知识获取。它既可避免由系统自己进行分析、归纳从而产生新知识所带来的困难,又无需领域专家了解系统内部知识表示和组织的细节,因此目前应用得较多。

2.3　归纳学习(Inductive Learning)

归纳学习是应用归纳推理进行学习的一类学习方法,也是研究最广的一种符号学习方法,它表示从例子设想出假设的过程。归纳是指从个别到一般、从部分到整体的一类推论行为。归纳推理是应用归纳方法所进行的推理,即从足够多的事例中归纳出一般性的知识,它是一种从个别到一般的推理。由于在进行归纳时,多数情况下不可能考察全部有关的事例,因而归纳出的结论不能绝对保证它的正确性,只能以某种程度相信它为真,这是归纳推理的一个重要特征。在进行归纳学习时,学习者从所提供的事实或观察到的假设进行归纳推理,获得某个概念。归纳学习也可按其有无教师指导分为示例学习以及观察与发现学习。

2.3.1　示例学习(Learning from Examples)

示例学习(图2)又称概念获取或从例子中学习,它是通过从环境中取得若干与某概念有关的例子,经归纳得出一般性概念的一种学习方法。在这种学习方法中,外部环境(教师)提供的是一组例子(正例和反例),这些例子实际上是一组特殊的知识,每一个例子表达了仅适用于该例子的知识,示例学习就是要从这些特殊知识中归纳出适用于更大范围的一般性知识,它将覆盖所有的正例并排除所有反例。

其学习过程是:①从示例空间(环境)中选择合适的训练示例;②经解释归纳出一般性的知识;③再从示例空间中选择更多的示例对它进行验证,直到得到可实用的知识为止。

在示例学习系统中,有两个重要概念:示例空间和规则空间。示例空间就是我们向系统提供的训练例集合。规则空间是例子空间所潜在的某种事物规律的集合,学习系统应该从大量的训练例中自行总结出这些规律。可以把示例学习看成是选择训练例去指导规则空间的搜索过程,直到搜索出能够准确反映事物本质的规则为止。这就是1974年,Simon和Lea提出的通过示例学习的双空间模型(图3)。

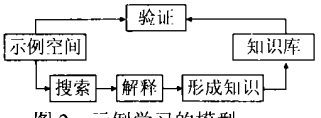

图2　示例学习的模型

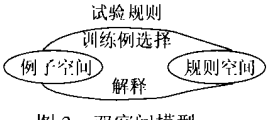

图3　双空间模型

2.3.2 观察与发现学习(Learning from Observation & Discovery)

观察与发现学习分为观察学习与机器发现两种。前者用于对事例进行概念聚类,形成概念描述;后者用于发现规律,产生定律或规则。

概念聚类是观察学习研究中的一个重要技术,基本思想是把事例按一定的方式和准则进行分组,如划分为不同的类、不同的层次等,使不同的组代表不同的概念,并且对每一个分组进行特征概括,得到一个概念的语义符号描述。

机器发现是指从观察的事例或经验中归纳出规律或规则,这是最困难、最富有创造性的一种学习。它可分为经验发现与知识发现两种。前者是指从经验数据中发现规律和定律;后者是指从已观察的事例中发现新的知识。

归纳学习方式是根据一些具体的现象形成并归纳出一些规律,学习结果供其他同类使用。该方式在协助获取专家知识方面起到很好的作用,由于专家多年来积累的经验通常是"隐性知识",甚至只是一种直觉,因此难以表述和提取。但专家经验来源于实践,是对大量实例和现象的归纳。因此,用归纳学习方法来获取专家知识恰到好处,它为解决专家系统的知识获取这个瓶颈问题提供了重要的手段。归纳学习仅通过实例之间的比较来提取共性与不同,难以区分重要的、次要的和不相关的信息,因此常常出现跳步问题。此外,归纳学习要求必须有多个实例,对有些领域来说给出多个实例并非易事,且得出的归纳结论的正确性问题进一步限制了其使用的范围。

2.4 类比学习(Learning by Analogy)

类比是人认识世界的一种重要方法,亦是诱导人们学习新事物、进行创造性思维的重要手段。类比学习(图 4)就是通过类比,即通过对相似事物进行比较所进行的一种学习。

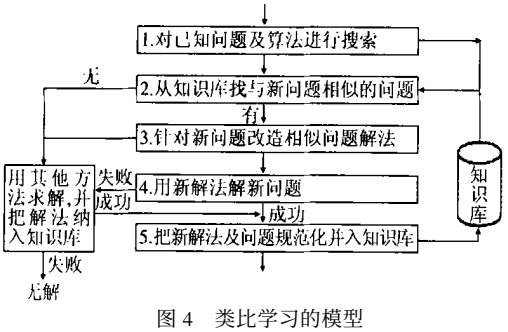

图 4 类比学习的模型

类比学习的基础是类比推理。所谓类比推理,就是指由新情况与记忆中的已知情况在某些方面类似,从而推出它们在其他方面也相似。显然类比推理就是在两个相似域之间进行的:

