

第 3 章 医学决策支持系统

医学决策支持系统(Clinical Decision Support System, CDSS)是医学信息技术中非常重要的应用形式,它结合了医学信息系统中的信息获取、医学知识的抽取、医学决策的推理生成等,是计算机软硬件技术在医学领域的综合应用,其目的是提升医学领域的决策水平、决策效率等。本章介绍医学决策支持系统的相关概念、技术及其应用。

3.1 医学决策支持系统概述

3.1.1 医学决策支持系统基本概念

1. 医学决策

医学决策(Clinical Decision)或称临床决策,就是所做的与治疗方案、医学处置和公共卫生政策等有关的一些重要决定。医学决策是复杂的,需要在对相关医学信息的收集、整理、加工、分析的基础上,结合决策者头脑中存储的知识,然后做出决定。因此,决策的关键是充分掌握信息并根据信息做出正确判断。

与其他科学领域的决策相比,医学决策的对象是人,所以有显著的不确定性,即决策往往要在不确定的情况下做出,这种不确定性表现在许多方面。例如,病人对自身症状的描述,“头痛”和“非常痛”的界限因各人的感觉阈值不同而界限模糊;不同医生的观察能力和记录病情的习惯不同,因此在对同一临床征象做判断和记录时,会有差异;每个人都是独特的个体,同样的治疗方案用在不同病人身上,治疗效果也是不确定的,对大多数病人有效的成熟治疗措施,可能对某些病人却得不到预期疗效,而事先,医生是很难把这部分特殊的病人区分出来。

医学决策尤其是临床医学决策的另一特点是需要进行风险值判断。心脏手术具有死亡风险,但手术成功可以延长生命和治愈心脏。在服用治疗高血压药物产生不良反应的同时,具有降低脑卒中和心肌梗死发生的可能性。根据治疗的可能结果来判断和权衡各种风险值贯穿整个医学决策分析过程。

2. 医学决策支持系统

医学决策支持(Clinical Decision Support, CDS)也称临床决策支持,是为临床医生、工作人员、患者或其他个人提供知识和特定于个人的信息,并在适当时进行智能过滤或呈现。医学决策支持系统(CDSS)也称为临床决策支持系统,是辅助医学工作人员、病人及其他潜在用户智能化地获取或筛选医学数据信息和知识,进行专项问题的辅助判断,达到提高决策水平和质量目的的系统。CDSS 最早出现于 20 世纪 50 年代,主要是医学专家

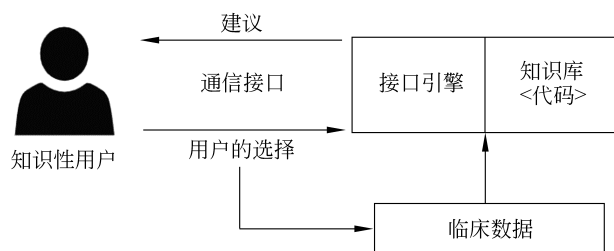
系统的开发。CDSS 的功能定义通常采用 Robert Hayward 的定义：“将医学观察与医学知识联系起来,以影响医生的诊疗决策,改善医疗卫生效果。”

医学决策支持系统可有广义、狭义之分。广义的医学决策支持系统是在医学信息系统基础上发展起来的,以支持各级医疗卫生人员辅助决策为目的,如公共卫生信息系统基础上的公共卫生决策支持系统、社区卫生信息系统基础上的社区卫生决策支持系统、临床信息系统基础上的临床决策支持系统等。狭义的医学决策支持系统是通过计算机进行模型计算、知识推理以及从医学数据获取诊断信息和诊断知识,达到支持医学诊断辅助决策的目的。也就是本章重点介绍的临床决策支持系统。

3.1.2 CDSS 的核心要素

CDSS 按系统结构可分为基于知识库的(Knowledge-based)和非基于知识库的两类。简单而言,前者依赖于包含若干规则(通常为 If-Else 结构的判定)的知识库,而后者通过机器学习去提取隐含规律。整体来说,CDSS 包含 3 部分核心要素:知识库(对 Knowledge-based 的 CDSS)或机器学习算法(对 Nonknowledge-based 的 CDSS)、推理机、人机交互。图 3.1 给出基于知识库和非基于知识库的 CDSS 结构。

基于知识库的单系统CDSS



基于非知识库的单系统CDSS

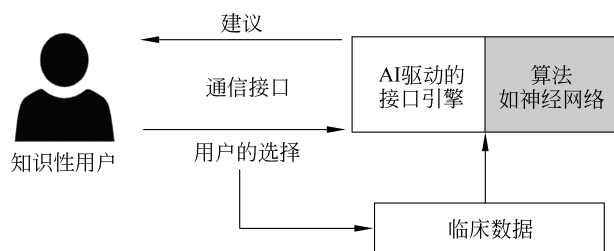


图 3.1 基于知识库和非基于知识库的 CDSS 结构

1. 知识库

知识库(Knowledge Base)是基于知识库的 CDSS 进行推理的基础。知识库的内容必须是基于证据支持的、权威的,如各种诊断标准、临床指南等,且内容需要根据证据的最新进展进行更新和维护。从来源看,医学知识库中的知识可以是原始知识和中间知识两种。

所谓原始知识是指直接从外界进入知识库的知识;中间知识则是由推理机构生成后追加入知识库的知识。从类型看,知识库中的知识可分为科学知识和经验知识两类:前者来源于医学书本、文献、专利等,如某个生物学过程的原理等,后者记录完整的病人数据和行之有效的指导原则,如临床医生根据已知的某些症状识别疾病的经验等;从医学知识库中包含知识的内容看,又可以将其分为文献数据库和事实数据库:前者存储全文、摘要等形式的文献,后者存储与人物、机构、事物等相关的情况、过程、现象、特征等方面的事实性信息,如某些病例数据库。用于 CDSS 的知识库还用于存放各种规则、因果关系、决策人员的经验等,包含经程序化语言编译的数据规则和关联,通常采用 If-Then 规则的形式。CDSS 的核心是持续更新的循证医学知识库和质量管理规则库。

