

# 基于数据挖掘技术的卫健委档案管理研究

张丹玲

台州市疾控中心 浙江台州 318000

**摘要:** 研究旨在探讨基于数据挖掘技术的卫生健康委员会档案管理策略,以应对当前档案管理中存在的挑战与问题。通过深入分析卫生健康档案管理的重要性及数据挖掘技术在其中的潜在价值,我们致力于提升档案管理效能,同时对健康数据的使用与分析进行优化。接着,对数据挖掘在卫生健康档案管理方面的运用进行了详细研究,数据收集、整合、预处理、特征选取与提取,以及模型挑选与构建,皆为所涵盖的环节。最后,我们阐述了档案管理策略的优化和实践方法,涵盖数据挖掘结果解析及决策支持系统搭建两部分。该研究的目的在于为卫生健康档案管理提供基于数据的科学策略,提升卫生信息管理水平,优化决策支持系统。

**关键词:** 卫生档案管理;数据挖掘技术;数据分析

## 引言

在当今数字化时代,卫生健康档案管理成为促进医疗体系高效运作的核心要素。随着健康数据不断积累,数据挖掘技术作为一种强大的工具正逐渐受到关注。本研究旨在探讨如何运用数据挖掘技术优化卫生健康委员会档案管理策略,提高管理效率,实现健康数据的更深层次利用与分析。通过对卫生档案管理现状和数据挖掘技术在卫生领域的潜在价值进行综述,我们将深入研究数据收集、预处理、特征选择与提取,以及模型建立等关键环节。本研究旨在为卫生信息管理提供创新思路,推动卫生健康系统的数字化转型。

### 1. 背景概述

在当代社会,卫生健康委员会扮演着维护公众健康的关键角色。为了更有效地履行职责,档案管理成为卫生健康体系中不可或缺的一环。卫生健康委员会档案管理的重要性体现在信息的整合、决策的科学性和卫生服务的优化方面。然而,随着健康数据的急剧增加,传统的档案管理方式面临诸多挑战<sup>[1]</sup>。为了解决这些问题,数据挖掘技术作为一种强大的工具被引入,其在档案管理中的潜在价值逐渐凸显。

#### 1.1 卫生健康委员会档案管理的重要性

卫生健康委员会档案管理的核心价值并非仅在于保障医疗资料的完备性,关乎公共卫生政策科学制定与卫生服务精准提供。卫生健康委员会借助高效的档案管理,得以优化疾病流行趋势监测和资源规划,进而为公众提供更加个性化和全面的医疗保健服务。这一管理体系赋予了卫生工作者实

时数据支持,不仅满足了他们的需求,这为决策者赋予了科学依据,因此,整体卫生体系的运作效能得到提升。

#### 1.2 数据挖掘技术在档案管理中的潜在价值

医疗信息不断累积增长,传统档案管理方式无法应对信息处理需求。数据挖掘技术因其挖掘数据潜在规律与关联的特性而备受关注。数据挖掘技术在档案管理方面的应用,可以识别患者群体特征,揭示潜在疾病关联,预测健康未来走势。采用深度学习、机器学习等方法,数据挖掘技术有助于提升档案数据运用效率,为卫生健康决策提供更为科学的依据。

### 2. 目的与研究意义

卫生健康档案管理作为卫生健康体系中的重要组成部分,其效率直接关系到公众健康服务的质量和效果。本研究旨在通过引入数据挖掘技术,提高卫生健康档案管理的效率,并进一步优化健康数据的利用和分析,以实现更精准、个性化的医疗服务。

#### 2.1 提高卫生健康档案管理效率

当前,卫生健康档案管理面临着大量的数据积累、信息碎片化和流程不畅等挑战。传统的手工管理方式已经难以满足日益增长的信息处理需求。引入数据挖掘技术,可以通过自动化的方式对大规模医疗数据进行处理和分析,实现信息的快速整合和提取。这将大幅度提高档案管理的效率,使医护人员能够更专注于临床工作,为患者提供更及时、高效的医疗服务。

## 2.2 优化健康数据的利用和分析

健康数据蕴含着丰富的信息，包括个体的生理状况、医疗历史、生活习惯等。然而，传统的手动分析方法难以充分挖掘这些信息中的潜在规律。通过数据挖掘技术，我们能够更全面、深入地理解健康数据，发现隐藏在数据背后的模式和关联<sup>[2]</sup>。这不仅有助于及早发现患者的健康风险，还能为个性化治疗和预防提供科学依据，最终提高医疗服务的质量。

## 3. 卫生健康档案管理现状

卫生健康档案管理当前面临着来自信息量急剧增加、隐私安全、数据碎片化等多方面的挑战。本章将深入探讨这些挑战，并提出相应的解决方案。同时，将介绍数据挖掘技术在卫生健康领域的应用情况，包括数据挖掘的基本概念以及在档案管理中的具体应用案例。