(1)已经认识的域。它包括过去曾经解决过且与当前问题类似的问题及相关知识,称为源域或者基(类比源),记为 S。

(2)当前尚未完全认识的域。它是遇到的新问题,称为目标域,记为 T。类比推理的目的就是从 S 中选出与当前问题最近似的问题以及求解方法来求解当前的问题,或者建立目标域中已有命题间的联系,形成新知识。

类比学习方法通常有属性类比学习和转换类比学习两种。

于 1986 年,J. G. Carbonell 提出了派生类比学习(Derivational Analogy)系统,它解决多步任务,在传统类比方法中,不记录解法的生成过程,只记录最后生成的算法。而在派生类比系统中,不仅记录最后的解法,而且记录与解法有关的信息,这样系统遇到困难时,可以对某个解法的生成过程进行深层分析,查明原因,根据新的条件,改造新解法。所以,派生类比方法比传统方法具有更强的智能,这种技术也为在专家系统中进行基于事件推理(Case-Based Reasoning)提供了一个有利机制。

但是派生类比方法需要对问题解法生成过程进行分析,因此必然要降低解体效率。同时,这种分析涉及领域知识,增加了问题的复杂性。所以需要在进一步实现中对这种方法进行检验、修改和完善。

当前类比学习模拟的主要困难是基(类比源)的联想,即给定一个目标域,再从无数个错综复杂的结构中找出一个或数个候选的基。在当前实际应用中,基都是由用户给出的,这实际上决定了机器只能重复人们已知的类比,而不能帮助人们学到什么。

2.5 基于解释的学习(Explamation-Based Learning, EBL)

基于解释的学习是通过运用相关领域知识,对当前的实例进行分析,从而构造解释并产生相应知识的一种学习方法。基于解释的学习不是通过归纳或类比进行学习的,而是通过运用相关领域知识及一个训练实例来对某一目标概念进行学习,并最终生成这个目标概念的一般描述。该一般描述是一个可形式化表示的一般性知识。

在进行解释学习时,要向学习系统提供一个实例和完善的领域知识。在分析实例时,首先建立关于该实例是如何满足所学概念定义的一个解释。由这个解释所识别出的实例的特性,被用来作为一般性概念定义的基础;然后通过后继的练习,期待学习系统在练习中能够发现并总结出更一般性的概念和原理。在这个过程中,学习系统必须设法找出实例与练习间的因果关系,并应用实例去处理练习,把结果上升为概念和原理,并存储起来供以后使用。

基于解释的学习方式可以理解成通过一个具体的结果和对它的解释过程,对具体的例子进行普化,从而得到一个普遍的原理。基于解释的学习可以提供更多的东西。这种学习方式好像只记录现有因果链,并把这些因果链重新装入更为直接有用的重聚,而不增加任何新东西。只提供加速作用,因为原则上总能回到原来的实例。

理论修正(Theory Revision)也称知识求精。在基于解释的学习系统中,系统是通过应用领域知识逐步进行演绎,最终构造出训练实例满足目标概念的证明(即解释)的。其中领域知识对证明的形成起着重要的作用,这就要求领域知识是完善的,它足够充分,可以解释被处理的所有例子。但是在现实世界,大多数领域不具备这个特征,不完善是难以避免的。因此,必须研究如何使 EBL 在不完善的领域理论(Imperfect Domain Theory, IDT)中依然有效;同时,还要研究如何使修改不完善的领域理论,使之具有更强的解释能力,而这件事似乎比前者更为重要。在理论修正方面可以修正弥补领域知识的各种缺陷。

需要寻求 EBL 与其他方法的结合,来解决领域知识不完善的问题。目前应用得较好的就是 EBL 与 SBL(Similarity-Based Learning)的结合,它们的结合可以用来比较多个例子的解释以找到公共部分,处理含有噪声和错误的例子。

2.6　基于神经网络的学习

一个连接模型(神经网络)是由一些简单的类似神经元的单元以及单元间带权的连接组成。每个单元具有一个状态,这个状态是由与这个单元相连接的其他单元的输入决定的。连接学习的目的是区分输入的模式的等价类。连接学习通过使用各类例子来训练网络,产生网络的内部表示,并用来识别其他输入例子。学习主要表现在调整网络中的连接权,这种学习是非符号的,并且具有高度并行分布式处理的能力,近年来获得极大的成功与发展。比较出名的网络模型和学习算法有单层感知器(Perceptron)、Hopfield 网络、Boltzmann 机和反向传播算法(Back Propagation, BP)。

人工神经网络是在现代神经科学的基础上提出和发展起来的,旨在反映人脑结构及功能的一种抽象数学模型。一个人工神经网络是由大量神经元节点经广泛互连而组成的复杂网络拓扑,用于模拟人类进行知识和信息表示、存储和计算行为。