2. 推理机

在 Knowledge-based 的 CDSS 中,推理机(Inference Engine)将逻辑规则应用到知识库中,按照一定的推理策略给出判断结果或是推导出新知识。由于知识库中每个新的事实都可以触发推理引擎中的附加规则,因此知识库与推理机相结合的推理过程将迭代进行。推理机的目标是模拟医学专家进行工作。人有多种思维方式,相应也有多种推理方式。CDSS 中常用到的是正向推理和反向推理、确定性推理和不确定性推理。正向推理从已知的事实出发,正向使用规则断言新的事实;反向推理从某个假设目标开始,寻找支持该假设的证据,向后推论以确定哪些事实必须被断言来实现目标。推理的确定性与否,决定了 CDSS 对于决策的支持程度:确定性推理给出决策结论,而不确定性推理则提供决策参考。在 Nonknowledge-based 的 CDSS 中,推理机通常以机器学习算法模型的形式提供。

3. 人机交互

人机交互是系统与用户之间数据输入、输出的通信及其操作界面。人机交互是 CDSS 中不可或缺的部分,用于将“人”(即使用 CDSS 的决策者)的需求和信息传输给“机”(即 CDSS 本身),并把 CDSS 所得出的决策结论反馈给使用者。系统根据用户手工输入或者从其他系统获得的条件进行判断,从知识库中抽取对应的相关信息(词条或句子)以屏幕显示等方式反馈给用户。决策支持系统应与医生的工作流程相融合,为此,良好的通信与交互形式对于 CDSS 发挥积极作用尤为重要。

3.1.3 CDSS 的工作流程

CDSS 可以应用于临床诊疗流程的诊前、诊中、诊后各阶段的多个应用场景,CDSS 可应用于临床诊疗的全流程如图 3.2 所示。诊前环节患者可通过患者的人机问答、图片上传等方式进行简单的问诊,完成病症的初步评估,实现智能分诊、导诊、转诊等;在诊中环节,系统可初步帮助医生推送合理的检查项目、明确诊断,同时给出优选的治疗方案建议;在诊后环节,系统可通过预后评估,预测病情发展情况,并对患者病情进行智能追踪,从而

实现个性化的健康可持续管理。

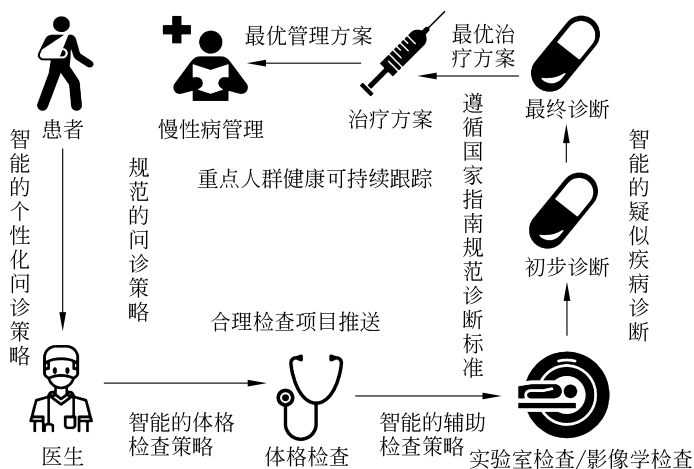


图 3.2 CDSS 可应用于临床诊疗的全流程

CDSS 的技术结构如图 3.3 所示,主要包括基础层、技术层和应用层。CDSS 应用建立在基础层和技术层之上才能够完成。其中基础层包括大数据、云计算等硬件加速,以及神经网络芯片等计算能力;技术层需要自然语言处理、认知技术、自动推理、机器学习、信息检索等技术,以及各种深度学习算法等。

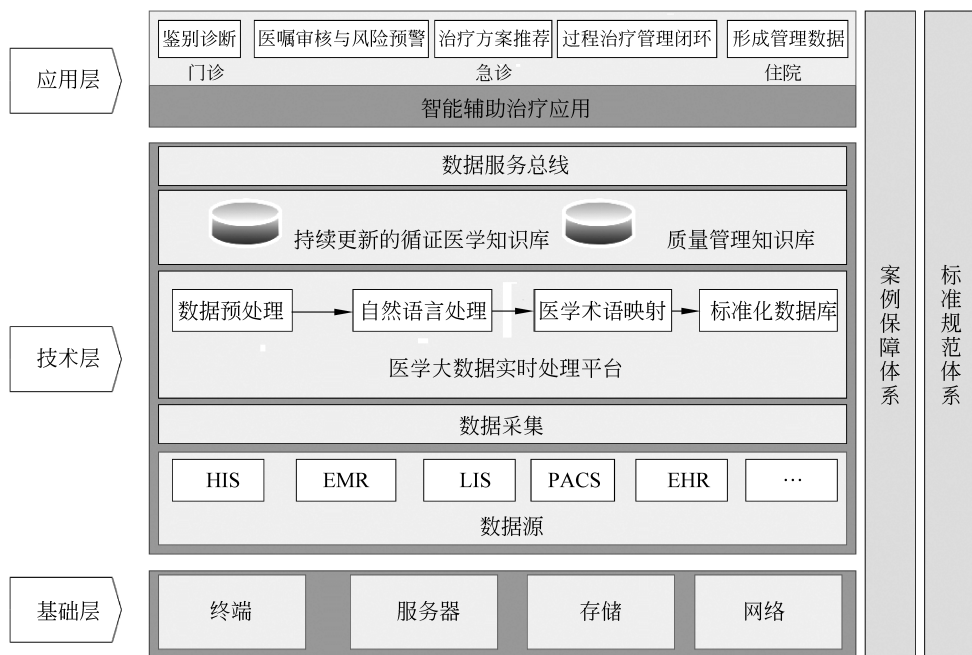


图 3.3 CDSS 的技术结构

3.2 医学决策支持系统的关键技术

3.2.1 知识库构建

1. 知识来源

知识获取是指从知识源获得知识来建造知识库的工作。知识获取的方法主要有 3 种：①知识工程师从医学专家那里获取知识，然后以正确的形式存储到知识库；②医学专家通过知识编辑器直接将自己的知识和经验存入知识库；③通过知识学习器从数据库中自动获取知识，在这种方法中，知识编辑器提供一个具有一定格式的对话界面，医学专家按照对话要求输入知识。

