381

O15
1137

河北省自然科学基金
河 北 省 博 士 资 金　资助项目

模糊示例学习
与模糊控制

王熙照　哈明虎　编著

河北大学出版社

责任编辑:韩　勇
封面设计:赵　谦
责任印制:闻　利

图书在版编目(CIP)数据

模糊示例学习与模糊控制/王熙照,哈明虎编著.保定:河北大学出版社,2002.5
ISBN 7-81028-829-6

Ⅰ.模… Ⅱ.①王…②哈… Ⅲ.①模糊数学②模糊控制 Ⅳ.①O159②TP13

中国版本图书馆 CIP 数据核字(2002)第 026861 号

出版:河北大学出版社(保定市合作路 88 号)　经销:全国新华书店
印制:徐水县印刷厂　规格:1/32(880mm×1230mm)
印张:8.75　字数:210 千字　印数:0001～1000 册
版次:2002 年 5 月第 1 版　印次:2002 年 5 月第 1 次

定价:16.00 元

序　　言

示例学习(Learning from examples)是一种以归纳推理为基础的学习,它是从多个示例中归纳出一般性概念或一般规律的学习方式,故示例学习又称概念获取,它已被公认为是专家系统发展的瓶颈。示例学习过程对领域知识要求较少,但须有多个实例,是一种从特殊到一般的学习过程,所以又称为“一般化”学习方法。示例学习的最高形式称为发现学习,它通过学习可高度概括出一般性的“定理”和“定律”。示例学习是目前机器学习中最重要的分支,属机器学习的核心领域。

传统的示例学习中,所考虑的示例空间都有以下几个特点:①每一示例的取值是确定的;②每一示例的分类情况也是确定的;③如果出现两个示例的取值相同但分类不同的情况(称为噪音或歧异值)应进行过滤处理;④一般情况下若例子的属性值出现丢失现象则无法参与训练;⑤归纳出的规则是确定的清晰规则。这些特点体现出了传统示例学习具有的共同特征,那就是所处理的对象(示例)必须能够精确描述。

随着人的思维、感觉中的不精确性出现在学习过程中以及示例学习研究的深入,具有精确描述特征的示例学习已不能适应一个系统中不精确知识自动获取的要求,研究不确定环境中的示例学习已非常必要。模糊集合理论为这种研究提供了有效的工具。

模糊集理论是研究和处理模糊性现象的理论。所谓模糊性主要指客观事物的差异在中间过渡时所呈现的“亦此亦彼”性。对立的事物似乎是“非此即彼”的,但绝对的突变是不存在的。在自然和社会现象中,差异往往要通过一个中介过渡的形式。处于中介过渡的差异便具有“亦此亦彼”的性质。例如高个子与矮个子、稳定与不稳定、

健康与不健康等这样一些对立的概念之间都没有绝对的分明的界限。这种中介过渡性造就出划分上的不确定性即为模糊性。一个带有模糊性的概念即模糊概念。模糊概念不能用普通集合论来刻画，从而产生了模糊集合论，它是由美国人 L. A. Zadeh 于 1965 年创立的。Zadeh 利用隶属程度来描述差异的中介过渡，它是用精确的数学语言对模糊性的一种描述。例如，“年轻人”、“中年人”、“老年人”是三个模糊概念，一个 30 岁的人应属于哪一种呢？Zadeh 提出使用 0 与 1 之间的数表示该成员属于某模糊集的程度，比如可认为 30 岁的人属于“年轻人”的程度为 0.8，“中年人”为 0.2，而“老年人”为 0。

对绝大多数系统来说，其最重要的信息来源有两种：提供测量数据的传感器、记录器等和提供系统性能描述的专家。我们称来自传感器的信息为数据信息，来自专家的信息为语言信息。数据信息通常可以用数字(如 0.12，0.8 等)表示，而语言信息则用文字或称模糊术语(如大、小、很大等)来表示。数字信息在传统的学习过程中已被广泛使用，但语言信息通常是作为符号值处理的，这种处理过程中并没有将这些术语作为一种模糊子集而是作为一种起替代意义的符号。如何在一个学习系统中有效地使用这些术语的模糊特征自然是一个重要的研究课题。