人工神经网络学习的工作原理是:一个人工神经网络的工作由学习和使用两个非线性的过程组成。从本质上讲,人工神经网络学习是一种归纳学习,它通过对大量实例的反复运行,经过内部自适应过程不断修改权值分布,将网络稳定在一定的状态下。在神经网络中,大量神经元的互连结构及各连接权值的分布就表示了学习所得到的特定要领和知识,这一点与传统人工智能的符号知识表示法存在很大的不同。在网络的使用过程中,对于特定的输入模式,神经网络通过前向计算,产生一个输出模式,并得到节点代表的逻辑概念,通过对输出信号的比较与分析可以得到特定解。在网络的使用过程中,神经元之间具有一定的冗余性,且允许输入模式偏离学习样本,因此神经网络的计算行为具有良好的并行分布、容错和抗噪能力。

基于神经网络的学习策略主要有两种:刺激-反应论和认识论。

刺激-反应论把自学习解释为习惯的形成。认为经过练习可在某一刺激与个体的某种反应之间建立一种关系,学习就是要建立这样一种关系,即确定神经网络中各个神经元之间的连接权值。成功的学习需要找到一组连接权值,而这组连接权值不能与单元激活值之间的相关性成正比,这由学习规则即最小均方规则(LMS)来进行约束。它利用目标激活值与实际所得的激活值之差进行学习,通过调整连接强度使得这个差减小。当这个差满足预先设定的值,学习过程便结束,学习所得就是神经网络的各个神经元之间的连接权值。

认识论学习策略强调理解在学习过程中的作用,认为学习是个体在其环境中对事物间的关系认识的过程,个体行为取决于其对刺激的知觉与否。构成学习的必要条件是个体对刺激的了解,即个体对符号与符号、符号与目的之间关系的认识,只有对情景有所认识,才能使它的行为变得有目的。

另外,BP 神经网络在非线性控制系统中虽被广泛运用,但作为有导师监督的学习算法,要求批量提供输入/输出对对神经网络训练,而在一些并不知道最优策略的系统中,这样的输入/输出对事先并无法得到。另一方面,强化学习从实际系统学习经验来调整策略,并且是一个逐渐逼近最优策略的过程,学习过程中并不需要导师的监督,因此提出了神经网络与强化学习的结合应用。其基本思想是通过强化学习控制策略,经过一定周期的学习后再用学到的知识训练神经网络,以使网络逐步收敛到最优状态。

自适应谐振理论(Adaptive Resonance Theory, ART)是一类重要的竞争型神经网络学习模型,其记忆模式与生物记忆形式类似,记忆容量可以随学习模式的增加而增加,不仅可以进行实时在线学习,还可以在动态环境下学习,具有较好的性能。

域理论是一种弛豫模型(Relaxation Model),是目前唯一的一种只需进行一遍学习的神经网络模型。其学习速度极快,可以进行实时监督学习,而且记忆容量大,是一种很好的异联想模式分类器。南京大学的周志华等在自适应谐振理论和域理论的基础上提出了一种基于域理论的自适应谐振神经网络算法 FTART/ FTART2,并且对其进行了改进,取得了很好的效果。这些都是值得研究的领域。

神经网络已经在很多领域得到了成功的应用,但由于缺乏严密理论体系的指导,在实际应用中,因为缺乏问题的先验知识,往往需要经过大量费力费时的试验摸索才能确定合适的神经网络模型、算法以及参数设置,其应用效果完全取决于使用者的经验。基于此原因,于 1990 年,Hansen 和 Salamon 开创性地提出了神经网络集成(Neural Network Ensemble)方法。该技术来源于机器学习界目前极热门的 Boosting 方法,也已成为当前研究的热点。

神经网络的另一大缺陷就是其典型的"黑箱性",即训练好的神经网络学到的知识难以被人理解,神经网络集成又加深了这一缺陷。从神经网络中抽取规则来表示其中隐含的知识是解决这个问题的一个有效手段。目前,从神经网络中以及从神经网络集成中抽取规则已成为研究的热点。

2.7　支持向量机(Support Vector Machines, SVMs)

支持向量机是 Vapnik 等人提出的一类新型的机器学习算法。由于其出色的学习性能尤其是泛化能力,从而引起了人们对这一领域的极大关注。该技术已成为机器学习界的研究热点,并在很多领域都得到了成功的应用,如人脸检测、手写数字识别、文本自动分类、机器翻译等。

SVMs 是一种基于统计的学习方法,它是对结构风险最小化归纳原则的近似。它的理论基础是 Vapnik 创建的统计学习理论。统计学习理论研究始于 20 世纪 60 年代末,在其后的 20 年内,涉足这一领域的人不多。SVMs 是统计学习理论中最年轻也最实用的内容,目前,有关这一理论以及应用的研究正在快速发展。不夸张地说,就像信息论为信息技术的崛起开辟道路一样,统计学习理论强带来机器学习领域一场深刻的变革[6]。统计学习理论就是研究小样本统计估计和预测的理论,主要内容包括四个方面:

(1)经验风险最小化准则下统计学习一致性的条件;

(2)在这些条件下关于统计学习方法推广性的界的结论;

(3)在这些界的基础上建立的小样本归纳推理准则;

