

基于电子病历的肿瘤患者临床路径变异因素的机器学习识别

刘一可

暨南大学附属第一医院病案科, 中国·广东 广州 510630

摘要: 论文针对基于电子病历的肿瘤患者临床路径变异因素机器学习识别展开论述, 先讲述了肿瘤患者临床路径变异因素识别的意义以及传统方法存在的不足之处。接着探讨了电子病历可给予大量数据的优势以及机器学习算法在处理繁杂数据和发掘潜在模式方面的能力。利用机器学习模型对电子病历里的肿瘤患者数据加以处理和分析, 进而识别出临床路径变异因素。通过研究得出结论, 机器学习可以有效地从电子病历当中找到影响肿瘤患者临床路径变异的因素, 给临床路径改良和肿瘤患者精准救治赋予依照。

关键词: 电子病历; 肿瘤患者; 临床路径变异因素; 机器学习; 识别研究

Machine Learning Recognition of Clinical Pathway Variation Factors in Cancer Patients based on Electronic Medical Records

Yike Liu

Medical Record Department of the First Affiliated Hospital of Jinan University, Guangzhou, Guangdong, 510630, China

Abstract: This paper discusses the machine learning recognition of clinical pathway variation factors in tumor patients based on electronic medical records. Firstly, it explains the significance of identifying clinical pathway variation factors in tumor patients and the shortcomings of traditional methods. Then, the advantages of electronic medical records in providing large amounts of data and the ability of machine learning algorithms to handle complex data and discover potential patterns were discussed. Using machine learning models to process and analyze tumor patient data in electronic medical records, in order to identify clinical pathway variability factors. Through research, it has been concluded that machine learning can effectively identify factors affecting clinical pathway variation in cancer patients from electronic medical records, providing guidance for clinical pathway improvement and precise treatment of cancer patients.

Keywords: electronic medical records; tumor patients; clinical pathway variability factors; machine learning; identification research

0 前言

肿瘤是危害人类健康的一种疾病, 治疗肿瘤的过程复杂且个体差异较大, 临床路径是一种标准化的医疗管理方式。它试图规范化肿瘤患者的治疗流程, 改善医疗质量和效率, 不过在临床实际应用过程中, 肿瘤患者的临床路径经常会出现变异情况, 造成肿瘤治疗效果不佳, 医疗资源被浪费等现象。所以, 精准识别肿瘤患者临床路径变异因素对于改进临床路径, 提升肿瘤治疗效果有着非常重要的意义。传统临床路径变异因素识别方法大多依靠医生的经验和主观判断, 存在识别效率低, 准确性差的问题。随着信息技术不断发展, 电子病历累积了大量肿瘤患者的临床数据, 给机器学习算法的应用供应了丰富数据来源, 机器学习算法可以从数据中自动学习模式和规律, 从而更准确地识别临床路径变异因素。

1 肿瘤患者临床路径变异因素识别的重要性与传统方法的局限

1.1 肿瘤治疗中临床路径变异因素的识别意义

肿瘤病人接受治疗时包含很多环节, 如诊断、手术、化疗、放疗等, 临床路径的制定目的在于给肿瘤病人给予标准, 规范的治疗计划, 保证病人能在恰当的时候得到恰当的治疗。不过, 因为肿瘤病人的个体差异, 病情复杂性, 医疗环境的不确定性等诸多因素, 临床路径在实际执行过程中常常会表现出变异, 这种变异也许会造成病人治疗延迟, 引发更多并发症, 降低治疗成效等等, 所以正确找出临床路径变异因素, 有益于医生及时调整治疗方案, 提升肿瘤治疗的精准度和有效率, 而且可以改善医疗资源的调配, 削减医疗花费。

1.2 传统临床路径变异因素识别方法的缺陷

传统的临床路径变异因素识别方法主要依靠医生的临

床经验和主观判断,医生根据患者的病历资料,凭借自身的专业知识,分析可能导致临床路径变异的因素,这种方法存在着明显的不足之处。一方面,医生的经验和知识水平各不相同,不同的医生对于同一个病例的分析结果可能存在差别,从而影响识别结果的准确性与可靠性。另一方面,传统方法难以应对大量的临床数据,无法充分而深入地挖掘出数据里潜藏的因素,而且,传统方法通常是事后分析,不能及时对临床路径变异予以预警并加以干预,这就影响了临床路径的执行效果。

1.3 电子病历为机器学习识别提供数据支撑

电子病历是患者在医疗机构就诊时所形成的电子化医疗记录,包含着患者的个人信息、病史、诊断结果、治疗方案以及检查检验报告等丰富的临床数据,这些数据具备全面性、即时性且可追查等特性,给机器学习算法应用给予了理想的资料根基。通过对电子病历数据加以挖掘并剖析,机器学习算法就能找出数据里的潜藏模式和规律,进而更为精确地判定肿瘤患者临床路径变异的因素,同传统的病历资料比较起来,电子病历数据更容易储存,管理并分享,给大规模的临床研究供应了方便。