### 3.1 挑战与方案

3.1.1 信息量急剧增加：随着医疗科技的发展，医疗数据呈爆炸式增长，传统的手工管理方式已无法胜任。解决方案包括引入信息技术，采用先进的数据存储和处理系统，以应对不断增长的数据流。

3.1.2 隐私安全：医疗数据的敏感性使得隐私安全问题变得尤为突出。加强隐私保护机制，采用先进的加密技术，建立健全的权限管理体系，是有效应对之道。

3.1.3 数据碎片化：卫生健康信息通常散布在多个系统和平台上，导致信息碎片化和难以整合。推行标准化数据格式，建设统一的信息平台，有助于整合各类数据，提升数据的一致性和可比性。

3.1.4 管理流程不畅：传统的手工管理流程往往效率低下，容易导致信息滞后。引入先进的信息管理系统和工作流程优化，能够提高管理效率，加速信息的流动和更新。

### 3.2 数据挖掘技术在卫生健康领域的应用

3.2.1 疾病风险预测：利用数据挖掘技术，可以分析患者的病历数据，识别患病风险的关键因素，并建立预测模型，从而实现对患者未来可能发生疾病的提前干预。

3.2.2 患者群体划分：通过聚类分析等技术，可以将患者分为不同的群体，有助于更精准地制定个性化的治疗方案和健康管理策略。

3.2.3 医疗资源优化：数据挖掘技术可以分析医疗服务的利用情况，识别病患就诊的规律，从而合理配置医疗资源，

提高医疗服务效率。

3.2.4 预防性措施制定：基于历史数据的挖掘，可以发现患者的特定健康趋势，有助于制定更有效的预防性措施，降低患病风险。

通过数据挖掘技术的应用，卫生健康档案管理能够更加深入地挖掘医疗数据的价值，提高管理水平，实现更科学的决策支持。未来，随着技术的不断发展，数据挖掘在卫生健康领域的应用前景将更为广阔，为卫生健康管理体系的升级提供更多可能性。

## 4. 数据挖掘技术在卫生健康档案管理中的应用

在卫生健康档案管理中，数据挖掘技术的应用涉及多个环节，包括数据的收集与整合、预处理、特征选择与提取，以及模型的选择与建立<sup>[3]</sup>。本部分详细探讨这些环节的具体操作和技术手段，旨在为卫生健康档案管理提供更深入的数据挖掘应用方法。

### 4.1 数据收集与整合

4.1.1 医疗记录：医疗记录是卫生健康档案中最基础、最重要的数据之一。通过整合病患的医疗历史、就诊记录、检验报告等信息，形成完整的病患档案。数据挖掘技术可应用于医疗记录的智能提取和整合，使得医护人员能够更全面地了解患者的健康状况。

4.1.2 健康统计数据：除个体医疗记录外，卫生健康档案还需要考虑整体的健康统计数据。这包括人口统计学信息、疾病发病率、医疗资源利用情况等。数据挖掘技术可以从这些统计数据中挖掘出潜在的关联性和规律，为决策提供更科学的依据。

### 4.2 数据预处理

4.2.1 数据清洗：医疗数据的质量直接关系到数据挖掘的效果，因此数据清洗是至关重要的步骤。去除重复数据、处理异常值和纠正错误是数据清洗的主要任务。通过去除冗余信息，处理异常值和错误，确保数据的一致性和准确性，为后续数据挖掘提供干净、可信的数据基础。例如，通过算法识别和处理异常生理参数，提高医疗数据的可靠性，为后续分析提供可信的基础。

4.2.2 缺失值处理：在医疗数据中，常常会出现部分数据缺失的情况，这可能影响数据挖掘的准确性。缺失值处理的方法包括使用插值方法填充缺失值，或者在分析中丢弃缺失值。选择合理的处理方法能够保持数据集的完整性和可靠

性，确保在数据挖掘过程中不受到缺失数据的影响。例如，通过基于相似患者的插值方法填充缺失的生理指标数据，提高数据的完整性和可用性。

#### 4.3 特征选择与提取

4.3.1. 重要健康指标的选择：在医疗数据中，识别对健康状况判定至关重要的指标是数据挖掘的关键任务。通过数据挖掘技术，能够准确找出影响患者健康的核心指标，为医护人员提供更有针对性的信息。例如，通过分析患者病历，数据挖掘可明确关键的生理参数或疾病标志物，帮助医护人员聚焦于关键信息的监测和管理。

4.3.2 特征工程的应用：在卫生健康档案管理中，特征工程是提高模型性能的有效手段。通过构建新特征，原始数据能够更好地反映患者的整体健康状态。例如，将医疗记录中的时间序列数据转化为趋势特征，有助于更准确地捕捉健康变化趋势<sup>[4]</sup>。特征工程的应用使得数据挖掘更为精细化，提高了模型对患者健康状态的准确性和预测性。