(4)实现新的准则的实际方法(算法)。

其中,最有指导性的理论结果是推广性的界,与此相关的一个核心概念是 VC 维。

模式识别方法中 VC 维的直观定义是:对一个指示函数

集, 如果存在 h 个样本能够被函数集中的函数按所有可能的 2h 种形式分开, 则称函数集能够把 h 个样本打散; 函数集的 VC 维就是它能打散的最大样本数目 h。若对任意数目的样本都有函数能将它们打散, 则函数集的 VC 维是无穷大。有界实函数的 VC 维可以通过用一定的阈值将它转换成指示函数来定义。VC 维反映了函数集的学习能力, VC 维越大则学习机器越复杂(容量越大)。

支持向量机(SVMs)是从线性可分情况下的最优分类面发展而来的。基本思想可用图 5 的两维情况说明。图 5 中, 实心点和空心点代表两类样本, H 为分类线, H_1, H_2 分别为各类中离分类线最近的样本且平行于分类线的直线, 它们之间的距离叫做分类间隔(Margin)。所谓最优分类线, 就是要求分类线不但能将两类正确分开(训练错误率为 0), 而且使分类间隔最大。分类线方程为 $x\times w+b=0$, 我们可以对它进行归一化, 使得对线性可分的样本集 (x_i, y_i), $i=1, \cdots, n$, $x\in R^d$, $y\in\{+1, -1\}$, 满足

$$y_i[w\cdot x_i+b]-1\geq 0 \quad i=1, \cdots, n \tag{1}$$

此时分类间隔等于 $2/\|w\|$, 使间隔最大等价于使 $\|w\|^2$ 最小。满足条件(1)且使 $\|w\|^2/2$ 最小的分类面就叫做最优分类面, H_1, H_2 上的训练样本点就称作支持向量。使分类间隔最大实际上就是对推广能力的控制, 这是 SVMs 的核心思想之一。

总的来说, 支持向量机就是首先通过用内积函数 $K(x_i, x_j)$ 定义的非线性变换将输入空间变换到一个高维空间, 在这个空间中求(广义)最优分类面。SVMs 分类函数形式上类似于一个神经网络, 输出是中间节点的线性组合, 每个中间节点对应一个支持向量, 如图 6 所示。

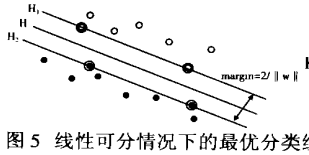

图 5 线性可分情况下的最优分类线

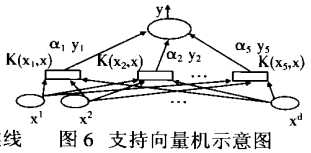

图 6 支持向量机示意图

由于统计学习理论和支持向量机建立了一套较好的有限样本下机器学习的理论框架和通用方法, 既有严格的理论基础, 又能较好地解决小样本、非线性、高维数和局部极小点等实际问题, 因此成为 20 世纪 90 年代末发展最快的研究方向之一, 其核心思想就是学习机器要与有限的训练样本相适应。

统计学习理论虽然已经提出多年, 但从它自身趋向成熟和被广泛重视到现在毕竟才只有几年的时间, 其中还有很多尚未解决或尚未充分解决的问题, 在应用方面的研究更是刚刚开始。

2.8 基于遗传算法的学习

在 60 年代, 美国 Michigan 大学的 Holland 受自然进化论的影响, 注意到学习不仅可以通过单个生物体的适应实现, 而且可以通过一个种群的许多代的进化适应发生。受达尔文进化论思想的影响, 他逐渐认识到在机器学习中, 为获得一个好的算法, 仅靠单个策略的建立和改进是不够的, 还要依赖一个包含许多候选策略的群体的繁殖。考虑到他们的研究想法起源于遗传进化, Holland 就将这个研究领域取名为遗传算法(Genetic Algorithm, GA)。

遗传算法是建立在自然选择和群体遗传学机理基础上的随机迭代和进化, 具有广泛适用性的搜索方法, 具有很强的全局优化搜索能力。它模拟了自然选择和自然遗传过程中发生的繁殖、交配和变异现象, 根据适者生存、优胜劣汰的自然法则, 利用遗传算子选择、交叉和变异逐代产生优选个体(即候选解), 最终搜索到较优的个体。遗传算法本质上是基于自然进化原理提出的一种优化策略, 在求解过程中, 通过最好解的选择和彼此组合, 则可以期望解的集合将会愈来愈好。

遗传算法是一种种群型操作, 该操作以种群中的所有个体为对象。具体求解步骤如下:

(1)参数编码。遗传算法一般不直接处理问题空间的参数, 而是将待优化的参数集进行编码, 一般总是用二进制将参数集编码成由 0 或 1 组成的有限长度的字符串。

(2)初始种群的生成。随机地产生 n 个个体组成一个群体, 该群体代表一些可能解的集合。其任务是从这些群 GA 体出发, 模拟进化过程进行择优汰劣, 最后得出优秀的群体和个体, 满足优化的要求。

(3)适应度函数的设计。遗传算法在运行中基本上不需要外部信息, 只需依据适应度函数来控制种群的更新。根据适应度函数对群体中的每个个体计算其适应度, 为群体进化的选择提供依据。设计适应度函数的主要方法是把问题的目标函数转换成合适的适应度函数。