2. 知识表示

知识表示是指在数据库中表示医学知识，这些知识经过了结构化，能够被计算机处理，并以一种人类能理解的方式告知处理结果。知识表示形式将直接影响知识库系统的性能，目前主要的知识表示方法有逻辑表示模式、语义网、过程表示和产生式系统、特性表、框架和脚本等。因为医学知识库中的知识来源广泛，来自不同领域、不同的专家，执行不同的功能，不同来源的知识结构和表示往往有很大的差别，所以往往采用多个子知识库，每个子知识库中尽量使用一种知识表示方法，从而使系统中的知识易于处理、解释和管理。自然语言处理（Natural Language Processing，NLP）及知识图谱（Knowledge Graph，KG）技术都是知识表示中常用的技术。

3. 医学知识库的特点

知识库设计人员、医学专家以及大量的知识构成了开发和改进知识库系统的 3 个基本要素。DSS 知识库所涉及的知识广泛，既有知识库特定的医学知识，又有建模知识和求解方法知识。医学知识尤其是医学专家长期积累的经验性知识，是系统开发人员不熟悉且不可能在短期内学到的；而知识库实现所需要的大量计算机专业知识，特别是人工智能和知识工程方面的知识，医学专家即使经过长时间学习，也难掌握。所以知识库的建立和维护必须依靠系统设计人员与医学领域多方面的专家密切合作。

医学知识库中的知识很多是经验性知识，是医学专家根据某些重复出现的因果联系或凭借某些直觉而获得的，缺乏研究的理论基础，因此医学专家在描述这些知识时很难做到准确无误，建造知识库自然避免不了反复测试、扩充及修改的过程。知识的冗余是指获取和利用各具不同优点的多来源知识解决问题，用知识的冗余是一种弥补医学知识的不完整和不精确的有效方法。

3.2.2 推理机技术

1. 临床检验和贝叶斯定理

临床治疗和临床数据充满大量不确定性和不完整性的因素,建立在这些因素上的临床决策自然是相当复杂的。临床医生可以凭直觉来确定最终结果,但直觉既不充分也不可靠;他们还能够单凭演绎推理的方法得到结果;但更多的时候,医生的诊断借助于医学知识和长期实践工作的经验积累,判断病症与疾病之间的关联,得出诊断结果,决定治疗方案。这就是以概率的方法来解决非确定性问题,其中主要使用贝叶斯定理。

1) 临床检验的概率问题

临床医生在诊断时往往先初步观察病人,结合医学知识和个人经验,判断病人是否患病以及患什么病;接下来进行进一步的检查、化验等,尽可能消除疑点,缩小误差,得到较为准确的疾病类型判断和相应概率。在初步观察时确定的疾病概率,称之为先验概率(Pretest Probability)。经过检验一般可以降低诊断的不确定性,得到对疾病的新的判断,这个概率称之为后验概率(Posttest Probability)。

理想的检验结果应该是明确的,正常或异常,即病人没有患病或患病,界限清晰而没有重叠。但实际上,检验结果呈正态分布,正常和异常的结果之间有重叠的部分。也就是说,健康的人群会得到患病的结果,而患病的人群可能检查出来是健康的。于是,接受检查的人员分成4种情况。

- (1) 患病,且检查结果为阳性异常——真阳性(TP)。
- (2) 无疾病,且检查结果为阴性正常——真阴性(TN)。
- (3) 患病,但检查结果为阴性正常——假阴性(FN)。
- (4) 无疾病,但检查结果为阳性异常——假阳性(FP)。

所有接受检验人员的总数应该是这4种情况之和,即 $TP+FN+TN+FP$; 检验呈阳性的人数为 $TP+FP$; 检验呈阴性的人数为 $TN+FN$; 真正患病的人数等于 $TP+FN$; 理想化的检验结果是消除 FN 及 FP。

这样,我们可以用以下公式来衡量某种决策的有效性。

- (1) 灵敏度(Sensitivity, TP rate): $TPR = TP / (TP + FN)$ 。
- (2) 特异性(Specificity, TN rate): $TNR = TN / (FP + TN)$ 。
- (3) 患病率(Prevalence) = 患病人数 / 受测人员总数 = $(TP + FN) / (TP + FN + TN + FP)$ 。
- (4) 阳性预测率(Positive Predictive Value, PV+) = 患病人群检验呈阳性的概率 = $TP / (TP + FP)$ 。
- (5) 阴性预测率(Negative Predictive Value, PV-) = 健康人群检验呈阴性的概率 = $TN / (TN + FN)$ 。

另外还有一个衡量检验有效性的概念——决策效能 T , T 可以表示为: $T = (TP + TN) / (TP + FN + TN + FP) = (TPR + TNR) / 2$, T 值介于 0 与 1 之间,越接近于 0, 检验