模糊集合理论已广泛应用于计算机科学、自动控制、地震工程、系统工程、土木工程、环境保护、机械、管理科学、思维科学、社会科学、医疗卫生、气象预报以及文学艺术、体育、心理等领域，但模糊技术最成功的应用领域应该说是模糊控制。特别是日本率先将模糊技术大量用到工业界和家用电器上如模糊洗衣机、模糊空调机、模糊控制的地铁运行系统等等，由于其良好的性能，在国际上受到广泛关注。另外特别需要指出的是由于智能作为人脑思维的模拟而带有明显的模糊性，模糊理论的应用正逐步渗入人工智能的各研究领域，并已取得了很大进展。国际模糊系统协会曾创立了《Fuzzy Sets and Systems》杂志，IEEE 每两年召开一次模糊理论及应用的专门会议，在其大部分会刊杂志中，都刊登关于模糊应用的成果。

通过学习产生模糊规则是模糊控制领域和模糊专家系统发展的瓶颈。产生模糊规则的方式早期是由专家凭经验给出的。近十余年人们开始研究如何从训练数据中自动抽取模糊规则，这些研究曾仅限于对数字信息，它们仅能够对精确描述出的示例集抽取模糊规则。科学的深化意味着研究对象的复杂化，而复杂的对象又难于精确化，当考虑的示例集不能精确描述时传统的学习方法就显得无能为力。本书正是为了弥补这种缺陷而写的。它讨论了不能精确描述的示例集上模糊规则的抽取问题，即示例中带有不确定性的学习问题。

一般认为，不确定性有两大类，一类是源于统计的随机性，另一类是源于人认识、思维、感觉、推理等的模糊性。随机性在示例学习中的处理早在80年代就开始了，例如Quinlan引入的概率决策树，Valiant引入的可学习模型等都使用了处理随机性的概率统计思想。Quinlan曾指出："归纳学习(决策树)的分类结果是分明的，它不能处理分类过程中潜在的不确定性(随机性)。当属性的取值有微小变化时，有可能导致分类结果明显不合适的突变。生成的决策树一般不具有稳健性，数据信息的不精确或缺少可能完全阻止了示例的分类"。Quinlan曾建议使用概率决策树，属性值的不精确性被认为是一种噪音，分支的阈值被软化，分类结果以一种概率分布的形式出现。

在示例学习中关于模糊性的研究是近十余年才开始的。模糊性不同于随机性是由于事物本身不能精确描述造成的。对它研究的意义至少体现在如下几个方面：

(1)示例的属性取值及分类值不能精确描述。由于模糊性与随机性本质上的差异，已无法使用处理随机性的传统的示例学习系统进行处理，而这种示例在模糊知识库中是一种最基本的形式，许多现实系统中信息的短缺及不精确均会导致这种形式。处于这种不精确知识自动获取的需要，进行其学习算法及理论的研究有着实际意义。

(2)现实数据的复杂性也导致不精确性的增加，噪音数据(包括两个示例的属性取值完全相同但分类不同这种情况)在现实数据中

是经常出现的。已有专门处理这种情况的几种噪音模型出现研究相关问题，采用的观点是一致的即依据概率统计模型，将噪音视为一种随机因素等。从模糊性的角度，噪音可得到另外形式的描述，并形成相应的处理方法。鉴于处理噪音在实际学习系统中的重要性，模糊化方法至少提供了一种行之有效的解决途径。

(3)模糊规则一般也是以"IF…THEN…"的形式出现的，但它的条件与结论都不是精确、清晰的概念而是模糊集。采用模糊学习产生的规则一般都具有这种形式，它一般不能认为100%准确，而带有一定的可信度。在匹配过程中，同一个示例可以与多个这样的模糊规则相匹配，自然得到了多个结论，这些结论可能是相近的也可能是相反的，这在精确的示例学习模糊中是不可能的。从决策的角度，它可以向决策者提供多种决策方案，从而有着较大的选择余地，模糊学习能对规则起一种软化作用使其更适合人们的习惯，故对它的研究在这种意义上也是很必要的。