(4)选择复制。按一定概率从群体中选择 M 对个体, 作为双亲用于繁殖后代, 产生新的个体加入下一代群体。即适应于生存环境的优良个体将有更多繁殖后代的机会, 从而使优良特性得以遗传。选择是遗传算法的关键, 它体现了自然界中适者生存的思想。

(5)杂交(交叉)。对于选中的用于繁殖的每一对个体, 随机地选择同一整数 n, 将双亲的基因码链在此位置相互交换。交叉体现了自然界中信息交换的思想。

(6)变异。按一定的概率从群体中选择若干个个体。对于选中的个体, 随机选择某一位进行取反操作。变异模拟了生物进化过程中的偶然基因突变现象。对产生的新一代群体进行重新评价、选择、杂交和变异。如此循环往复, 使群体中最优个体的适应度和平均适应度不断提高, 直至最优个体的适应度达到某一界限或最优个体的适应度和平均适应度值不再提高, 则迭代过程收敛, 算法结束。GA 的搜索能力主要是由选择和杂交赋予的, 变异算子则保证了算法能搜索到问题解空间的每一点, 从而使算法达到全局最优。

遗传算法受到研究人员广泛重视是由于它采用随机搜索方法, 其特点是几乎不需要所求问题的任何信息而仅需要目标函数的信息, 不受搜索空间是否连续或可微的限制就可找到最优解, 具有强的适应能力和便于并行计算。人们相信随机算法可以解决非线性全局优化问题, 自适应方法可以解决机器学习问题, 并行算法有极高的计算效率。因此, 遗传算法广泛地应用于自动控制、计算科学、模式识别、工程设计、智能故障诊断、管理科学和社会科学领域, 适用于解决复杂的非线性和多维空间寻优问题。与此同时, 经典遗传算法的缺点也显现出来: 有时计算时间过长, 不能保证解是全局最优的。

遗传算法尚存在很多问题, 其原因是 GA 自身的一些缺

陷:①遗传算法没有有效措施来保证所进行的是全局搜索;②变异可消除基因缺陷,但同时会产生新的基因缺陷,因而如何有效地消除基因缺陷又是一个重要的问题;③进化的终止判定,严格地说,遗传算法的迭代是不能完全收敛的,这样终止判定就成了一个亟待解决而又举足轻重的问题。

基于遗传算法和进化计算,提出基于思维进化的机器学习(Mind-Evolution-Based Machine Learning, MEBML)。MEBML 主要由趋同和异化算子构成。趋同策略的优劣直接影响着进化的效率与最优性。

GA 与人工神经网络的融合:模拟脑神经系统的人工神经网络和模拟遗传系统的 GA 既有相同之处又有不同之处。相同的是两者都是关于学习和适应的数学模型,都是基于机体信息处理机制的搜索算法,都具有高度的并行处理功能;不同的是人工神经网络处理的是单个个体的学习,而 GA 则处理的是种群的适应。人工神经网络基本上是局部搜索,即解空间上的现在点的附近是下一个搜索点,而 GA 基本是广域搜索,在解空间上设定了多个搜索点,后续的搜索点由选择、淘汰、交叉和变异来决定。不难看出,两者具有互补的特征,若把两者融合为一体,具有强大的局部搜索能力的人工神经网络可以搜索到更理想的解,而遗传算法的广域搜索能力则可以避免陷入局部最优解。经典遗传算法的计算流程如图 7 所示。

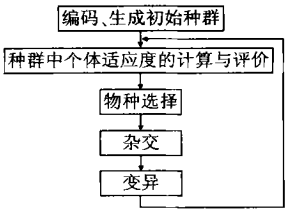

图 7　遗传算法的结构图

2.9　强化学习

在 20 世纪 80 年代,基于试错方法、动态规划和瞬时误差方法形成了强化学习(Reinforcement Learning)理论。1984 年,Sutton 提出了基于 Markov 过程的强化学习;1996 年,Kaelbling 在总结强化学习的研究时指出,实现这种学习的手段就是自适应机制;1998 年,Sutton 和 Barto 将这些研究统称为适应性计算。根据 Simon 的说明,这也是一种学习,但是在机制上,这类机器学习理论不同于人工智能意义下的学习。其主要区别是:这类机器学习强调对变化环境的适应,这意味着,它们需要建立一种基于反馈机制的学习理论。目前,强化学习理论在智能控制、机器人学、导弹制导及分析预测等领域的研究中有着大量的应用。但在国内,强化学习的研究还处于起步阶段。

强化学习,又称奖励学习、评价学习。在传统的机器学习分类中没有提到过强化学习。而在连接主义学习中,把学习算法分为三种类型,即非监督学习(Unsupervised Learning)、监督学习(Supervised Learning)和强化学习。强化学习一词来自于行为心理学,这一理论把行为学习看成是反复试验的过程,从而把环境状态映射成相应的动作。所谓强化学习,就是智能系统从环境到行为映射的学习,以使奖励信号(强化信号)函数值最大。强化学习不同于连接主义学习中的监督学习主要表现在教师信号上。强化学习中由环境提供的强化信号是对产生动作的好坏作一种评价(通常为标量信号),而不是告诉强化学习系统(Reinforcement Learning System, RLS)如何去产生正确的动作。由于外部环境提供的信息很少,RLS 必须靠自身的经历进行学习。通过这种方式,RLS 在行动—评价的环境中获得知识,改进行动方案以适应环境。其基本模型如图 8 所示。