的鉴别力越小;越接近于 1,则检验的鉴别力越大。

2) 贝叶斯定理

贝叶斯定理用于描述诊断和临床表现的关联,通过不同临床表现的概率得出不同疾病诊断的概率。假设 S 是某种临床表现,它可能由好几种不同的疾病引起。用 D_i 表示 S 对应的第 i 种疾病, $P(D_i)$ 为疾病 D_i 的发生概率(即先验概率),往往通过流行病学方法得到。例如,某个地区高血压的发病率是 30%,那么在这个地区随机抽取样本,得到高血压罹患的可能就是 30%,这个值也就是患病率。条件概率 $P(S|D_i)$ 为疾病 D_i 发生的情况下,症状 S 出现的概率,例如高血压病人头晕头痛的症状发生率为 70%。条件概率 $P(D_i|S)$ 表示症状 S 出现时疾病 D_i 发生的概率(即后验概率),如病人有头晕头痛的临床表现,那么他患高血压的概率有多少,这可以通过以下贝叶斯定理求得。

$$P(D_i | S) = \frac{P(D_i) \cdot P(S | D_i)}{\sum_{j=1}^n P(D_j) \cdot P(S | D_j)}$$

贝叶斯定理决策法就是利用贝叶斯定理修正先验概率,求得后验概率,据此进行决策的方法。

下面是一个具体的例子。用某种方法检查癌症,根据临床记录,癌症病人施行该项检查结果为阳性的概率为 95%,非癌症病人该项检查结果为阴性的概率为 90%,又由以往的统计,某地区癌症的发病率为 0.0005,如果用该方法在此地区进行癌症检查,效果如何?

在这里,我们通过求检验的真阳性,即检验结果呈阳性时病人患癌症的概率,就可以判断该检验方法的效果,这个概率越大,检验方法就越有效。设 S 为检验结果是阳性的情况, D 是癌症病人,则: $P(D)=0.0005$, $P(S|D)=0.95$,此地区人群非癌症 $\bar{D}$ 的比例为 $P(\bar{D})=1-P(D)=0.9995$,假阳性(非癌症但检验呈阳性)的概率为 $P(S|\bar{D})=1-90\%=0.10$,由贝叶斯定理得:

$$P(D | S) = \frac{P(D) \cdot P(S|D)}{P(D) \cdot P(S|D) + P(\bar{D}) \cdot P(S|\bar{D})} = \frac{0.0005 \times 0.95}{0.0005 \times 0.95 + 0.9995 \times 0.10} = 0.0047 \ll 1, \text{所以这种检验方法准确度很低。}$$

贝叶斯定理考虑到了阳性和阴性的信号,其结果比较可靠。但使用的难点在于估计先验概率和条件概率,要求同时满足:模型中的 i 种疾病互斥,先验概率之和为 1(即要构成一个完整的疾病群);用于鉴别诊断的症候指标互相独立无关;当计算出各后验概率 $P(D_i|S)$ 后,作为临床判断的依据只有当各 $P(D_i|S)$ ($i=1,2,\dots,n$) 间差距达 5 倍以上时方可下结论,或是当某一后验概率值达 0.85 才下结论。实际工作中很难满足这样苛刻的条件,因此贝叶斯定理决策法的应用受到限制。

美国犹他大学的 ILIAD 决策诊疗模型采用了完整意义上的贝叶斯决策模型。

2. 决策树与决策分析

当医生需要做出临床决策,例如选择某种治疗方案或分析医疗风险时,常遇到这样的问题:有若干个治疗方案可供选择,分析一下病人的情况和这些方案,大部分条件是已知的,但还存在一定的不确定因素,每个方案的执行都可能出现几种结果,各种结果的出现

有一定的概率,临床决策存在着一定的胜算,也存在着一定的风险。这时,决策的标准只能是期望值,即各种状态下的加权平均值。针对上述问题,用决策树法来解决不失为一种较好的选择。

决策树(Decision Tree)一般都是自上而下生成的。每个决策或事件(即自然状态)都可能引出两个或多个事件,导致不同的结果,把这种决策分支画成图形很像一棵树的枝干,故称决策树。

决策树的构成有4个要素:①决策结点;②方案枝;③状态结点;④概率枝,如图3.4所示。

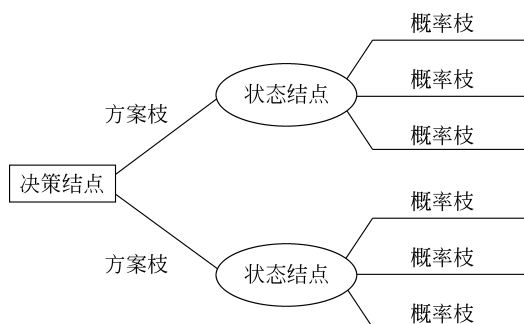


图 3.4 决策树的构成

图3.4中,决策结点用方框表示,它表示一个时间点,从这点出发决策者需主观选择一种行动方案。由决策结点引出若干条细支,每条细支代表一个方案,称为方案枝,也称决策枝。每个方案枝导向一个状态结点,也称机会结点,代表执行此方案产生的不受决策者控制的可能性,如手术后是否感染等。由状态结点引出若干条细支,称为概率枝,也称机会枝,每条概率枝代表一种自然状态,在每条细枝上标明状态的内容和其出现概率。在概率枝的最末梢标明该方案在该自然状态下所达到的结果,如死亡或康复。这样树状图由左向右、由简到繁展开,组成一个树状网络图。

决策树法采用这种树状网络图作为分析工具,利用概率论的原理,用决策结点代表决策问题,用方案枝代表可供选择的方案,用概率枝代表方案可能出现的各种结果,经过对各种方案在各种结果条件下损益值的计算比较,为决策者提供决策依据。

常用的具有决策树功能的软件有 SAS/EM、SPSS Answer Tree、CART、See5-demo、KnowledgeSEEKER、KnowledgeSTUDIO、Business Miner、Decision series 等。

3. 人工神经网络

人工神经网络(Artificial Neural Network, ANN)是理论化的人脑神经网络的数学模型,是基于模仿大脑神经网络结构和功能而建立的一种信息处理系统。随着神经网络技术的发展,其用途日益广泛,应用领域也不断拓展,在医学领域的应用也越来越广泛,如用于临床诊断、预后研究、临床决策分析、医学信号分析处理等。