2 基于电子病历的机器学习识别肿瘤患者临床路径变异因素的可行性

2.1 电子病历数据的丰富性与多样性特征

电子病历包含肿瘤患者从就诊到治疗结束的所有数据,患者基本信息、症状表现、影像学检查结果、病理诊断、治疗方案、用药情况、不良反应等,这些数据既丰富又具有多样性,能够反映出肿瘤患者间的个体差异以及病情的变化情况。而机器学习算法恰好能够利用这些数据,通过学习数据中的特点和模式来识别与临床路径变异有关的因素,以大量肿瘤患者的电子病历数据为例,其中包含着一些基因突变同某种治疗方案的临床路径变异之间存在联系,这样就可以为个性化的治疗提供一定的参考。

2.2 机器学习算法处理复杂数据的能力分析

肿瘤患者的临床数据有着高维度、非线性以及复杂的特点,这些特性用传统数据分析方法难以应对,机器学习算法具备很强的数据处理实力,它能应对高维度的数据,深入挖掘数据中的非线性联系和繁杂之处,像深度学习算法通过形成多层神经网络,自然而然地去学习数据里的数据,从而改进对复杂的数据的处理程度,在肿瘤患者临床路径变异因素识别当中,机器学习算法对电子病历中的各种数据实施分析,找到隐藏在背后的因素,加强识别的精确性和可信度。

2.3 机器学习在医疗领域的应用潜力

随着人工智能技术的突飞猛进,机器学习在医学上有着越来越多的应用,在医疗领域中的诊断、治疗选择以及预后评价等方面都取得了显著的成果,在肿瘤病人的临床路径的变异因素辨识方面,机器学习可以参考机器学习算法在其

他医学应用的经验不断地改善机器学习算法与模型以得到更好的辨识效果。机器学习也可以联合临床决策支持系统给予医生及时的临床路径变异警告及建议,协助医生做出更加准确和合理的医疗决策,从而提升肿瘤患者的临床治疗效果以及生存质量。

2.4 机器学习能挖掘潜在关联与未知变异因素

肿瘤病人临床路径存在多种变异因素,有些因素之间可能存在潜在联系,也有些因素是目前未知的未知因素,传统方式很难捕捉到这些潜在联系和未知因素,机器学习算法具备很强的模式识别和特征挖掘能力,可以从海量电子病历数据中找出变量间的复杂联系。例如,利用深度学习算法对肿瘤病人基因数据,临床指标数据以及治疗记录数据实施综合分析,也许会找到某些基因表达模式同某种治疗方案的疗效之间存在着某种潜在联系,此类联系在传统研究当中或许未曾被察觉。而且,机器学习还能够通过大量病例学习,找出一些稀有的,却会引发临床路径变异的未知因素,给临床大夫给予新的思路和研究方向,从而改善肿瘤病人的临床路径。

3 基于电子病历的肿瘤患者临床路径变异因素机器学习识别策略

3.1 推进电子病历数据规范化与标准化建设

在肿瘤患者临床路径变异因素识别里,电子病历数据是机器学习算法的关键基础,但当下电子病历数据存在格式不统一,内容不规范,编码混乱等情况。这对机器学习算法的精准度和可靠度造成了重大影响,所以提升电子病历数据的规范化,标准化创建非常迫切,政府,医疗机构要起主要作用,制订全面而详细的电子病历数据标准和规范,这些标准不但包含数据格式,各种不同种类的数据,如文本、图像、数值等。该怎样存储,怎样展现,而且还要包含数据内容,哪些信息是一定要记录的,哪些是可选记录的,怎样详细地记录等,编码方面也要形成统一的编码体系,保证各个医疗机构,不同系统之间的数据能够互相兼容和识别。数据采集过程中,医护人员需严格按照标准录入数据,从源头上保证数据的准确性和完整性,为保证标准的严格执行,要建立严格的数据质量监控体系。建立定期的数据审核机制,对数据的完整性和准确性以及一致性进行审核,及时清除数据中的错误和缺失值,对明显错误的数据进行修正,对缺失值使用合理的办法进行填充或者标注,不同医疗机构之间的电子病历数据存在壁垒,这影响数据的共享和利用,打破数据壁垒,让电子病历数据互联互通,共享,给机器学习算法赋予更全面,更丰富的数据来源。