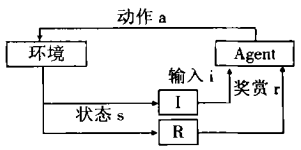

图 8　强化学习的基本模型

强化学习理论是从动物学习、参数扰动自适应控制等理论发展而来。其基本原理可简述如下:如果 Agent 的某个行为策略导致环境正的奖赏(强化信号),那么 Agent 以后产生这个行为策略的趋势便会加强。

强化学习的主要算法有瞬时差分方法(Temporal Difference Method)、Q-学习算法(Q-Learning Algorithm)、自适应启发评价算法(Adaptive Heuristic Critic Algorithm)。另外,还有基于神经网络的强化学习。

目前,强化学习在国际上是十分活跃的研究领域。强化学习的机理比较符合人及生物的学习过程,其思想与 Brooks 提出的行为主义思想是完全一致的。虽然强化学习应用的范围比较广泛,但强化学习比较适合应用于智能控制及智能机器人领域。在智能控制方面,对于具有不确定模型的控制问题,一直是控制理论和控制工程实践中的难题,由于系统具有复杂的非线性和不确定性,使得基于数学模型的传统控制方法难以奏效。随着控制系统复杂程度的增加和控制技术的进展,学习控制正成为一种切实可行的控制手段。强化学习为我们提供了一条有效的途径,采用强化学习方法可以构成一个实时学习控制系统。在智能机器人方面,一方面可以采用强化学习实现智能机器人底层的基础控制;另一方面,也可以采用强化学习实现智能机器人的高层的行为学习,如机器人的路径规划、动作学习等。从国内的研究状况看,强化学习的应用研究还不广泛,尤其是在实际系统中应用得更少,因此,应加大这方面的研究力度。

2.10　多 Agent 学习

随着网络和分布式计算技术的发展,一些现实系统往往异常复杂、庞大,并呈现出分布式特性,以至于单 Agent 因个体所拥有的知识、计算资源和视图的限制而力不能及,因此对多 Agent 的研究迅速发展,逐渐成为人工智能研究的热点。多 Agent 学习已经成为人工智能和机器学习研究方向发展最迅速的领域之一。它和传统体系的显著区别在于其自主性、反应性、协作性。自主性是指在没有人的介入下,它可以持续运行,并能控制自身的动作和内部状态;反应性是指它能感知环境,并采取适当的动作改变环境;协作性是指在多 Agent 环境中协同工作和消解冲突,完成某些互相受益且自身无法独立求解的复杂任务。这也就是说,多 Agent 系统在运行时,无需环境知识的完备性,并且具有自恢复能力。由于环境的未知性,所以多 Agent 的学习过程是不可避免的,采用强化学习作为多 Agent 系统的学习方法。

Agent 是运行于动态环境中具有较高自制能力的实体。目前,由于对 Agent 研究的侧重点不同,Agent 尚没有一个统一和权威的定义。但在有关 Agent 的特性方面,由 Wooldridge 等人提出的 Agent 的“弱定义”和“强定义”最经典并被广为接受。即一个 Agent 最基本的特性应包括:自治性、反应性、社会性、能动性,其行为符合理性要求,这五条判定准则构成了 Agent

的弱定义;而强定义下的 Agent 还应具有如移动性、自适应性、通信能力(包括协作和协调等)等特性以及一些如知识、信念、承诺、意图、义务等人类才具有的特性。Bratma 从哲学上对行为意图进行了研究,提出了关于 Agent 特性的形式表示方法——BDI 理论,即信念(Belief)、愿望(Desire)和意图(Intention),被公认为 Agent 的理论基础之一。

多 Agent 系统(Multi-Agent System, MAS)的概念是由多个智能 Agent 组成的系统。它一般具有个体行为独立自制、个体信息不完全、能力有限、无全局控制、数据分散化和计算异步等特点。MAS 作为解决复杂系统的一个有效方法,能够利用并行分布式处理技术和模块化设计思想,把复杂系统划分成相对独立的 Agent 子系统,通过 Agent 之间的合作与竞争来完成对复杂问题的求解。实际系统中的 Agent 可由不同开发者在不同时间运用不同的工具和技术来实现,因此它们各自具有不同程度的问题求解能力。在 Agent 问题求解过程中,要考虑到其他 Agent 的存在,要与其他 Agent 进行交互,而不是在完全个体封闭的状态下工作。

面向共同目标的问题求解,针对共同的目标任一个 Agent 难以独立完成或多个 Agent 一起完成更有效。有效是指共同完成的效用大于独立完成的效用,如移动一个大物体、演奏交响乐和建房子等通常一个 Agent 难以独立完成,需要明显的多 Agent 合作求解。而像做饭、运输货物等有时一个 Agent 也可以完成,但是效用有时会比多 Agent 合作求解的效用低。