(4)连续值属性的学习问题一直是示例学习中的一个重要部分。传统的示例学习系统对连续值属性的处理一般采用离散化的方式，即将属性按其值域分为几个区间作为几个符号以转化为符号值示例学习问题，考虑到离散过程分点选取的困难(实际上分点选取带有很大的模糊性即没有分明的边界)，模糊化连续值属性为解决连续值属性的学习问题又提供了一种有效途径。

(5)实际应用领域如模糊逻辑系统、模糊控制、模糊模式识别、图像处理等不精确知识获取的需要。

模糊逻辑系统是指那些与模糊概念和模糊逻辑直接有关的系统，以往的文献资料中常见的有两类：纯模糊逻辑系统和具有模糊产生器和模糊消除器的逻辑系统。其中的模糊规则库由若干"如果-则"规则构成，模糊推理机在模糊逻辑原则的基础上来决定如何将输入模糊集与输出模糊集的对应。纯模糊逻辑系统是模糊逻辑系统的核心部分，它提供了一种量化专家语言信息和在模糊逻辑原则下系统地利用这类信息的一般化模式。然而，纯模糊逻辑系统有一个缺

陷即输入和输出均为模糊集合，而大多数工程系统中输入和输出变量为连续真值，为克服这一缺点出现了具有模糊产生器和模糊消除器的逻辑系统。

模糊示例学习理论和算法作为经典归纳学习的扩充与完善，使得示例学习的应用范围扩大到了能处理不确定性。某些传统的归纳学习算法将作为这里的一种特殊形式。模糊示例学习的归纳结果与传统的比较由于合理地处理了不精确信息、噪音数据等问题从而有着更强的分类能力及稳健性，使得知识表示的方式更为自然。由于能生成不同水平和不同置信度的推理规则，故而提供了一种构造专家系统的有效途径，并可为决策者提供丰富的决策信息。

由于作者学识和水平所限，书中不妥甚至错误之处在所难免，敬请同仁及读者批评指正。此外，研究生杨兰珍，王瑞省，牛奔，杨宏伟，孙娟，王金凤，赵明华，何强，邢红杰参加了书稿的整理、校对工作，在此一并致谢。

作　者

2002 年 4 月

目　录

第 1 章 模糊集合

自从美国加州贝克莱大学的 L. A. Zadeh 教授于 1965 年发表“Fuzzy Sets”一文之后,模糊集合论就成为一门新的学科诞生了。在短短的三十多年中,模糊集合的理论和应用都得到了飞速的发展,愈来愈引起人们的高度重视。模糊数学的发展进一步丰富了经典数学的理论,为人们处理模糊信息提供了众多巧妙方法,现已广泛应用于计算机科学、人工智能、信息处理等领域,特别是模糊技术的产业化,模糊产品的问世,取得了显著的经济效益,充分显示出模糊集合论在知识经济时代具有强大的生命力。本章参考文献[1-6]介绍了模糊集的一些基本概念,旨在为后面几章的模糊示例学习技术提供基础。

1.1 引言

经典集合是现代数学最基本的概念之一,它不仅自身已成为一门独立的数学分支,而且集合的概念已广泛渗透到数学的各个领域。

下面考察几组对象:

(1) 所有不大于 5 的自然数;

(2) 所有身高 1.65 以上的人;

(3) 某地区所有的食品商店;

(1′)10 以内所有近似于 5 的自然数;

(2′)所有高个子的人;

(3′)某地区所有效益好的食品商店。

比较上述几组对象,前三组是经典集合。集合中的元素是确定的,即元素与集合之间的关系只有“属于”或“不属于”两种,二者必居其一,而且只居其一。而对于后三组,显然并非是经典集合。哪些自