3.2 机器学习模型选择与训练的优化策略

在基于电子病历的肿瘤患者临床路径变异因素识别任务里,选取恰当的机器学习模型是达成较好识别效果的关键,不同的机器学习模型有着各自独特的原理与适用场合,

决策树模型由于其直观易懂的特性,可以清晰地体现决策过程,适合用于处理那些具备明确分类规则的数据,支持向量机模型在处理高维数据以及非线性问题时表现不错,可以找到最佳的分类超平面,神经网络模型具备很强的非线性拟合能力,可以处理较为复杂的模式识别任务,针对回归问题而言,线性回归模型比较简单而且有效,适用于变量之间存在线性关系的情形,岭回归模型通过加入正则化项,化解了线性回归也许会出现的过拟合状况,在模型训练阶段,恰当地划分训练集和测试集十分关键。训练集用来训练模型,调整参数,测试集用来测试模型的泛化能力,使用交叉验证等方法能更好地测试模型的性能,防止因为数据划分不合理导致的测试偏差,还要对模型进行细致的调整,调整模型的参数,如神经网络的层数、节点数、支持向量机的核函数参数等,以此来提升模型的泛化能力和识别准确率,而且集成学习也是一种有效的提升模型性能的方法,把好几个不同的机器学习模型结合起来,用投票,加权平均等方法得到最后的预测结果,集成学习可以发挥各个模型的优点,减少单个模型的误差,从而进一步改善对肿瘤患者临床路径变异因素的识别效果,给临床给予更准确的决策根据。

3.3 临床专业知识与机器学习的融合应用

机器学习算法在处理数据和挖掘模式上有着很强的能力,但肿瘤患者临床路径变异因素识别时,临床专业知识是不可或缺的,医生经过多年的临床实践,对肿瘤患者的病情、治疗方案以及可能出现的变异情况有较深的认识,可以从大量的电子病历数据中筛选出哪些因素与临床路径变异有关,哪些因素可能被忽视,所以在机器学习识别过程中,要重视临床专业知识与机器学习的结合。一方面,医生要参与机器学习模型的设计与特征选择,在模型设计阶段,医生可以凭借自己的临床经验,对模型的结构和算法提出要求,使模型符合临床实际需要。在特征选择阶段,医生可以给出专业的建议,选取与临床路径变异有关的重要特征,排除掉与临床路径变异无关或者重复的特征,提升模型的效率和精准性。另一方面,机器学习算法的识别结果也要结合临床专业知识来加以解读和检验,机器学习算法可能会产生一些表面上看起来比较合理但实际上并不符合临床实际状况的识别结果,医生凭借专业知识就能够对这些识别结果作出判断并加以修正,从而防止出现误判和漏判的情况,通过临床专业知识同机器学习相结合的方式,可以充分发挥二者各自的优势,进而改善肿瘤患者临床路径变异因素识别的准确度和可靠性,给临床治疗给予更为科学和恰当的提议。

3.4 构建临床路径变异因素的动态监测反馈机制

基于电子病历的肿瘤病人临床路径变异因素识别不是

一次性完成的工作,而是一个持续性、动态的过程。因为肿瘤病人的病情总是在变化,治疗方案也会有所改变。所以,需要一套动态监测反馈的体系去及时了解肿瘤病人的临床路径变异情况。医院可以通过先进的机器学习算法对肿瘤病人进行动态监测,快速的对大量数据进行处理,及时发现肿瘤病人的临床路径变异情况,并及时找到肿瘤病人的临床路径变异因素。例如在某项指标有明显异常的时候,可以迅速找出并分析其可能存在的原因。同时,还要把这些信息及时反馈给医生及管理人员,并且反馈出的结果中还应该包括变异的具体信息以及可能的原因还有相关的建议措施等。医生以及相关的管理人员可以针对这些反馈信息及时调整治疗方案以及临床路径,从而让患者接受到最适合自己的治疗,而且对于临床路径变异因素进行持续跟踪分析也非常重要,通过持续跟踪变异因素的变化情况以及影响因素,可以为后续优化机器学习模型和临床路径提供参考,不断优化机器学习模型可以提升其对临床路径变异因素的识别能力,优化临床路径可以提升肿瘤患者的治疗效果以及医疗质量,建立动态监测与反馈机制可以实现对肿瘤患者临床路径的精细化管理,从而为肿瘤患者的精准治疗提供有力支持,使患者能够获得更好的治疗效果和生活质量。

4 结语

基于电子病历的肿瘤患者临床路径变异因素的机器学习识别有着重要的研究价值和应用前景,可以加强电子病历数据的规范化和标准化建设,选择和训练机器学习模型,重视临床专业知识和机器学习的融合,建立临床路径变异因素的动态监测和反馈机制等策略,从而提升肿瘤患者临床路径变异因素识别的准确度和可信度。不过,当前这个领域研究才刚开始,碰上数据品质,模型可解释性等难题,将来要深入探究,改进机器学习算法和模型,加强临床操作和研究的联系,推动基于电子病历的肿瘤患者临床路径变异因素机器学习识别技术发展,为肿瘤病人的诊治,改善医疗品质作出更大贡献。

参考文献:

- [1] 颜剑峰.临床路径结合电子病历在医院管理中的作用[J].中国管理信息化,2013.
- [2] 胡少轶.基于电子病历的临床路径系统设计与应用,2011,22(3):487.
- [3] 顾亚明,季宝轶,张连丰.临床路径管理对改进医疗质量的作用分析[J].中国卫生质量管理,2012,19(3):33-34.
- [4] 许彬.医院电子病历档案管理的应用策略[J].移动信息,2021(8):1-2.
- [5] 赵统堂.医院电子病历档案管理及共享策略探讨[J].兰台内外,2022(1):34-36.