

·医学信息学·

数据仓库技术在糖尿病健康教育中的应用研究

陈翔,刘春,杨锐

(安徽中医药大学第一附属医院,安徽 合肥 230031)

摘要:阐述糖尿病健康教育对糖尿病病情治疗的作用,介绍了数据仓库技术和数据挖掘的概念及工作原理,探讨数据仓库技术在糖尿病健康教育系统中的智能分析与挖掘,以便为患者寻找最适合的健康教育方案,增强患者对病情的自我管理能力和提高糖尿病患者的生活质量。

关键词:糖尿病;健康教育;数据仓库;智能分析;挖掘

中图分类号:TP311.13

文献标识码:B

DOI:10.3969/j.issn.1006-1959.2018.20.001

文章编号:1006-1959(2018)20-0001-03

Application Research of Data Warehouse Technology in Diabetes Mellitus Health Education

CHEN Xiang,LIU Chun,YANG Rui

(the First Affiliated Hospital of Anhui University of Traditional Chinese Medicine,Hefei 230031,Anhui,China)

Abstract: Explain the role of diabetes mellitus health education in the treatment of diabetes, introduce the concept and working principle of data warehouse technology and data mining, and explore the intelligent analysis and mining of data warehouse technology in diabetes mellitus health education system, in order to find the most suitable health education for patients. The program enhances the patient's self-management ability to improve the quality of life of diabetic mellitus patients.

Key words: Diabetes mellitus; Health education; Data warehouse; Intelligent analysis; Mining

糖尿病(diabetes mellitus, DM)是一种典型的慢性终生疾病。为了全面、有效地控制该病,除了常规的临床和药物治疗外,还必须对患者进行适当的健康教育^[1]。目前,糖尿病患者的自我管理能力和水平仅处于中、低水平,患者在日常生活中经常出现各种急慢性并发症,严重影响患者的生活质量。如何提高糖尿病患者自我管理的能力,减轻糖尿病对患者身心健康的影响,已成为国内外糖尿病相关学者的重要研究目标。积极开展糖尿病健康教育是提高患者自我管理能力和水平的最有效手段^[2]。目前,国内外学者都在积极探索糖尿病健康教育的有效模式。健康教育在糖尿病治疗中的作用越来越重要。通过让患者掌握糖尿病的基本知识,可以有意识地进行饮食控制、相关运动治疗、药物治疗和疾病监测。根据患者的实际情况,制定相应的健康教育计划是控制糖尿病最有效、最长远的科学措施^[3]。本文基于数字化技术,通过糖尿病健康教育数据仓库的构建,对患者的基本信息、临床表现特征、所采用的教育手段、疗效等一系列相关数据进行清理、导入与规范化处理,并使用现代化的数据统计与挖掘技术,实施智

能分析与挖掘,寻找出对患者最适合的健康教育方案,增强患者对糖尿病自我管理的能力,从而提高患者的生活质量。

1 数据仓库

1.1 概念 数据仓库(Data Warehouse, DW)的定义首先 W.H.NMOM 提出^[4]。DW 是面向主题的、集成的、相对稳定的、依赖于时间的、历史的、持久的、支持决策的数据集合,其存储数据和管理资源的基本手段仍然沿用传统的数据库技术,分析数据和提取信息最为行之有效的办法仍然是统计分析技术,当今最为流行的分析技术是人工智能挖掘。DW 与计算机网络技术、并行计算技术、人工智能技术等相结合、相互渗透和集成应用,并应用于对相应决策制定过程的支持^[5]。

1.2 DW 的功能 DW 在功能上应该具备多层次、多方位的数据挖掘与辅助决策的能力,能够很好地助力于临床与科研能力的提高。数据是进行糖尿病健康教育智能分析与挖掘的基础,因此 DW 是支撑糖尿病健康教育智能分析与挖掘系统最基础的工具。DW 通过对 HIS、LIS、PACS 与 EMR 等系统中与糖尿病患者健康教育治疗相关的数据提取出来,也就是将分散在各系统中难于访问的数据,集中转化为统一的且随时可用的糖尿病健康教育信息。