#### 4.4 模型选择与建立

4.4.1 机器学习算法的选择：选择适当的机器学习算法至关重要。在卫生健康档案管理中，可根据任务特性和数据分布选择合适的算法。例如，对于分类任务，决策树可提供直观的解释性；支持向量机适用于复杂的非线性关系；神经网络能处理大规模数据和复杂模式识别。综合考虑算法特点和需求，选择最适合的算法可提高模型效果。

4.4.2 模型建立与优化：在选择算法后，需建立并优化模型以获得最佳性能。首先，进行数据划分，分为训练集和测试集。然后，对模型进行训练，并通过调整参数来优化模型。交叉验证是评估模型泛化性能的重要方法，确保模型在新数据上表现良好。优化模型有助于提高实际应用效果，从而更好地支持卫生健康档案管理的决策和实践。

通过以上一系列数据挖掘技术的应用，卫生健康档案管理得以更全面、深入地分析患者的健康信息，实现对大数据的智能化处理，为决策提供科学支持。这一过程将有助于提高卫生健康管理的水平，为患者提供更加个性化、高效的医疗服务。未来，随着数据挖掘技术的不断发展，其在卫生健康档案管理中的应用前景将更为广泛。

### 5. 档案管理策略的优化与实践

在数据挖掘技术的支持下，卫生健康档案管理的策略能够得到有效的优化与实践。本章将深入探讨数据挖掘结果

的解读和决策支持系统的构建，以实现更智能、科学的档案管理。

#### 5.1 数据挖掘结果的解读

健康趋势分析：数据挖掘技术可以对患者的历史健康数据进行分析，发现潜在的健康趋势。例如，通过挖掘生理指标、生活习惯和疾病发病率等数据，可以识别出不同人群的健康变化趋势。解读这些趋势有助于制定更加精准的健康管理计划，提前预防潜在风险。

#### 5.2 决策支持系统的构建

5.2.1 基于数据挖掘的决策支持模型：在档案管理策略中，数据挖掘的结果需要转化为实际可操作的决策。建立基于数据挖掘的决策支持模型是关键一步。模型可以根据患者的个体特征和整体健康状况，提供个性化的医疗建议。这种模型可以是规则引擎、机器学习模型或深度学习模型，根据具体需求进行选择。

5.2.2 实际应用案例：为了更具体地说明决策支持系统的实际应用，考虑一个患有慢性疾病的患者群体。通过数据挖掘技术，系统分析了这一群体的医疗记录、生活习惯和遗传信息，发现了不同个体之间的疾病发展趋势存在差异。基于这一分析，建立了个性化的治疗建议模型。在实践中，医生可以根据患者的档案信息，系统生成的模型建议制定更为个性化的治疗方案，包括药物的选择、剂量的调整以及生活方式的调整等。

患者 A，通过数据挖掘分析显示其患病风险较高，但基于个体特征，系统建议采用更频繁的健康监测和个性化药物调整。而患者 B，由于其生活方式较为健康，系统建议强调保持健康生活习惯，并在日常生活中进行定期的锻炼。这样的实际应用案例充分展示了数据挖掘技术在决策支持系统中的实际应用效果。通过以上实际案例，卫生健康档案管理得以更为精细化和个性化，医疗决策更加科学合理，提高了患者的医疗体验和治疗效果<sup>[5]</sup>。同时，决策支持系统的不断优化也为医护人员提供了更科学、高效的工作手段，全面推动了卫生健康档案管理策略的实际落地和优化实践。

#### 总结：

本文深入探讨了数据挖掘技术在卫生健康档案管理中的应用。从数据收集与整合、预处理、特征选择与提取，到模型选择与建立，我们详细介绍了每个环节的操作和技术手段。同时，我们还探讨了数据挖掘结果的解读和决策支持系

统的构建,通过实际应用案例展示了数据挖掘技术在档案管理策略优化与实践中的价值。这些讨论为卫生健康领域的档案管理提供了科学、智能化的解决方案,为提高医疗服务质量、优化资源配置、改善患者健康状况提供了有效支持。

#### 参考文献

[1] 赵辉. 数据挖掘技术在档案管理中的应用研究 [J]. 南国博览, 2019(8):1.

[2] 甘璐. 基于数据挖掘技术的档案馆信息快速分析算

法研究. 现代电子技术 42.7(2019):3.

[3] 费文奇. 数据挖掘技术在档案管理中的应用研究. 兰台内外 27(2019):2.

[4] 索向峰. 数据挖掘技术在档案管理系统中的应用研究 [J]. 黑河学院学报, 2013, 4(2):113-115.

[5] 王静伦. 探讨数据挖掘技术在档案管理中的应用 [J]. 经济与社会发展研究, 2020(17):1.