面向不同目标的问题求解,每个 Agent 有各自不同的目标,每个 Agent 难以独立完成自己的目标或通过将部分(或全部)委托给其他 Agent 来完成更有效。

这样多个 Agent 之间通过相互之间的委托形成一种问题求解格式。队工作模型是目前 Agent 合作的研究热点,是一种面向共同目标的特殊问题求解形式。在现实生活中,充满大量队工作的例子,如机器人足球、部队战斗、交响乐演奏等。

多 Agent 技术主要涉及复杂和并发系统的建立与管理、流动访问与管理、信息搜集与处理、语言处理、工业制造、飞行器控制、监控、分布式计算与协同工作、电子商务、用户界面和中间件以及机器人等。

3 机器学习的发展与展望

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元的学习机制、自组织系统和学习算法的研究;70 年代中期以来,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,如示例学习、解释学习、类比学习、概念聚类等;80 年代又提出多层网络的学习算法,从而使机器学习进入了连接学习的研究阶段。但是随着研究的不断深入,不难发现,无论是基于连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,通过其内部连接机制的变化,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。

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(即基于连接机制),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。它涉及到连接理论、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。

传统科学追求简单性、必然性、决定论的目标,而现实世界大多是非线性的、非平衡的系统,复杂自适应理论(CAS)则扬弃了传统科学的目标和范式,探索复杂性、偶然性、非决定论,强调了综合性和整体性的研究范式。人的学习与认识充满了大量的联想和顿悟(突变序),人的社会行为是大量个体彼此交互、自适应的结果,在这个自适应过程中人的学习和认识得到了增强。这就等同于复杂自适应系统理论中的突变和自学习、自适应进化。伽利略晚年致力于研究潮汐运动,他逝世的那一年圣诞,牛顿出生了,苹果掉下来打在牛顿头上,他发现了重力。若是苹果打在计算机上面,它将怎样呢?可见知识的继承与人的联想、顿悟有关,人类在自适应进化中学习,机器学习还是望尘莫及。遗传算法试图模拟生物的遗传进化,神经网络模拟人脑结构、人的思维模型,强化学习则基于反馈机制,多 Agent 系统强调自主、协作与交互,这些理论因为更加接近真实自然,因而取得很大的发展和成功。复杂自适应系统(非平衡自组织理论)把生物学和物理学重新装到一起,把必然性和偶然性重新装到一起,把自然科学和人文科学重新装到一起,研究从混沌到有序。它可能是理解天体、太阳系、地球、生命等"四大起源"的钥匙,也可能是解决各种复杂问题的工具和手段。

4 结束语

本文对机器学习领域中各种方法进行了一次较全面的介绍,同时指出了一些研究的热点。机器学习是一个十分活跃、充满生命力的研究领域,同时也是一个困难的、争议较多的研究领域。在这个领域中,新的思想、方法不断地涌现,研究继续向纵深防线发展,取得了令人瞩目的成就,但是还存在大量未解决的问题。当前人工智能研究的主要障碍和发展方向之一就是机器学习,因此机器学习有着广阔的研究前景。另外,由于机器学习与其他各种学科有着密切的联系,研究者应该从不同的研究环境和领域寻找多种学习体制和方法,同时机器学习的研究也在等待着有关学科的研究取得进展。

从目前的研究趋势来看,估计机器学习今后将在以下几个方面做更多的工作:① 人类学习机制的研究;② 发展和完善现有的学习方法,并开展新的学习方法的研究;③ 建立实用的学习系统,特别是多种学习方法协同工作的集成化系统的研究;④ 机器学习有关理论及应用的研究。 (下转第 13 页)

的时间内(如 NTP 多播)。这种方法只需要把核心模块放到 Honeypot 上, 修改网关上的嗅探器就可以了。

4 Honeynets 的最新技术及发展

Honeynets 工程组开发了β 码来实现远程移动数据的性能。Sebek 程序分为以下两部分:

(1)核心模块。它要先配置(如配置当 NetBios 分组送到电缆上时源 MAC 和 IP 要伪装成多少等), 然后安装到 Honeypot 中。

(2)一个特殊的嗅探器。它要安装在 Honeynets 网关上, 以积极地捕获、解密和重建攻击者的行为。

工程组研究的另一个领域是虚拟(Virtual)Honeynets[9,10]的应用。虚拟 Honeynets 支持 GenⅠ和 GenⅡ两种技术, 把 Honeynets 配置的三个必要条件: 数据控制、数据捕获与数据收集以及实际的 Honeypot 本身全部合成为一个物理的系统。Honeypot 上装有真实的操作系统, 没有任何仿真东西。优点就是价值和效率, 也更容易配置和维护。

5 结论

Honeynets 只是一种用来收集黑客信息的 Honeypot, 主要用于研究。Honeynets 的管理工作非常复杂, 因此管理者要有责任保证一旦 Honeynets 被攻破, 不能被用作跳板去攻击其他任何系统。没有正确的管理, 使用的风险可能超过收益。当然这个工具不是安全上的万能, 也不可能是每个单位的合适的安全方案。网络安全是一个复杂的课题, 首先要保证自己的计算机和网络是安全的, 才能用 Honeynets 作为一种工具来主动向攻击者和自己学习更多的知识。同时, 在配置 Honeynets 之前也应向自己的法律顾问了解一些存在的法律问题。

网络入侵诱骗系统在不断发展, Honeynets 作为一种新的比较完善的研究黑客的工具必将得到充分的发展和使用。

参考文献:

[1] 夏春和, 吴震, 赵勇, 等. 入侵诱骗模型的研究与建立[J]. 计算机应用研究, 2002, 19(4): 76-79.

