

团体标准

T/CI 403-2024

中医智慧辨证论治模型构建技术规范

Technical specification for the construction of wisdom syndrome differentiation
model of traditional chinese medicine

2024-07-01 发布

2024-07-01 实施

中国国际科技促进会

发布

目次

前言 II

引言 III

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 基本要求 2

 4.1 目标明确性 2

 4.2 方法和流程规范 2

 4.3 数据质量和隐私保护 2

5 智慧中医辨证模型 2

 5.1 基于辨证要素的中医证型匹配策略 2

 5.2 基于特征加权与动态组合的证候分类策略 4

 5.3 融合知识图谱的多通道中医诊断策略 5

 5.4 基于四诊信息智能提取的中医辨证模型 5

6 智慧中医论治模型 6

 6.1 基于中医证候的法则治法适配策略 6

 6.2 基于深度交叉神经网络的处方推荐策略 6

 6.3 基于图数据模式匹配的方药检索策略 8

 6.4 融合个体特征的中医论治模型 8

7 中医诊疗优化模型 10

 7.1 基于序贯治疗的疗效预测模型 10

 7.2 基于疗效反馈的诊疗优化机制 10

 7.3 基于深度强化学习的诊疗优化方法 11

 7.4 中医疗效评价体系的构建 12

8 智慧辅助诊疗系统构建技术 13

 8.1 基于证型的中医处方推荐 13

 8.2 基于数据驱动的中医诊疗方法 14

 8.3 中医智慧辅助诊疗模型评估 14

9 模型测试 15

 9.1 测试目的 15

 9.2 数据采集测试 15

 9.3 数据预处理测试 15

 9.4 特征提取测试 16

 9.5 处方推荐测试 16

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由河南科技大学提出。

本文件由中国国际科技促进会归口。

本文件起草单位：河南科技大学、河南科技大学第一附属医院、同济大学、东南大学、河南省洛阳正骨医院（河南省骨科医院）、广州市香雪制药股份有限公司、慧医谷中医药科技（天津）股份有限公司、河南群智信息技术有限公司、洛阳理工学院、卢氏县中医院。

本文件主要起草人：朱军龙、吴庆涛、何良华、王琳、张明川、陈阳、王娟、康志英、董玉舒、张虹、郭珂珂、夏丽晔、王斌、赵亮、葛又铭、李美雯、王国勇、张娟梅、李明政。

本文件为首次发布。

引 言

当前社会医疗资源供给不足与人民群众日益增长的健康需求之间矛盾突出,且由于缺乏可用于研究的标准化数据集,跨科室、跨机构协同困难,导致医疗信息的共享性较低。利用大数据、物联网、人工智能、计算机视觉等理论和技术的兴起与发展,同时结合中医独特的诊疗方法,研究符合中医特色诊疗模式的智能辅助诊疗系统,不仅可以践行“互联网+智慧医疗”产业的发展政策,而且有望推动医疗服务走向真正意义的信息化、智能化和标准化进程。

为了拓展中医辨证论治模型构建技术的应用,实现中医辨证论治与信息技术紧密结合,同时顺应我国中医辨证论治的智慧化提升和升级改造需求,本文件制定了中医智慧辨证论治模型构建技术的规范。

中医智慧辨证论治模型构建技术规范

1 范围

本文件规定了中医智慧辨证论治模型构建的智慧中医辨证模型、智慧中医论治模型、中医诊疗优化模型、智慧辅助诊疗系统构建技术、模型测试等内容。

本文件适用于中医智慧辨证论治模型构建，也可作为智慧中医设计与研究的参考。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

证型库 syndrome database

一个集中存储和管理中医证候类型信息的知识库，其中包含各类临床中医证候的定义、特点、症状表现、治疗原则等相关信息。

3.2

关联异质图 heterogeneous association graph

用于描述患者个体特征、中药以及二者之间关系的图论表达形式，由多种类型的节点和不同类型的边组成；其中不同的节点表示患者个体特征和中药，边表示患者个体特征与中药之间的关系。