然数属于“近似于5的自然数”;哪些人属于“高个子的人”;哪些商店属于“效益好的食品商店”,这些都难以说清楚,即元素与集合之间的关系不再是“属于”或“不属于”这种绝对情形。究其原因,关键在于像“近似于5”、“高个子”、以及“效益好”这些概念,并非是清晰明确的。诸如上述这些概念,我们称之为“模糊概念”;这些模糊概念所反映的现象,称之为“模糊现象”。

一般说来,模糊集合是对模糊现象或模糊概念的刻划。所谓模糊现象就是没有严格的界限划分而使得难用精确的尺度刻划的现象,而反映模糊现象的概念称之为模糊概念[1]。例如“高个子的人”,我们既不能认为身高1.65的人是“高个子”,也不能认为身高1.95的人是“高个子”,因为“高个子”是没有严格界限划分的,因而很难只根据身高(即精确的尺度)断然认定某个人是或不是“高个子”。

事实上,模糊现象(模糊概念)处处存在于我们的周围。例如,“拂晓”,天气“很冷”,空中降着“大雪”,行人穿着“厚棉衣”,商场人“非常少”,“老人”坐在“温暖”的房间里,“青年人”喝着“热茶”等等。其中“拂晓”、“很冷”、“大雪”、“厚棉衣”、“非常少”、“老人”、“温暖”、“青年人”、“热茶”,都是一些模糊现象(模糊概念)。经典数学很难对上述这些模糊现象(模糊概念)进行处理,但是1965年,美国加利福尼亚大学自动控制教授扎德(L. A. zadeh)首次提出了“模糊集合”,从而实现了对模糊现象(模糊概念)进行有效的定量描述,标志着模糊数学的诞生。

模糊数学是以模糊集合为基础,对模糊现象(模糊概念)进行研究和处理的一门学科。目前其应用已涉及到国民经济的各个领域及部门,并且在经济管理、自动控制、系统分析、知识描述、图象识别、医学诊断等不确定、非线性系统决策方面有了明显的实际效果。例如,现代人工智能研究的一个领域就是如何使机器像人脑那样分析和处理模糊性问题:

不妨举一个形象的例子。一天,我们在路上行走,远处走来一个

人。尽管与那人相距一段较远的距离，一般情形下，我们仍然可以判断此人是否为自己熟悉的人。其实，我们对此人只掌握一些模糊信息，例如走路姿势、速度（快慢），年龄大小（老或年轻），体型（瘦弱、中等、肥胖）等等。这说明人脑具有分析和处理模糊信息的能力。模糊数学的诞生，搭起了人与机器之间的桥梁，使机器的智能化成了一种可能。

模糊数学是具有强大生命力的学科，相信其在 21 世纪会有更大的进展。

1.2 模糊集的基本概念

模糊集合是对模糊现象或模糊概念的刻划，那末它是如何刻划呢？—— 这是本节研究的一个重要问题。

我们知道经典集合有多种表达方式，例如，列举法、描述法等等。其中特征函数法是其中的一种表达方式。

设 A 为论域 U 中的一个子集，即 $A \subseteq U, u \in U$。则集合 A 的特征函数可表示为

$$\chi_A(u) = \begin{cases} 1 & u \in A \\ 0 & u \notin A \end{cases}$$

由于元素 u 与集合 A 的关系只有 $u \in A$、$u \notin A$ 两种情形，则特征函数取值只为 0 与 1，即值域为 $\{0,1\}$。另外，特征函数有下述三个运算性质：

(1) $\chi_{\overline{A}}(x) = 1 - \chi_A(x)$

其中是 $\overline{\mathbf{A}}$ 是 A 的补集。

(2) $\chi_{A \cup B}(x) = \max(\chi_A(x), \chi_B(x))$

(3) $\chi_{A \cap B}(x) = \min(\chi_A(x), \chi_B(x))$

仿照用特征函数表示经典集合的方法，扎德提出用隶属函数表示模糊集合，即把特征函数的值域由 $\{0,1\}$ 扩大到 $[0,1]$。定义如下：

定义 1.1 给定论域 U，$\mathbf{A}$ 是论域 U 到 $[0,1]$ 的一个映射，即

$$
\begin{aligned}
\boldsymbol{A}: U &\longrightarrow [0,1] \\
u &\longmapsto \boldsymbol{A}(u)
\end{aligned}
$$