1.3 DW 的设计与创建 DW 总体上分为基本功能

基金项目:国家中医药管理局项目(编号:JDZX2015140)

作者简介:陈翔(1985.12-),男,安徽安庆人,硕士,工程师,主研方向:数据挖掘,中医药信息化

通讯作者:杨锐(1991.10-),男,安徽滁州人,本科,工程师,主研方向:中医药信息化,数据挖掘

层、管理层和环境支持层三个层次。通过设计过程的循环,实现对数据的选择、变换、建模、评估、解释、运用和巩固,见图 1。

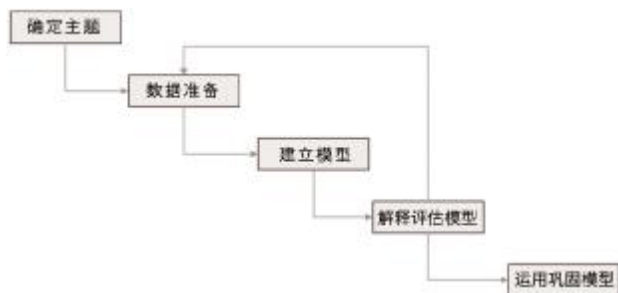


图 1 DW 的结构模型

创建基于 HIS、LIS、PACS、EMR 等医院信息系统的糖尿病健康教育数据仓库,是从现有数据出发的一种设计方法,即一种“数据驱动”的系统设计方法。其基本思想是利用医院的主要业务系统数据库中的数据,按照糖尿病健康教育这个研究方向所对应的分析领域对数据与数据之间的联系重新进行考虑,确定数据仓库中的主题,并利用数据模型有效地识别数据与数据仓库中的主体数据的“公共性”。

从数据仓库的定义可以看出,构建一个糖尿病健康教育数据仓库应该经历以下几个步骤:提取主题、组织数据、获取并集成数据和构建应用。随着数据仓库中的数据量的增加,积累的数据相互之间是否存在某些潜在的、未知的关联、模式或趋势?然而,这些关联、模式或趋势不能通过我们的肉眼或简单的计算方式就获得,而必须通过对这些数据的一系列再处理和加工过程,需要利用数据挖掘技术。

2 数据挖掘

2.1 概念 数据挖掘(data mining, DM)是指从一个大量的、随机的、不完整的、模糊的及有噪声的数据中提取隐藏在这些数据中未被人们事先获知但又潜在有用的信息和知识的过程^[6]。DM 在决策支持应用过程中起着非常重要的作用,它主要基于人工智能、机器学习、模式识别、统计、数据库、可视化等技术对现有数据进行自动分析,并利用归纳推理挖掘出潜在的、有价值的信息,从而辅助决策者作出正确的决策。

2.2 DM 常用的算法 DM 的常用算法有决策树、关联规则、粗糙集、神经网络、统计分析、支持向量机、基于事例推理、模糊聚类、可视化技术和贝叶斯预测。决策树和人工神经网络是辅助临床诊断中的经常用到的方法。其中,决策树是一种非常有效的机器学习分类算法。

决策树学习通常是从一组无序和不规则事例中

推测出决策树表示的分类规则。它采用自顶向下递归的方法,首先比较决策树内部节点的属性值,然后根据不同的属性值判断节点的向下分支,最后在决策树的叶节点得到最终结论。从根到叶节点的每个路径对应于一则连接规则,因此整个决策树对应于一组析取表达式的规则。决策树学习算法的最大优点是:在学习过程中,用户不需要了解对其太多的背景知识。它只需要知道怎么训练实例和用属性-结论的方式来表达就可以使用该方法来进行学习。本研究所采用的糖尿病健康教育挖掘算法就是决策树技术。