3.3

图卷积网络 graph convolutional network

一种处理图结构数据的深度学习模型，用于有效地捕获图中的节点关系和整体的网络结构信息。

3.4

多层感知机 multilayer perceptron

一种前馈人工神经网络模型，由一个输入层、一个或多个隐藏层以及一个输出层组成，层与层之间是全连接的，能够学习和模拟非线性和复杂的函数关系。

3.5

数据筛选 data filtering

一种根据特定的条件或标准，从数据集中选择出符合条件的数据子集的过程。

3.6

指标清单 indicators list

一种系统化的列表，用于描述和定义在特定领域或任务中需要测量和评估的各种指标。

4 基本要求

4.1 目标明确性

应明确规定本文件的目标和目的：确定模型的应用范围、目标疾病或病种，以及模型在临床实践中的预期效果和贡献。

4.2 方法和流程规范

应规范本文件的方法和流程：实现数据采集、特征提取、模型训练和验证等环节的规范，以向可靠性、准确性和可重复性方向发展为目标。

4.3 数据质量和隐私保护

应加强本文件构建过程中的数据质量控制和隐私保护：实现数据的准确性、一致性和标准化，以及采取适当的措施保护患者数据的隐私和安全。

5 智慧中医辨证模型

5.1 基于辨证要素的中医证型匹配策略

智慧中医辨证模型包括证型匹配策略、证候分类策略和中医诊断策略，通过动态提取四诊辨证要素，智能识别病位、病性及病势，并对辨病结果进行中医证候识别，从而确定主证、兼证。基于辨证要素的中医证型匹配策略可以抽象为“证素辨识”和“证型匹配”两步，具体如图1所示。

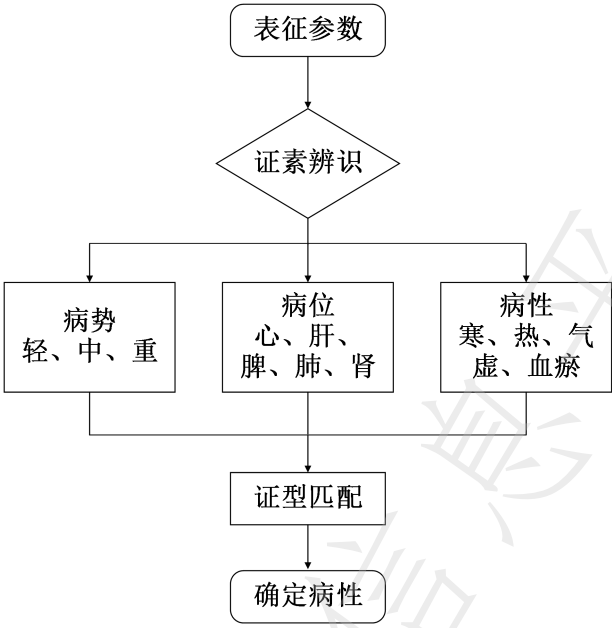


图 1 基于辨证要素的中医证型匹配原理

5.1.1 证素辨识

证素辨识如下：

- a) 症状收集：通过与患者交流、四诊（问诊、望诊、舌诊和脉诊）观察等方式收集患者的症状信息；
- b) 病势辨识：通过观察患者的症状表现、脉象等，判断疾病的轻重缓急程度，如轻症、重症等，即确定病情的严重程度和病势的走向，帮助医生做出合理的治疗决策；
- c) 病位辨别：通过对患者的症状进行分析和综合判断，确定疾病发生的部位或脏腑组织，如头面部、胸腹部、四肢等，常见的脏腑组织包括心、肝、脾、肺、肾等；
- d) 病性鉴别：综合患者的主观症状和客观体征，对疾病性质进行判断，如寒热虚实、气滞血瘀等。

5.1.2 证型匹配

依据证型拆解出的证素即可形成证型匹配规则，如表 1 所示；证型匹配的具体步骤如下：

表 1 证素-证型匹配对照表

序号	证素	证型
1	心、血瘀	心血瘀阻证
2	痰、热、脾	痰热蕴脾证
3	痰（湿）、气虚、脾	脾虚痰湿证
4	心、肾、阴虚、阳虚	心肾阴阳两虚证

- a) 证型分类：将已知的各种中医证候类型进行分类和整理，建立起证型库，包括各类证候的定义、特点、症状表现和治疗原则等；
- b) 证素对比：对收集到的患者的证素信息进行归纳和整理，找出其涉及的基本证素，如寒热、湿燥、气滞、血瘀等。然后将基本证素与证型库中的各种证候特点进行对比，找出与患者症状相匹配的证型；
- c) 证型确定：通过综合判断确定最符合患者症状的证型，并作为辨证施治的依据。

5.2 基于特征加权与动态组合的证候分类策略

针对中医电子病历文本数据中的重点症状进行权重加权，并且对四诊信息的重要字段进行动态组合。该策略结构如图 2 所示，主要包括以下部分：

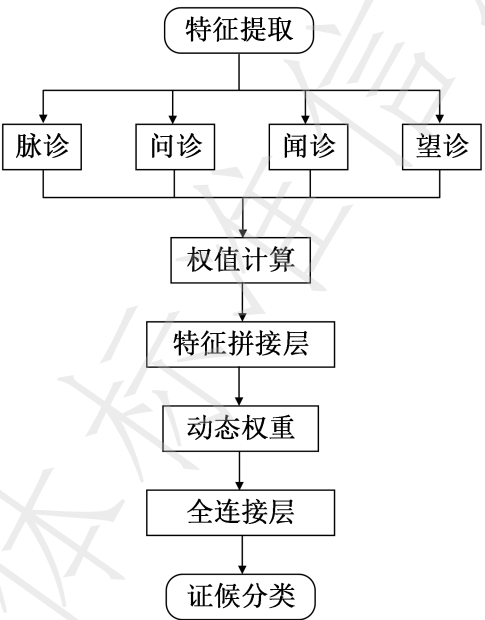


图 2 基于特征加权与动态组合的证候分类策略

5.2.1 特征加权

通过将中医电子病历四诊语料中望诊、闻诊、问诊、脉诊的数据分别输入到子网络中，并对四诊字段分别进行关键特征的提取并赋予权值，使得模型能够较快地从大量信息中筛选出高价值信息。