1) 人工神经网络的结构和工作原理

人体内各部分之间信息传递的基本单元是神经元,每个神经元都由一个具有简单处

理作用的细胞体、一个连接其他神经元的轴突和树突组成。人的大脑拥有约上百亿个神经元,是一个能够进行感受、记忆、联想及反应等复杂思维的庞大信息处理体系。ANN 是一种模仿人的神经结构发展起来的计算机网络系统,由许多相对独立的人工神经元彼此连接成网络,模仿生物神经处理信息的方式解决问题。

如图 3.5 所示,ANN 是一种由多个人工神经元以某种规则连接而成的层次网络结构,一般分为输入层、隐含层和输出层,各层包含多个并行操作的神元(结点)。输入层的每个结点(图中以序号表示)对应一个个的预测变量,输出层的结点对应目标变量,可有多。在输入层和输出层之间是隐含层(对 ANN 使用者来说不可见),隐含层可能不止一层而是多层,隐含层的层数和每层结点的个数决定了神经网络的复杂程度。除了输入层的结点,神经网络的每个结点都与很多它前面的结点(称为此结点的输入结点)连接在一起,每个连接对应一个权重 W_{xy} ,此结点的值就是通过它所有输入结点的值与对应连接权重乘积的和作为一个函数的输入而得到,这个函数称为活动函数或挤压函数。在图 3.5 中,结点 4 输出到结点 6 的值 W_{46} 可通过如下计算得到: $W_{14} \times \text{结点 1 的值} + W_{24} \times \text{结点 2 的值}$ 。

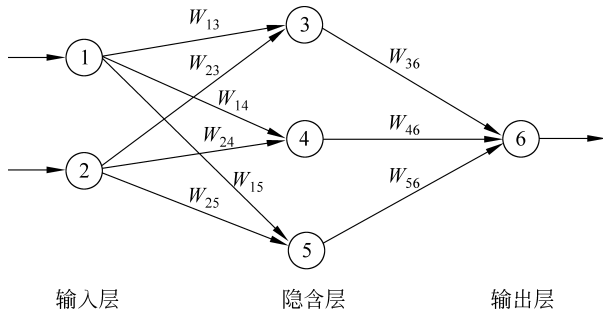


图 3.5 ANN 结构示意图

当一组数据(或称输入模式)输入 ANN,网络输入层的每个单元都接收到输入模式的一部分,对输入模式缓冲后,通过输入层与隐含层的连接权重将输入信息传至隐含层;隐含层中的各处理单元有的兴奋,有的抑制,经过隐含层处理过的信息传至输出层;输出单元将隐含层单元的输出作为自己的输入,同时也经过隐含层与输出层间权重矩阵及输出单元阈值的作用,在激活函数处理后也有的兴奋,有的抑制。输出层单元输出的模式就是网络对输入模式激活的总效应。

ANN 的一个显著特征是它通过向环境学习获取知识并改进自身性能从而来解决问题,所以要想用 ANN 解决实际问题必须先让它学习或者说对它训练,让它掌握输入样本的内在规律性,从而才能完成对新样本的正确识别、分类或做出某种响应。网络在学习过程中不是通过修改单元本身来完成训练,而是靠改变网络中的连接权重进行学习。

在诸多类型的神经网络中,最常用的是如图 3.5 所示的前向传播式神经网络,下面以此为例介绍 ANN 的学习过程,为讨论方便假定只含有一层隐含层。

(1) 前向传播。数据从输入到输出的过程是一个从前向后的传播过程,后一结点的值是通过与其前面相连的结点传过来,然后把值按照各个连接权重的大小加权输入活动函数再得到新的值,进一步传播到下一个结点。

(2) 回馈。当结点的输出值与预期的值不同,也就是发生错误时,神经网络就要“学习”(从错误中学习)。可以把结点间连接的权重看成后一结点对前一结点的“信任”程度(其向下一结点的输出更容易受其前面那个结点输入的影响)。学习的方法是采用惩罚的方法,过程如下:如果一结点输出发生错误,那么看其错误是受哪个(些)输入结点的影响而造成的,是不是其最信任的结点(权重最高的结点)使其出错,如果是则降低对该结点的信任值(降低权重),以示惩罚;同时提升那些做出正确建议结点的信任值。对那些受到惩罚的结点来说,它也需要用同样的方法来进一步惩罚其前面的结点。就这样把惩罚一步步向前传播直到输入结点为止。

用前向传播得到输出值,如果发生错误,则用回馈法进行学习,对训练集中的每一条记录都要重复这个步骤。把训练集中的每一条记录都运行过一遍,就是完成了一个训练周期。要完成神经网络的训练可能需要很多个训练周期,经常是几百个。训练完成之后得到的神经网络就是通过训练集发现的模型,描述了训练集中响应变量受到预测变量影响的变化规律。

由于神经网络隐含层中的可变参数非常多,如果训练时间足够长的话,神经网络很可能把训练集的所有细节信息都“记”下来,而不是建立一个忽略细节只具有规律性的模型,这种情况称之为训练过度。显然这种“模型”对训练集会有很高的准确率,而一旦离开训练集应用到其他数据上,准确度很可能会急剧下降。为了防止这种训练过度的情况,必须知道什么时候需要停止训练。在有些软件中会在训练的同时用一个测试集来计算神经网络在此测试集上的正确率,一旦这个正确率不再升高甚至开始下降时,就认为该神经网络已经达到最好的状态,可以停止训练了。

神经网络和统计方法在本质上有很大差别。神经网络中的参数可以比统计方法中的多很多,如图 3-5 中就有 13 个参数(9 个权重和 4 个限制条件)。由于参数多,而且参数通过各种各样的组合方式来影响输出结果,所以很难对一个神经网络表示的模型做出直观的解释。实际上神经网络也正是作为“黑箱”来使用的,可以不关心“黑箱”的内部结构,只要了解其输出的功能就可以使用。在大部分情况下,这种限制条件是可以接受的。就如银行需要一个笔迹识别软件,但使用者没有必要知道为什么这些线条组合在一起就是一个人的签名,而另一个相似的却不是。