2.3 DM 的应用 首先建立糖尿病决策树,然后根据患者信息、临床表征等判断出患者所患糖尿病的类型,最后根据对以往同类患者的治疗方案及疗效等推算出最适合的健康教育方案。如果起点是高血糖,第一个分支决定是否是否妊娠,答案“YES”,决策树就直接归为妊娠期糖尿病,然后根据以往对妊娠期糖尿病的历史治疗方案及疗效推算出对当前患者最适合的健康教育方案;答案是“NO”,则将其归属于其他类型糖尿病,然后再根据的年龄、是否有自发性酮症、病因和病情轻重缓急等条件判断该患者是 1 型糖尿病还是 2 型糖尿病,再根据以往对该类型糖尿病治疗所取得较好疗效的健康教育方案,使用智能挖掘算法来找出最适合患者的方案。

3 糖尿病健康教育系统结构设计

3.1 设计思路 遵循着“在合适的时间对当前的患者提供正确的健康教育方案”的设计理念,利用数据挖掘技术对以往历史积累的海量糖尿病临床表征、健康教育方案以及治疗效果等信息进行分析与挖掘,为糖尿病患者提供科学智能的健康教育方案。

3.2 总体架构 系统总体框架结构分为 3 个层次:支撑层、分析层和应用层。①支撑层:由糖尿病患者数据和糖尿病健康教育知识库构成。其中糖尿病患者的数据主要包括患者的基本信息、主要自诉、临床辅助检查信息、临床诊断等;糖尿病健康教育知识库主要是针对不同的糖尿病提供与之相对应的健康教育方案,为系统提供数据来源和业务规则支撑。②分析层:对患者信息、临床表征、教育手段、疗效等数据进行清理、导入与规范化处理,并使用现代数据统计、挖掘技术,实施智能分析与挖掘。③应用层:根据分析层提供的智能分析与挖掘的结果,最后为患者寻找到辅助治疗效果最佳的糖尿病教育方案,并提高给患者,见图 2。

(下转第 5 页)

(上接第 2 页)

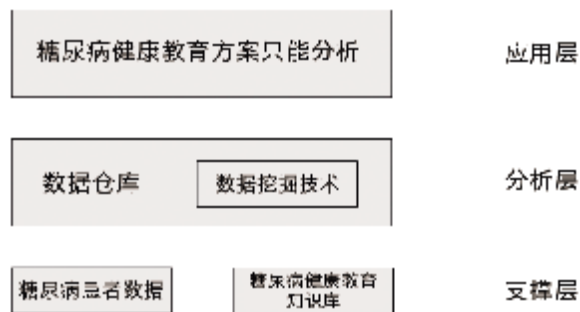


图 2 糖尿病健康教育系统智能分析与挖掘总体架构示意图

4 结论

糖尿病健康教育系统采用了当下较为先进的数据仓库及数据挖掘技术,实践表明:该系统能够从海量的糖尿病治疗的相关信息中提出有价值的信息,帮助临床医生快速准确地为患者提供最适合的健康教育方案,增强患者对糖尿病自我管理的能力,提高患者生活的质量,充分体现了“以病人为中心,以

健康教育为指导”的理念,从而为广大糖尿病患者带来了福音。

参考文献:

- [1]王一东,王庆梅.中医糖尿病健康教育对 2 型糖尿病患者血糖控制情况的影响[J].糖尿病新世界,2014(7):261-262.
- [2]刘莹.健康教育在糖尿病临床治疗中的实施效果分析[J].中国医药指南,2017,15(21):29-30.
- [3]宋伟,刘春,忻凌.基于真实世界的糖尿病健康教育数据仓库的设计[J].医学信息,2017,30(2):1-2.
- [4]李韬,冯佳莉,张屹,等.数据仓库技术在医院信息化建设方案实现中的应用价值[J].中国医药导报,2016,13(4):160-163.
- [5]赵妍,王颖,闫国涛,等.基于数据仓库的临床决策支持系统在我院的应用[J].中国医疗设备,2016,31(7):95-97.
- [6]李文莹,李文.数据挖掘技术在医院信息系统中的应用研究[J].中国管理信息化,2018(4):143-144.

收稿日期:2018-6-2;修回日期:2018-6-11

编辑/李桦