5.2.2 动态组合

根据中医电子病历中望诊特征、闻诊特征、问诊特征、切诊特征的文本字段的重要程度提供对应的权值，形成各字段的动态组合。不同的特征类型对于模型存在不同的定量影响，结合不同特征的动态加权可以提高模型分类的效果。

5.2.3 融合与预测

将经过子网络动态组合后的特征向量输入一个全连接层中，进行汇合与证候分类，通过不同特征的

信息融合实现对于患者证候的分类预测。

5.3 融合知识图谱的多通道中医诊断策略

5.3.1 多通道特征提取

采用多通道方法对中医电子病历中丰富的四诊信息进行文本处理，设置四个通道，对输入症状进行多种特征表示，学习不同的症状特征。

5.3.2 知识图谱嵌入

- a) 图谱构建：实体抽取后得到八类实体和十五类关系，首先通过脉诊表现、舌诊表现、闻诊表现、望诊表现、查体表现、证候结论表现、护理宜忌表现、中医诊断表现来对图谱中八个类别的实体进行汇聚；然后在已有八类实体基础上结合关系：脉诊-证候结论、舌诊-证候结论、闻诊-证候结论、望诊-证候结论、查体-证候结论、证候结论-护理宜忌、证候结论-中医诊断，使用 Neo4j 构建知识图谱；
- b) 图谱嵌入：选用 ComplEx 模型作为知识图谱嵌入模型，将知识图谱中的实体和关系进行向量表示，主要用于补全知识库的知识，也可用于知识问答、推荐、语义检索、文本信息增强。

5.3.3 模型融合与预测

将多通道特征向量和知识图谱嵌入向量融合后输入一个全连接层中，进行算法的学习与预测，如图 3 所示。使用随机加权平均算法进行模型的集成与优化，提高模型的泛化能力。

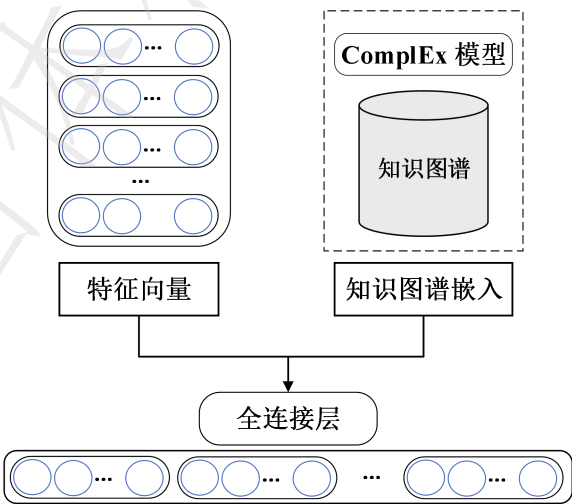


图 3 知识图谱融合示意图

5.4 基于四诊信息智能提取的中医辨证模型

5.4.1 望诊信息的智能化提取

可以使用舌面象仪、便携式面诊仪、眼象健康成像仪等设备获得高质量图像，然后采用图像分割模

型获取感兴趣部位，再针对关键部位基于知识规则设计特征或利用机器学习进行特征选择。

5.4.2 闻诊信息的智能化提取

智能提取设备包括中医闻诊仪、嗅诊仪等，其中声诊技术可以采用传统的小波变换或深度学习中的卷积网络等方式进行特征提取，通过分析声音信号中的频率、振幅等特征，以及声音信号的时域和频域特征，来获取与患者病情相关的信息。

5.4.3 问诊信息的智能化提取

问诊信息的分类诊断模型采用极限学习机、症状关联网络、深度森林等进行建模，其中深度森林算法在多标记评价指标和单个证型的分类准确率方面占有优势，能够有效解决中医问诊证候分类问题。