则称 $\boldsymbol{A}$ 是 U 上的模糊集，$\boldsymbol{A}(\cdot)$ 称为模糊集 $\boldsymbol{A}$ 的隶属函数，$\boldsymbol{A}(u)$ 称为 u 对 $\boldsymbol{A}$ 的隶属度。在不混淆的情况下，模糊子集也称为模糊集合。

此后为讨论方便，经典集合用 $A,B,C,\cdots$ 等表示；而模糊集合用 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\cdots$ 等表示。

例 1.1　设 $\boldsymbol{A}$ 表示"青年人集合"，其隶属函数如下

$$
\boldsymbol{A}(x) = \begin{cases} 1 & 0 \leqslant x \leqslant 25 \\ [1 + (\frac{x-25}{5})^2]^{-1} & 25 \leqslant x \leqslant 100 \end{cases}
$$

相应曲线如图 1.1

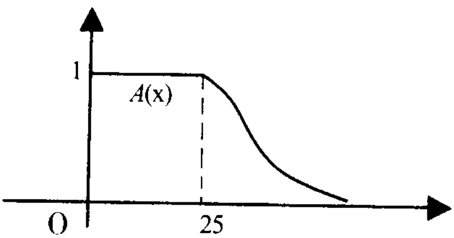

图 1-1　青年人的隶属函数

按上式计算，$\boldsymbol{A}(20) = 0.5, \boldsymbol{A}(35) = 0.2$，故可认为 20 岁的人属于青年人的程度为 0.5，35 岁的人属于青年人的程度为 0.2。显然这种表述方式比经典集合更合乎实际。

注意：1）$\boldsymbol{A}(u)$ 的值越接近于 1，表示 u 属于 $\boldsymbol{A}$ 的程度越高；当 $\boldsymbol{A}(u) = 1$ 时，表示 u 完全属于 $\boldsymbol{A}$。$\boldsymbol{A}(u)$ 的值越接近于 0，表示 u 属于 $\boldsymbol{A}$ 的程度越低；当 $\boldsymbol{A}(u) = 0$ 时，表示 u 完全不属于 $\boldsymbol{A}$。若对于任一 $u \in U$ 都有 $\boldsymbol{A}(u) = 1$ 或 $\boldsymbol{A}(u) = 0$，则模糊集合 $\boldsymbol{A}$ 退化为清晰集。由此可表明经典集合是模糊集合的特例，模糊集合是经典集合的

推广。

2) 模糊集合完全由其隶属函数来描述。正由于隶属函数的引入,才使我们有可能利用精确的数学方法去分析和处理模糊信息。

模糊集合通常的表达方式有以下几种。

当论域 U 为有限集合时,即 $U=\{u_1,u_2,\cdots u_n\}$,有如下三种方式

(1) 扎德表示法,即

$$\boldsymbol{A}=\frac{\boldsymbol{A}(u_1)}{u_1}+\frac{\boldsymbol{A}(u_2)}{u_2}+\cdots+\frac{\boldsymbol{A}(u_n)}{u_n}$$

其中右端并非分式求和,只是一个记号。$\frac{\boldsymbol{A}(u_i)}{u_i}$ 表示论域 U 中的元素 u_i 与其隶属度 $\boldsymbol{A}(u_i)$ 之间的对应关系。此外隶属度为 0 的项可以略去不写。

例 1.2 设论域 $U=\{1,2,3,...,10\}$。讨论"大"、"小"这两个模糊概念。

根据经验,我们定量地给出它们的隶属函数,模糊集"大"、"小"分别表示为

$$\boldsymbol{A}=\frac{0.2}{4}+\frac{0.4}{5}+\frac{0.5}{6}+\frac{0.7}{7}+\frac{0.9}{8}+\frac{1}{9}+\frac{1}{10}$$

$$\boldsymbol{B}=\frac{1}{1}+\frac{0.9}{2}+\frac{0.6}{3}+\frac{0.4}{4}+\frac{0.2}{5}+\frac{0.1}{6}$$