[2] Lance Spitzner. Definitions and Value of Honeypots[EB/OL]. http://www.enteract.com/~lspitz, 2002-05-17.

[3] Loras R Even. What is a Honeypot[EB/OL]. http://project.honeynets.org, 2000-07-12.

[4] Honeynets Project. Know Your Enemy: GenⅡ Honeynets[EB/OL]. http://www.honeypot.org, 2003-04-12.

[5] Honeynets Project. Know Your Enemy: Honeynets[EB/OL]. http://project.honeynets.org/papers/honeynets, 2002-09-08.

[6] David Klug. Honeypots and Intrusion Deception[EB/OL]. http://www.org/infosecFAQ/honeypots.htm, 2000-09-13.

[7] William W Martin, et al. Honeypot and Honeynets-Security through Deception[EB/OL]. http://project.honeynets.org/, 2001-05-25.

[8] Michael Sink. The Use of Honeypots and Packet Sniffers for Intrusion[EB/OL]. http://rr.sans.org/intrusion/honey_pack.php/, Detection, 2001-04-15.

[9] Michael Clark. Virtual Honeynets[EB/OL]. http://online.securityfocus.com/infocus/1506/, 2001-11-07.

[10] Honeynets Project. Know Your Enemy: Defining Virtual Honeynets [EB/OL]. http://project.honeynets.org/papers/honeynets/, 2002-09-08.

作者简介:

马传龙(1973-), 男, 山东潍坊人, 讲师, 硕士研究生, 研究方向为计算机网络与通信; 邓亚平(1948-), 男, 重庆铜梁人, 教授, 硕士生导师, 研究方向为计算机网络、网络安全。

(上接第 10 页)

参考文献:

[1] 杨炳儒. 知识工程与知识发现[M]. 北京: 冶金工业出版社, 2000.

[2] 王永庆. 人工智能原理与方法[M]. 西安: 西安交通大学出版社, 1998.

[3] 史忠植. 知识发现[M]. 北京: 清华大学出版社, 2002.

[4] 周志华, 等. 神经网络集成[J]. 计算机学报, 2002, 25(1): 1-8.

[5] 孙晨, 周志华, 陈兆乾. 神经网络规则抽取研究[J]. 计算机应用研究, 2000, 17(2): 34-37.

[6] 王国胜, 钟义信. 支持向量机的理论基础—统计学习理论[J]. 计算机工程与应用, 2001, (19): 19-20, 31.

[7] 王国胜, 钟义信. 支持向量机的若干新进展[J]. 电子学报, 2001, (10): 1397-1400.

[8] 崔伟东, 周志华, 李星. 支持向量机研究[J]. 计算机工程与应用, 2001, (1): 58-61.

[9] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.

[10] 张汝波, 顾国昌, 刘照德, 等. 强化学习理论、算法及应用[J]. 控制理论与应用, 2000, 17(5): 637-642.

[11] 陆鑫, 高阳, 李宁, 等. 基于神经网络的强化学习算法研究[J]. 计算机研究与发展, 2002, 39(8): 981-985.

[12] 周志华, 陈兆乾, 陈世福. 基于域理论的自适应谐振神经网络分类器[J]. 软件学报, 2000, 11(5): 667-672.

[13] 孙承意, 谢克明, 程明琦. 基于思维进化机器学习的框架及新进展[J]. 太原理工大学学报, 1999, 30(5): 453-457.

[14] 王继成. 基于认知模拟的自适应机器学习算法研究[J]. 软件学报, 2001, 12(8): 1205-1211.

[15] 刘新宇, 洪炳[illegible] 基于 BDI 框架的多 Agent 动态协作模型与应用研究[J]. 计算机研究与发展, 2002, 39(7): 798-801.

[16] 黄敏, 佟振声. 分布式多 Agent 系统的研究[J]. 电力情报, 2002, (1): 65-70.

[17] 董红斌, 孙羽. 多 Agent 系统的现状和进展[J]. 计算机应用研究, 2001, 18(1): 54-56.

[18] 许国志, 顾基发, 车宏安. 系统科学[M]. 上海: 上海科技出版社, 2000.

[19] James A Highsmith. Adaptive Software Development[M]. 北京: 清华大学出版社, 2003.

作者简介:

闫友彪(1974-), 男, 湖南常德人, 硕士, 主要研究方向为机器学习、计算机网络、计算机图形学; 陈元琰(1961-), 男, 福建仙游人, 副教授, 博士, 主要研究方向为计算机图形学、计算机网络。