5.4.4 脉诊信息智能化提取

采用支持向量机提取脉诊信息特征，通过有效的构建超平面来对不同类别的信号进行准确分类。

6 智慧中医论治模型

6.1 基于中医证候的法则治法适配策略。

6.1.1 数据处理

读取中医证候文本数据集并将其转换为统一格式，同时对中医证候治法数据集进行切词、停用词等数据处理，最终形成预处理的语料库。

6.1.2 特征筛选

利用语料库建立词汇共现矩阵，然后基于共现矩阵和算法学习中医证候治法文本词向量矩阵，并将其输入到卷积层，使用不同尺寸的卷积核对输入的证候治法文本进行关键特征筛选。

6.1.3 特征提取

通过激活函数的非线性能力对特征进行保留并映射，得到证候治法数据集的关键特征，在池化层利用最大值采样提取中医证候的关键特征。

6.1.4 智能诊断

将池化层得到的中医证候向量传递到全连接层，根据输出结果预测治法类别，得到中医治法诊断结果。

6.2 基于深度交叉神经网络的处方推荐策略

基于深度交叉神经网络的处方推荐策略将处方特征与病案信息特征进行自动组合，得到最佳的组合

特征，从而实现处方推荐功能，结构如图 4 所示。

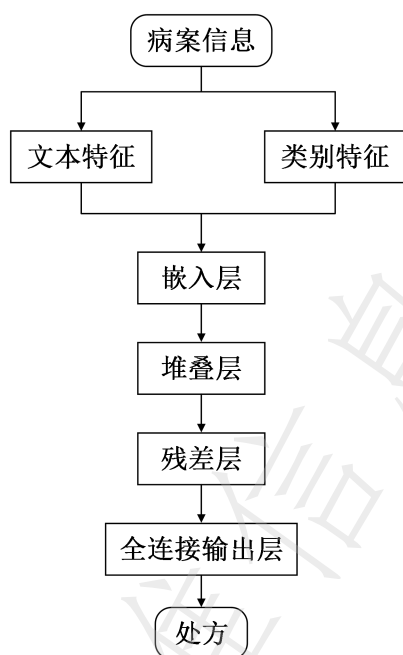


图 4 基于深度交叉神经网络的处方推荐策略结构

6.2.1 特征层

将输入的病案信息和处方信息进行特征表示，其中病案信息分为文本信息和类别信息，使得输入信息能够更好地进行特征提取以及向量表达。

6.2.2 嵌入层

对划分好的患者主诉、现病史、既往史、中医诊断结果等文本特征和科室、性别、各项数值指标等类别特征分别进行 word2vec 和 one-hot 嵌入，使得稀疏高维向量转为稠密低维向量。

6.2.3 堆叠层

将嵌入特征和部分原特征进行合并形成处方推荐的组合特征，并将其输入到下一层次。

6.2.4 残差层

主要结构是多个残差块，利用多层次残差网络将病案信息和处方信息的组合特征矢量的各个维度进行充分的交叉，提高模型的表达能力。

6.2.5 输出层

将残差操作后的组合特征与输入的原始组合特征进行相加，从而更好的将处方特征和病案信息特征进行交叉，同时最大限度地对原始组合特征进行特征保留，从而提高处方推荐的准确率。