2) 人工神经网络应用举例——疾病诊断

建立一个 ANN 疾病诊断的模型,输入变量就是病人的临床资料,如症状、体征以及各种检查结果等,而输出变量则为是否患某种疾病的诊断。

在这里,ANN 疾病诊断模型的作用就如同临床医生。临床医生在进行疾病诊断时,传递给大脑的信息就是病人的临床信息,结合大脑中已存储的临床经验进行决策。但是临床信息(输入变量)的来源很难完整,也无法排除假象和不确定的情况,在众多的信息中,哪些价值高的信息应着重采纳,哪些次要的只作参考,各个医生的意见并不一致,这也使医生的诊断效果不够理想。虽然专家系统的研制和应用取得了重大进展,但在实现类似人脑的学习、联想等方面存在着知识获取的“瓶颈”问题。而 ANN 的自学习、联想记忆、高度并行和容错等功能可以突破这一障碍,根据已学会的知识和处理问题的经验对复杂问题做出合理的判断。因此,ANN 很适用于医学领域的疾病诊断。ANN 疾病诊断模

型包括两个基本程序：训练和测试。

(1) 训练。将描述病人各种情况的数据作为输入变量加到模型的输入端,同时调整神经元之间的连接权值以使模型的输出和实际的病例情况相符,即当病人确实患有某种疾病时,网络的输出结果也恰好指示为该种疾病,反之亦然。若训练集模型的输出基本上与实际结果一致,则训练过程结束,认为 ANN 已建立起病人的各种因素与其是否为某疾病病人之间的函数映射关系。为了使函数关系具有普遍性,以便提高准确性,一般对某种疾病往往要输入成百上千的样本才能获得经过良好训练的 ANN 系统。

(2) 测试。对新的候诊病人,将该病人的各种情况输入到训练好的神经网络中去,根据网络的输出结果就可以知道该病人是否患有此种疾病。

3.2.3 人机交互

CDSS 中狭义的人机交互技术包括传统的命令行或图形化交互界面,用于将结果以文字或是图形的形式输出;以及借助键盘、鼠标等输入形式,用于将所需的患者临床数据等信息输入到系统中。随着信息技术在医疗领域的日益深入,CDSS 与临床原有工作流程的衔接越来越紧密,其智能化程度也越来越高。利用语音输入、自然语言理解、智能对话生成等技术,可以实现基于语音形式的人机交互;利用数据格式和数据接口的统一,可以实现 CDSS 从 HIS 直接获取数据,以及 HIS 中智能化的流程推送;利用自然语言理解与生成技术,还可以实现 CDSS 的决策决定的标准化报告生成。简而言之,技术的发展为 CDSS 的人机交互方式在传统形式的基础上提供了各种新的、便捷的可能。

3.3 医学决策支持系统的实例

3.3.1 CDSS 实例

本节以一个案例来说明运用决策树进行临床决策方案的选择。慢性肝炎和肝硬化是临床上常见的疾病,两者的发病概率分别为 20%和 80%,两者均有慢性肝衰竭的症状,但治疗方法有很大差异。用类固醇治疗慢性肝炎,可以使病人的两年生存率从 67%提高到 85%;但用其来治疗肝硬化,不但没有效果,还可能导致病人胃肠道出血和血管栓塞等并发症,并发症的风险使肝硬化病人的两年生存率从 50%降到 48%。临床上通常采用肝活组织检查(活检)来鉴别这两种疾病,活检存在 1%的死亡率。如果慢性肝衰竭的病人只考虑慢性肝炎和肝硬化两种可能的病情诊断,并且假设活检能够 100%确诊疾病,将病人两年的生存率作为评价指标,可以构建如图 3.6 所示的决策树。为了区分起见,图中概率枝上的发生概率以小数形式给出,评价指标(两年生存率)以省略百分号的形式给出。

构造决策树时,依据临床诊疗的思路,先做出决策是否进行活检,用序号 1 标记第一个决策结点。如果进行活检,其结果不受决策者控制,而是由机遇决定,所以这是个状态结点,由字母 A 标记。活检可能引起病人死亡或存活,所以状态结点 A 引出两条概率枝,概率枝上标明死亡或存活的概率,死亡概率已知为 0.001,存活概率等于 $1 - 0.001 =$