由上式可知 $\boldsymbol{A}(4)=0.2$, $\boldsymbol{B}(4)=0.4$,这说明 4 属于"大"的程度小于 4 属于"小"的程度。但不能说 4 属于 $\boldsymbol{B}$,不属于 $\boldsymbol{A}$。

(2) 序偶表示法

$$\boldsymbol{A}=\{(u_1,\boldsymbol{A}(u_1)),(u_2,\boldsymbol{A}(u_2)),\cdots(u_n,\boldsymbol{A}(u_n))\}$$

采用此法,例 1.2 中的 $\boldsymbol{A}$、$\boldsymbol{B}$ 可分别写成

$$\boldsymbol{A}=\{(4,0.2),(5,0.4),(6,0.5),(7,0.7),(8,0.9),(9,1),(10,1)\}$$

$$\boldsymbol{B}=\{(1,1),(2,0.9),(3,0.6),(4,0.4),(5,0.2),(6,0.1)\}$$

(3) 向量表示法

$$\boldsymbol{A} = (\boldsymbol{A}(u_1), \boldsymbol{A}(u_2), \cdots \boldsymbol{A}(u_n))$$

注意,此时隶属度为 0 的项不能略去。

例 1.2 中的 $\boldsymbol{A}$、$\boldsymbol{B}$ 可表示如下

$$\boldsymbol{A} = \{0,0,0,0.2,0.4,0.5,0.7,0.9,1,1\}$$

$$\boldsymbol{B} = \{1,0.9,0.6,0.4, 0.2,0.1,0,0,0,0\}$$

有时也将上述三种方式综合起来表示。

$$\boldsymbol{A} = (\boldsymbol{A}(u_1)/u_1, \boldsymbol{A}(u_2)/u_2, \cdots, \boldsymbol{A}(u_n)/u_n)$$

当论域 U 为无限集时,扎德给出其表达方式如下

$$\boldsymbol{A} = \int \frac{\boldsymbol{A}(u)}{u}$$

这里的"$\int$"只是一种记号,表示无限多元素合并的一种缩写。例如"青年人"此模糊集合 $\boldsymbol{A}$ 可表示为

$$\boldsymbol{A} = \int_0^{25} 1/x + \int_{25}^{100} [1 + (\frac{x-25}{5})^2]^{-1}/x$$

今后,论域 U 上的全体子集所构成的集合记为 $P(U)$;全体模糊子集所构成的集合记为 $F(U)$。显然 $P(U) \subseteq F(U)$。

1.3 模糊集合的运算

类似于经典集合论,这一节我们来研究模糊集合之间的序关系及其运算。

定义 1.2 设 $\boldsymbol{A}$、$\boldsymbol{B} \in F(U)$。若对任一 $u \in U$,都有 $\boldsymbol{A}(u) \leqslant \boldsymbol{B}(u)$,则称 $\boldsymbol{B}$ 包含 $\boldsymbol{A}$,或称 $\boldsymbol{A}$ 是 $\boldsymbol{B}$ 的子集,记作 $\boldsymbol{A} \subseteq \boldsymbol{B}$;若对任一 $u \in U$,都有 $\boldsymbol{A}(u) = \boldsymbol{B}(u)$,则称 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相等,记作 $\boldsymbol{A} = \boldsymbol{B}$。

根据定义 1.2 可知,模糊幂集 $F(U)$ 上的包含关系具有如下性质:

(1) $\boldsymbol{A} \subseteq \boldsymbol{A}$(自反性);

(2) 若 $\boldsymbol{A} \subseteq \boldsymbol{B}$ 且 $\boldsymbol{B} \subseteq \boldsymbol{A}$,则 $\boldsymbol{A} = \boldsymbol{B}$(反对称性);