6.3 基于图数据模式匹配的方药检索策略

基于图数据模式匹配的方药检索策略严格遵守“观其脉证，知犯何逆，随证治之”的原则，流程如图 5 所示。

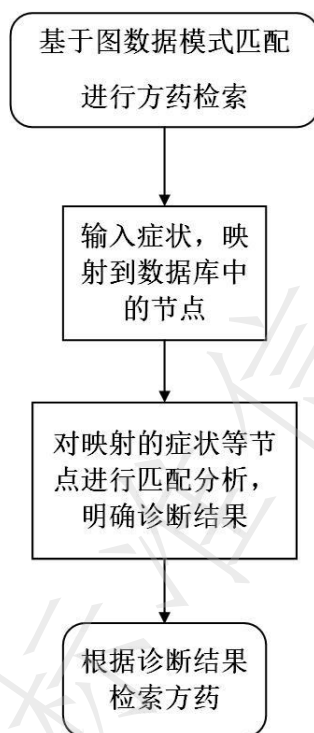


图 5 基于图数据模式匹配的方药检索流程

6.3.1 观其脉证

将输入症状与图数据库中的症状进行模式匹配，通过检索输出可能的候选诊断，模拟医生将收集到的病例信息与记忆中的医学知识进行匹配，初步分析病情的思维过程。

6.3.2 知犯何逆

通过计算候选诊断所对应的症状与输入症状的相似度，对候选诊断进行筛选，判断两者是否契合，以此模拟大脑对“症”与“证”之间契合度的联想与认知。

6.3.3 随证治之

通过检索输出相应的方剂和药物结果，并在确立主证对应主方的基础上，根据兼证的病机特点与主方中已有的药物进行加减化裁，达到最佳的治疗效果。

6.4 融合个体特征的中医论治模型

患者个人属性差异会导致在患有相同疾病的情况下可能有差异化的症状表现，即使是同一症状其药

方也因个体属性特点不同而存在异同。针对该问题，本文件提供如图 6 所示的融合患者个体特征的中医论治模型，主要包括以下部分：

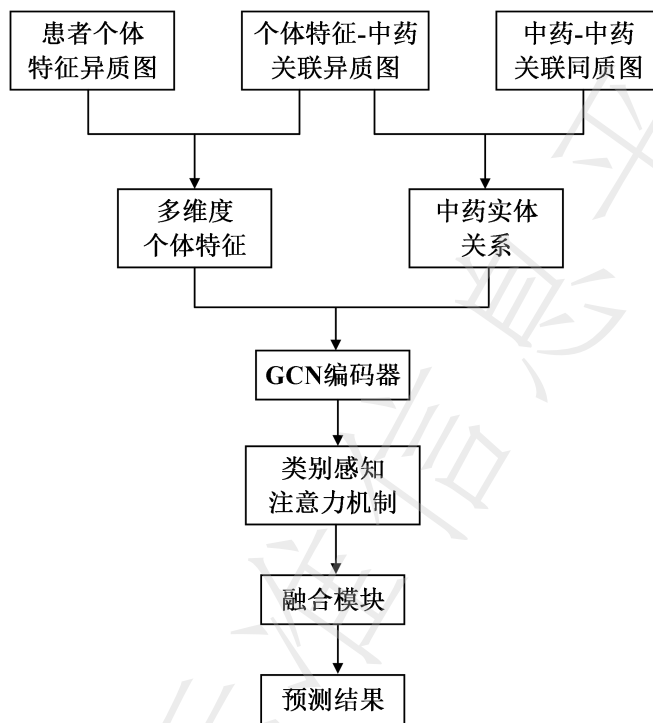


图 6 融合个体特征的中医论治模型结构

6.4.1 患者个体特征-中药关联建模

从患者个体化特征的角度出发，构建多维度个体特征与中药之间的关联异质图。

- 首先需要基于临床数据搭建患者个体特征异质图，患者个体特征-中药关联异质图，以及中药-中药关联同质图。
- 将患者的特征划分为三个维度，个人属性（包括年龄和性别）、症状表现（包括症状和证候）和诊断（包括中西医体系中的疾病诊断）。
- 根据医案中的共现信息，模拟中医基于用户个体特征、症状和诊断，关联到具体的药方组成中药集合，借助异质图描述这种关联映射关系。

6.4.2 构建多图卷积神经网络

针对异质图采用多类型感知的注意力机制来区分不同类型的节点对患者个体特征和中药实体表示学习的影响，多类型即图中对应的多种节点类型，如患者个体特征异质图中的症状、诊断与个体特征。在消息传递机制上，为每个图上的节点存储一个类型的属性信息，在实际的数据预处理过程中，为每一个节点专门存储单跳可达节点的信息，所有节点的邻居信息集合即存成邻接矩阵，以此来表示完整的图关联边集。

6.4.3 特征融合-预测

使用多层感知机（MLP）融合以上特征的嵌入来表示患者特征及中药特征，并将融合后得到的患者及中药的整体表示作为模型预测函数的输入，最终得到预测药方的向量表示，基于临床数据中的实际中药药方进行对比，中药集合推荐任务的准确度较高。

7 中医诊疗优化模型

7.1 基于序贯治疗的疗效预测模型

7.1.1 整理中医临床症状原始数据

按照发生的时间先后顺序，整理出病患的临床症状的状态时间序列。

7.1.2 统计状态

整理出状态空间，进行状态划分，确定由中医临床症状构成的时间序列的观察值状态，统计出每一个中医临床症状的频数。临床症状的状态有“有”和“无”两个状态，统计出每一个时间序列的状态个数。