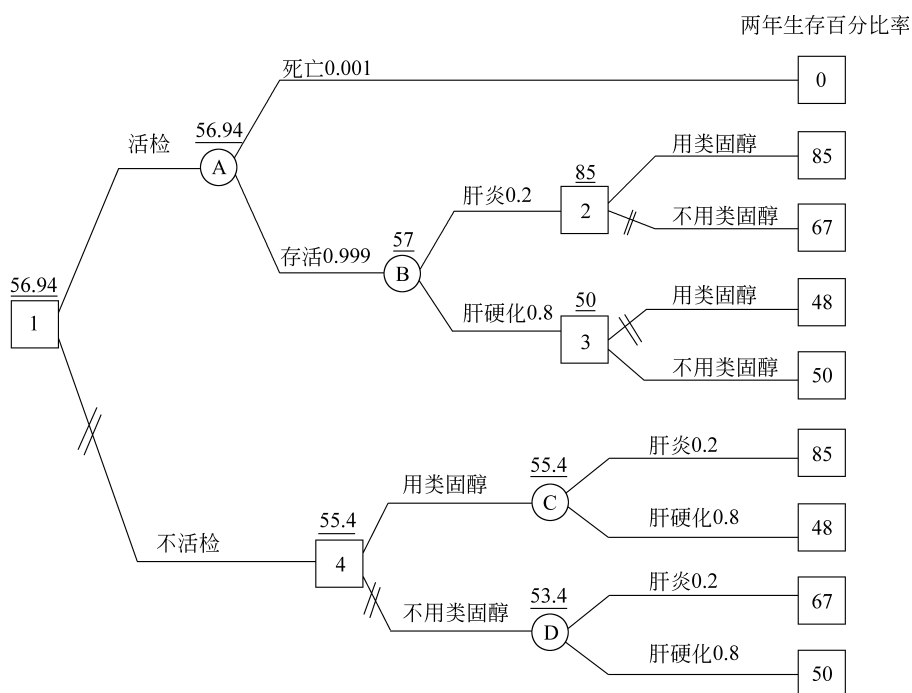


图 3.6 决策树示例

0.999。对于死亡的情况,两年生存的百分比自然为 0。活检存活的病人有两种可能,患慢性肝炎或肝硬化,这也不是决策者所能控制的,用状态结点 B 标记,肝炎和肝硬化的比例已知,分别为 0.2 和 0.8,写在 B 引出的概率枝上。对这两种病症,均有是否用类固醇治疗的决策,分别以序号 2 和 3 标记这两个决策结点。慢性肝炎用与不用类固醇治疗的两年生存率都已知,为 85% 和 67%,对应写在由决策结点 2 引出的概率枝后面;如果病人活检的结果是肝硬化,同样方法,将用与不用类固醇治疗的两年生存率 48% 和 50% 写在由决策结点 3 引出的概率枝后面。

在做序号 1 的决策时,还可以选择不对病人活检,接下来就是决策这部分不活检病人的用药问题,用序号 4 标记。如果对病人用类固醇治疗,得到的结果也是决策者所不能控制的,用状态结点 C 表示。这个结果取决于病人患病的类型,如果是肝炎(占病人比例的 20%,写在 C 引出的一个概率枝上),用药后两年生存率为 85%;如果病人患肝硬化(占病人比例的 80%,写在 C 引出的另一个概率枝上),用药后两年生存率为 48%。如果对不活检病人不用类固醇治疗(状态结点 D),同样方法引出概率枝和两类病人的两年生存率,分别为 67% 与 50%。

按上述思路,建立起所属问题对应的决策树。图中○的结点 A、B、C、D 称为状态结点;□的结点 1、2、3、4 称为决策结点。决策树方法要解决的问题,就是根据给定的决策树,找到使评价指标最优的决策结点的顺序路径。为此,一种方法是采用预期货币值(Expected Monetary Value, EMV)作为衡量标准,以计算各种决策方案的收益值,来判断哪个方案是最优的。此例中以决策树的最后结点(两年生存百分比)为起点,向前回溯

到决策树的起始点(决策结点 1)。根据已知概率,计算各个结点的 EMV,选择 EMV 大的结点,获得最佳治疗方案。

以下是各个结点 EMV 的计算方法,计算结果以加下画线的形式写在图 3.6 的每个结点上面。首先对最后结点的第一层前溯结点进行计算,即图中的结点 2、3、C、D。其中,对于状态结点和决策结点的 EMV 的计算方法不同。

对于状态结点,其 $EMV = \text{各概率枝后续结点的 EMV 对于概率值的加权和}$ 。因此,对于状态结点 C、D,EMV 为

○ 状态结点 C: $EMV = 85 \times 0.2 + 48 \times 0.8 = 55.4$ 。

○ 状态结点 D: $EMV = 67 \times 0.2 + 50 \times 0.8 = 53.4$ 。

而对于决策结点,其 EMV 取各概率枝上后续结点的 EMV 较大的值,而 EMV 较小的概率值则在图中以双斜杠形式表示删枝。因此,对于决策结点 2、3,EMV 为

□ 决策结点 2: $EMV = 85$ 。

□ 决策结点 3: $EMV = 50$ 。

同理,依次计算剩余的结点 B、4、A、1,各结点的 EMV 分别为

○ 状态结点 B: $EMV = 85 \times 0.2 + 50 \times 0.8 = 57$ 。

□ 决策结点 4: $EMV = 55.4$ 。

○ 状态结点 A: $EMV = 0 \times 0.001 + 57 \times 0.999 = 56.94$ 。

□ 决策结点 1: $EMV = 56.94$ 。

综合以上分析,每一个决策点都被保留了一个最优决策枝。对于本例中决策的起始点(即决策结点 1),能够得到的 EMV 最大为 56.94(当采用活检检查时达到),其次为 55.40(当不活检且采用类固醇治疗时达到)。因此,CDSS 系统得到的决策结论是对此案例进行活检检查。当活检结果区分了肝炎或肝硬化,再进一步选择相应是否采用类固醇的治疗方案。

3.3.2 CDSS 的发展现状与展望

CDSS 的研究开始于 20 世纪 50 年代,最早的研究方向是基于知识库的专家系统。如早期由美国斯坦福大学开发的 MYCIN 系统,包含了约 350 条规则,用于对血液检验的结果进行细菌诊断和抗菌素选用。我国人民卫生出版社推出的“人卫临床助手”包含了疾病知识和典型病例两个核心库,同时整合了国家临床路径、医疗相关标准、医患沟通、临床伦理思维、医疗损害防范、医学法律法规等知识。随着深度学习、人工智能技术的崛起,CDSS 的研究有了新的方向。首先是基于自然语言处理(NLP)模块,CDSS 可以从海量的医学文献库中获取知识。如 IBM 公司的 Watson 的突出技能之一就是在极短的时间内学习和理解海量的医学知识。此外,随着医疗行业信息化程度的提高,利用 HIS-CDSS-互联网数据库的实时对接,可在瞬间查阅海量文献资料,实现高效的学习。

CDSS 作为未来智慧医院电子病历评级的重要条件以及医院信息化迭代的需求,其市场正从大医院往基层医疗机构转移,用于帮助解决基层医疗医生数量少、诊疗水平弱、慢病管理压力大等问题。同时,CDSS 的应用场景也逐步从单一的诊疗辅助,扩展到医院