7.1.3 数据筛选

从时间序列中筛选出符合预测条件的中医临床症状序列进行预测，在临床实践中帮助医生根据患者的病情和治疗历史制定更加个性化和有效的治疗方案。

7.2 基于疗效反馈的诊疗优化机制

在这个机制中，通过建立健全的医学数据库，将患者的病历资料、医学影像、实验室检查结果等数据整合在一起。流程如图 7 所示。

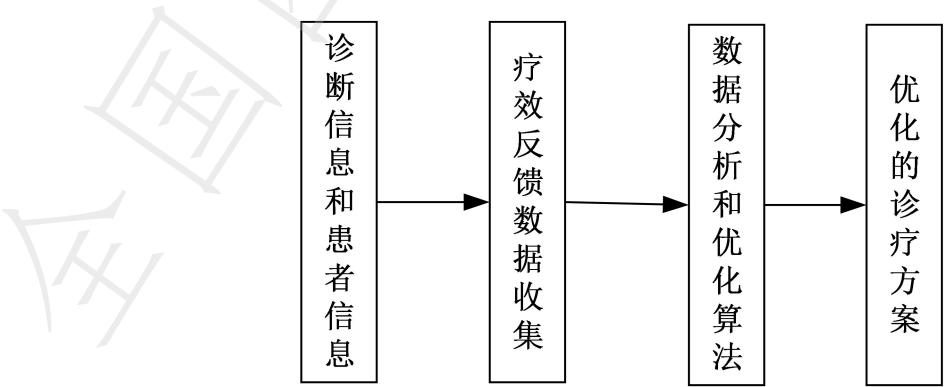


图 7 基于疗效反馈的诊疗优化方法

7.2.1 数据收集

建立健全的医学数据库，将患者的病历资料、医学影像、实验室检查结果等数据整合在一起。

7.2.2 数据挖掘和分析

通过数据挖掘和分析技术，从大量的医学数据中发现患者的疾病特点和治疗反应。

7.2.3 制定治疗指南和规范

根据数据分析的结果，可以制定相应的治疗指南和规范，为医生提供诊疗的参考依据。

7.2.4 诊疗方案优化

医生在诊断和治疗过程中，可以根据患者的具体情况，参考相关的治疗指南，并结合疗效反馈进行调整和优化治疗方案。

7.2.5 实时监测和健康管理

通过实时监测患者的生理参数和治疗效果，医生可以随时了解患者的病情变化，及时调整治疗方案，避免因治疗不当而导致的健康风险。

7.3 基于深度强化学习的诊疗优化方法

针对中医诊疗优化问题，通过设计离线虚拟环境、奖励函数、状态转移预测以及诊疗方案聚类分析，构建出一个基于深度强化学习的中医动态诊疗方案优化模型。将患者的症状和诊疗优化模型优化输出的诊疗方案作为输入，构建多层神经网络用于预测中药处方；最终将优化输出的诊疗方案和预测的中药处方作为治疗方案推荐给患者。具体流程如图 8 所示。

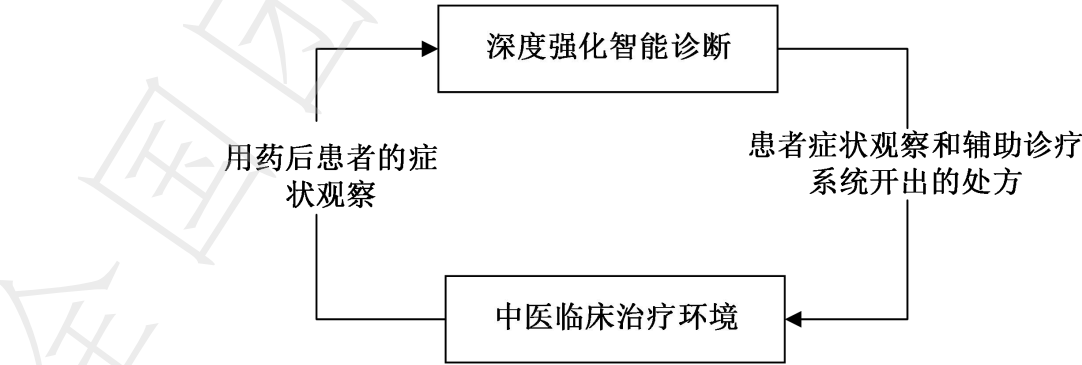


图 8 基于深度强化学习的诊疗优化流程

7.3.1 诊疗方案优化模块

通过结合 Q-Learning 和卷积神经网络的强化学习模型、通过融合全连接网络和 LSTM 神经网络的

强化学习模型进行诊疗方案优化模型训练，得到最佳的患者诊疗方案。

7.3.2 处方推荐模块

输入诊疗方案，构建多层神经网络用于预测中药处方，最终将优化输出的诊疗方案和预测的中药处方作为治疗方案推荐给患者。

7.4 中医疗效评价体系的构建

中医疗效评价体系的构建是为了科学客观地评估中医治疗的效果，以指导中医临床实践和提高治疗效率，该体系的流程图如图9所示。

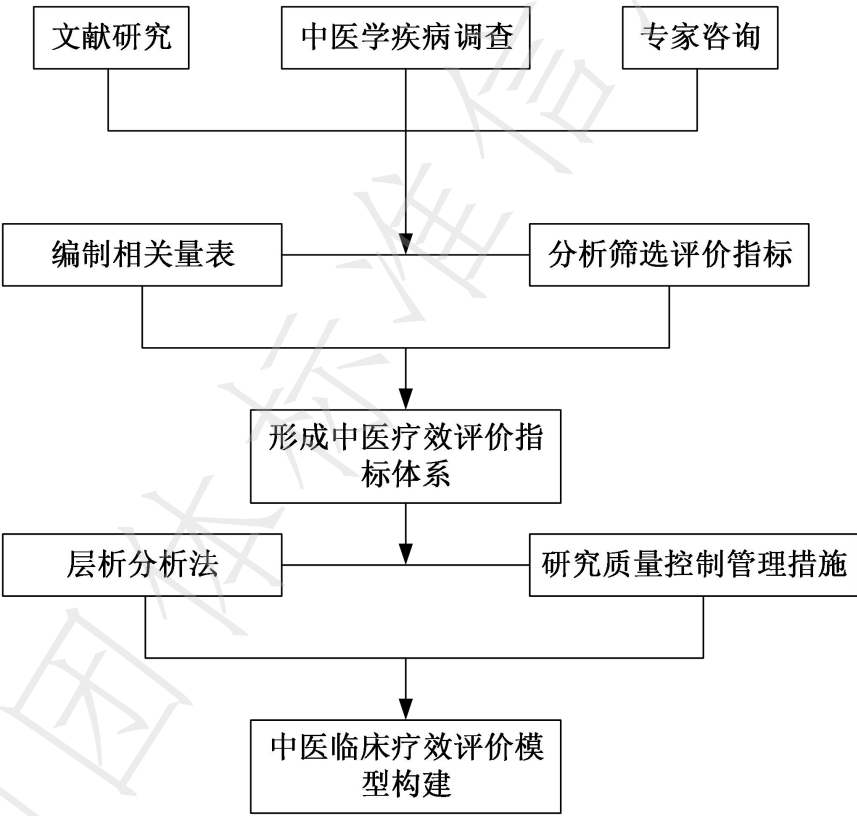


图9 中医疗效评价系统

7.4.1 确定指标清单

通过聚类分析统计临床常见的基本证候要素及症状、体征和四诊信息，形成病-证-症的结构框架，进而形成中医证候指标清单。

7.4.2 建立指标集

基于前期形成的中医证候指标清单，经过咨询-反馈-统计分析过程，对指标条目进行取舍，最终确定中医证候和疾病评价指标集。

7.4.3 确定指标权重

采用层次分析法，基于前期形成的中医证候和疾病评价指标集，构建由目标层、维度层、指标层组成的多层次结构模型，对各层次中的指标，分别建立成对比较矩阵，计算各指标权重，并进行一致性检验，最终确立各评价指标的权重系数。

8 智慧辅助诊疗系统构建技术

8.1 基于证型的中医处方推荐

中医辨证论治诊疗体系先采集病例数据并进行数据处理，提取病例中的关键因素，然后利用中医理论判断疾病证型，结合证型诊断思想和医生开放经验，得出具体的治疗方案，开出处方。中医通过望闻问切分析诊断后得到的病例数据后，将数据挖掘算法应用到这些数据上，建立模型得出结果并分析其中有价值的信息，即症状与证型、症状与药物、药物与用药剂量、药物和用药剂量与方剂等之间的关系。诊疗流程如图 10 所示。

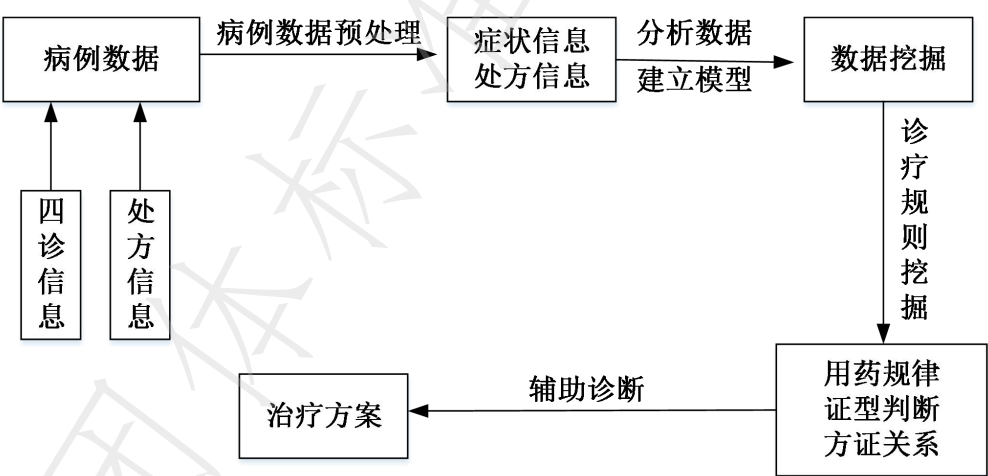


图 10 基于数据挖掘的中医诊疗模型

8.1.1 用药规律

采用数据挖掘技术对方剂数据进行分析，分析某种药物经常与哪些其他药物相结合，探究药物特性与方剂功效之间的关系，使用算法生成和归纳出一些有特定疗效的常用药物组合。

8.1.2 证型判断

筛选大量的症状信息，从症状指标数据中进行特征提取，获得有效信息，并对病情机理进行分析，通过统计医生的诊断结果建立的模型进行模拟判断，预测疾病证型。

8.1.3 方证关系

使用数据挖掘技术进行研究，从高维海量的诊疗数据中归纳出症状与药物的关系。

8.2 基于数据驱动的中医诊疗方法

8.2.1 数据采集与预处理

采集中医病、证的概念、定义、症状、体征、治法和方剂等信息。对采集的数据进行数据清洗和数据标准化，去除重复、错误或不一致的数据，将不同来源、不同格式的数据统一到相同的标准下，组建中医药多源异构数据集。