质控和管理等领域。目前 CDSS 正向医院/科室管理、科研协作平台搭建、结构化病历系统、患者交互及患者教育、医生继续教育、药物警戒、医疗控费等方向发展,走向了医疗管理的全流程。

3.4 医学决策支持实验

3.4.1 知识概要

在本章实验中,需要用到的相关知识点如下。

- (1) 决策树与决策分析。
- (2) 贝叶斯定理。

3.4.2 实验目的和实验内容

学习和掌握医学决策支持的基本流程;学习决策树及贝叶斯定理在医学决策支持方面的应用;掌握决策树和贝叶斯定理在实际案例中的计算。

具体实验内容:根据特定的患者数据,使用决策树进行决策分析,并结合贝叶斯定理进行计算,获得结果。

3.4.3 实验 基于决策树的医学决策支持实验

1. 实验任务

已知某一新的疾病 X 的常见症状是咳嗽、发烧和疲倦,而我们对此疾病 X 的认识目前仅限于 10 个病例,数据如表 3.1 所示。

表 3.1 疾病 X 的患者数据(表中 1 代表“是”,0 代表“否”)

是否咳嗽	是否发烧	是否疲惫	是否诊断为 X
1	0	1	0
1	1	0	0
0	1	1	1
1	1	1	1
0	1	1	1
0	1	1	0
1	1	1	1
0	0	0	0
1	1	0	1
1	1	1	1

构建根据症状诊断是否 X 的决策树,并得出某一新病例 A 关于疾病 X 的诊断。已知 A 有咳嗽、发烧,但不觉得疲惫。

2. 操作要点

(1) 统计各症状与诊断 X 之间的样本分布。

根据已有病例数据,针对某一症状分别统计发生或未发生该症状情况下被诊断为 X 或不被诊断为 X 的例数。以咳嗽为例,当发生咳嗽时($N=6$),被诊断为 X 和未被诊断为 X 的人数分别为 4 和 2;当未发生咳嗽时($N=4$),被诊断为 X 和未被诊断为 X 的人数分别为 2 和 2。同理,对发烧、疲惫做类似统计。

(2) 计算各症状的基尼不纯度,选取第一决策条件。

为了确定各特征(此例中即各症状)对决策的影响程度,可以计算各特征的基尼不纯度(Gini Impurity)。对某一结点(如发生咳嗽症状),其基尼不纯度的计算公式如下:

$$I_G = 1 - \sum_{i=1}^C p_i^2$$

其中, p 是该结点属性(如发生咳嗽症状)在某一分类(被诊断为 X 和未被诊断为 X)下的概率。例如发生咳嗽情况下的 $p_1=4/6, p_0=2/6$,则发生咳嗽对应结点的基尼不纯度 =

$$1 - \left(\frac{4}{6}\right)^2 - \left(\frac{2}{6}\right)^2 = 0.44。$$

根据贝叶斯后验概率的计算公式,对于咳嗽症状的结点,计算其总的基尼不纯度(或基尼增益)等于各分支结点的基尼不纯度的加权和。咳嗽症状的结点包括发生咳嗽和未发生咳嗽两个分支,则:

$$\begin{aligned} \text{咳嗽的基尼不纯度} &= \text{发生咳嗽的概率} \times \text{发生咳嗽对应结点的基尼不纯度} + \\ &\quad \text{未发生咳嗽的概率} \times \text{未发生咳嗽对应结点的基尼不纯度} \\ &= \frac{6}{10} \times 0.44 + \frac{4}{10} \times 0.50 = 0.47 \end{aligned}$$

同理,可以分别计算发烧和疲惫对应的基尼不纯度。选择其中基尼不纯度最低的症状,则该症状区分疾病 X 与否的效果最佳,可作为决策树的根结点(即第一决策结点)的决策条件。

(3) 重复(2),进一步确定决策树的完整结构,并绘制树状结构图。

(4) 对于实验任务中给定的新病例 A(有咳嗽、发烧,但不觉得疲惫),根据决策树给出你的诊断: 病例 A 是否患有疾病 X?

(5) 思考当病例数不断增多,可能对决策即诊断结果产生什么影响?

练习与思考

一、判断题

1. 临床医疗的实质是科学决策的过程。 ()

2. HIS 中的决策支持系统实现方法都来自于统计学、数据库、人工智能等技术。 ()
3. 贝叶斯定理决策法就是利用贝叶斯定理修正先验概率,求得后验概率,据此进行决策的方法。 ()
4. 一个完整的医学决策支持系统主要由知识库、推理机和人机交互构成。 ()
5. CDSS 只是对于诊疗的辅助。 ()

二、选择题

1. 决策树方法由结点和分枝组成,其中用小方框□代表()结点,表示决策者可控制的;用圆圈○代表机会结点,表示决策者无法控制的。
A. 决策 B. 控制 C. 决定 D. 根
2. 医学知识库系统的实现需要解决()、知识利用和知识获取 3 个关键技术。
A. 知识传播 B. 知识表示 C. 知识存储
D. 知识识别
3. 以下不属于医学决策的特点的是()。
A. 不确定性 B. 复杂 C. 经济因素是关键 D. 需进行风险评估
4. 以下不属于知识库系统的基本要素的是()。
A. 知识库设计人员 B. 患者 C. 医学专家 D. 大量的知识
5. ()是由多个人工神经元以某种规则连接而成的层次网络结构,一般分为输入层、隐含层和输出层。
A. DNN B. CNN C. BNN D. ANN

三、思考题

1. 什么是决策支持系统? 其基本结构如何?
2. 医学决策支持系统的主要目标是什么?
3. 医学决策支持系统中常用的方法有哪些? 在实践中应用了哪些技术?
4. 医学决策支持系统中知识库的主要作用体现在哪些方面?
5. 目前有关“大数据”的概念对医学决策支持会有怎么样的影响?