8.2.2 分词

采用分词工具对中医文本数据进行分词，实现文本数据从非结构化到文本向量的转变。

8.2.3 特征提取

提取文本向量中的关键词特征，并计算特征权重，并对提取计算结果进行评估选择。

8.2.4 特征选择

结合疾病的症状特征、各个疾病对应各个证型的症状特征以及其特征权重，通过特征选择和加权计算的方式，量化辨病和辨证的过程，从而实现辨证论治的客观化，得出精准的辨证论治结果。

8.2.5 构建模型

采用中医病证结合诊疗模式思想，对疾病、证型概念和症状表现进行特征提取和权重计算，得到疾病-特征相关关系和疾病-证型-特征相关关系，结合疾病-证型-治法-方剂联系分别构建中医辨病模型、中医辨证论治模型，最后结合中医诊疗特征相关性计算方法构建基于数据驱动的中医精准诊疗模型。

8.3 中医智慧辅助诊疗模型评估

8.3.1 确定评估指标

选择评估目标，如评估模型在处方推荐、治疗建议等方面的性能，根据评估目标，选择合适的评估指标，如准确率、召回率等，量化模型的性能。

8.3.2 数据收集与预处理

收集中医临床数据、挖掘中医古籍数据并通过数据挖掘从互联网上抓取与中医相关的文本、图片、视频等信息，提取出与特定疾病相关的症状、证候、治疗方法等信息。对数据进行清洗与预处理，去除重复数据、处理缺失值和异常值、标准化数据格式，并根据模型需要，对数据进行适当的特征提取和转换。

8.3.3 模型训练

使用经过预处理的数据对中医智慧辅助诊疗模型进行训练，根据模型的特点和评估的目标，设置合适的训练参数。

8.3.4 模型验证

将数据集划分为训练集、验证集和测试集，使用训练集训练模型，使用验证集调整模型参数和选择最佳模型，使用验证集评估模型的性能。通过比较模型在测试集上的预测结果与真实结果，计算评估指标的值，以评估模型的性能。

8.3.5 模型优化

根据模型验证的结果，对模型进行优化和调整，通过改进模型结构、调整参数、引入新的特征等方法来提高模型的性能。

9 模型测试

9.1 测试目的

对中医智慧辨证论治模型构建过程中数据采集、数据预处理、特征提取以及处方推荐进行证实，确定与本文件要求内容的符合性。

9.2 数据采集测试

9.2.1 测试方法

数据采集测试方法如下：

- 测试系统能够正确接收、存储和展示中医临床数据，如患者信息、诊断结果、治疗方案等；
- 验证系统能够有效管理中医处方数据，包括药物名称、用量、疗程等；
- 系统能按照预设的规则验证数据的合法性和完整性，如药物名是否存在、用量是否合理等。

9.2.2 测试指标

数据采集和记录应符合标准的比例应高于 98 %，数据记录的可追溯比例应高于 98 %。

9.3 数据预处理测试

9.3.1 测试方法

数据预处理测试方法如下：

- 测试经过数据清洗后数据中没有的错误、缺失或不一致的部分，确保数据准确性和完整性；
- 验证系统对数值型数据进行规范化处理，是否将其缩放到特定的范围内，如归一化到 0 到 1 之

间或者标准化成均值为 0，标准差为 1 的分布；

- c) 系统可以将来自不同数据源的中医临床数据整合为一个统一的数据集，以便于后续分析和建模使用。

9.3.2 测试指标

数据缺失值比例应小于 5%，重复数据比例应低于 1%，不同数据源之间数据一致性应高于 95%。

9.4 特征提取测试

9.4.1 测试方法

特征提取测试方法如下：

- a) 将提取的特征应用于机器学习模型中，进行交叉验证以评估模型的泛化能力和性能。验证特征提取的质量和效果，确保特征能够在实际应用中提升模型的预测能力；
- b) 使用统计方法如方差分析、相关系数、卡方检验等，评估特征与目标变量之间的关系；
- c) 使用数据可视化技术如散点图、箱线图、热力图等，对提取的特征进行可视化分析，观察其分布和与目标变量之间的关系。

9.4.2 测试指标

选择与目标变量高度相关的特征应达到 90 %，具有多重共线性的特征的比例应小于 5 %。

9.5 处方推荐测试

9.5.1 测试方法

处方推荐测试的方法如下：

- a) 根据患者的病史、体格检查和辅助检查结果，进行中医诊断，确定病名和证候。根据中医理论，进行辨证施治，即根据辨证论治的原则，确定治疗的重点和方向；
- b) 根据中医诊断和辨证施治的结果，选择适当的中药药材或中成药。根据中药药性、归经相克相宜的原则，进行药物的合理搭配和配伍；
- c) 将制定的处方提交给有经验的中医专家或团队进行审核。

9.5.2 测试指标

症状改善率应在 80 %以上，患者到处方的满意度应高于 90 %，推荐的处方的毒副作用应为 0 %。