

基于全域医疗大数据的精准预警监管系统研究与实践^{*}

杨丽静

陈育庆

徐 旭

(杭州市卫生健康事业发展中心
杭州 310002)

(杭州市卫生健康综合行政执法队
杭州 310021)

(杭州市卫生健康事业发展中心
杭州 310002)

廖维维

蒋奉锦

(杭州市卫生健康综合行政执法队 杭州 310021)

(杭州浩洋科技有限公司 杭州 310005)

〔摘要〕 从总体框架、数据层架构、服务层模式等方面介绍基于全域医疗大数据的精准预警监管系统构建, 阐述其应用实践, 指出该系统应用能够提高数据效能、监管效率以及执法精准性, 为相关研究提供参考。

〔关键词〕 诊疗行为; 预警; 精准

〔中图分类号〕 R-058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673-6036.2022.09.013

Study and Practice of an Accurate Early Warning and Supervision System Based on Regional Medical Big Data YANG Lijing, CHENYuqing, XU Xu, LIAOWeiwei, JIANG Fengjin

〔Abstract〕 The paper introduces the construction of an accurate early warning and supervision system based on regional medical big data from the aspects of overall framework, data layer architecture, service layer mode, etc., expounds its application practice, points out that the application of this system could improve data efficiency, regulatory efficiency and enforcement precision, providing references for related study.

〔Keywords〕 diagnosis and treatment behavior; early warning; accurate

1 引言

2019 年 9 月国务院出台《关于加强和规范事中事后监管的指导意见》(国发〔2019〕18 号)^[1], 提出要创新和完善监管方式, 深入推进“互联网+监管”。利用数据可视化、机器学习、模型预测、并行计算等大数据技术, 加强对风险防控动态预警管理。探索推行以远程监管、移动监管、预警防控为特征的非现场监管模式, 可以加强监督管控力度,

〔修回日期〕 2022-06-13

〔作者简介〕 杨丽静, 副研究员, 发表论文 9 篇; 通信作者: 陈育庆, 高级经济师。

〔基金项目〕 浙江省医药卫生科技计划项目“基于全域医疗数据的‘互联网+监管’技术应用研究”(项目编号: 2021KY269)。

提升监管规范化、精准化、智能化水平。之后各地、各部门尤其是市场监管部门开始开展“互联网+监管”探索和实践。四川省率先在全国范围内建立“医疗三监管”模式，建设具备医疗卫生机构、医务人员和医疗行为记录数据监管功能的政府数字化监管平台^[2-3]，充分整合医政医管、妇幼保健、基层卫生、综合监督等管理部门相关数据，采用大数据处理技术进行清洗、关联分析和深度应用，实现多角度、多维度分析，高效、准确定位医疗机构、医务人员及医疗行为情况，动态、实时、可视化地呈现监管结果。浙江省玉环市卫生监督所建设“智慧卫监”医疗机构诊疗行为预警云计算平台^[4]。该平台整合卫生监督协管系统、国家医师注册信息系统、基层医疗机构医院信息系统（Hospital Information System, HIS）等多方数据，围绕“大数据+风险预警”机制开发抗菌药物处方异常预警、诊疗科目超范围预警、处方异常预警、医师空挂预警、医护注册预警等功能，初步实现医疗卫生精准执法。上述相关系统数据获取较单一，仅对部分公立医疗机构及民营医疗机构进行监管，且无法实现跨场景应用。杭州市全民健康信息平台全域医疗大数据精准监管预警系统率先采用，与杭州市医保系统进行数据对接，实现对全市 80% 以上、近 5 000 家医疗机构进行实时在线智能监管的跨场景应用。本文针对该系统构建和实践进行研究。

2 精准预警监管应用系统构建

2.1 设计思路

2.1.1 系统业务架构（图 1） 包括数据交换层、数据中心层、应用层 3 部分。数据交换层通过 Http 接口、Socket、Web Service、数据导入、数据库读取等方式进行数据交互。数据中心层可跨场景采集杭州市医保、全民健康信息系统、医师注册信息、医疗机构注册信息等业务系统的诊疗行为数据，包括定点医疗服务机构、人员就诊登记、项目明细、手术、住院医嘱、病案主页、检查单、检验单等信息，并进行交叉分析、业务数据对比、关键字搜索、预警等级划分等数据处理，将预警结果推送至浙江省“互联网+监管”平台。应用层供医疗机构等监管单位进行快速预警数据查看，从而为监

管工作提供快速、精准的事实数据依据并保留处理记录。

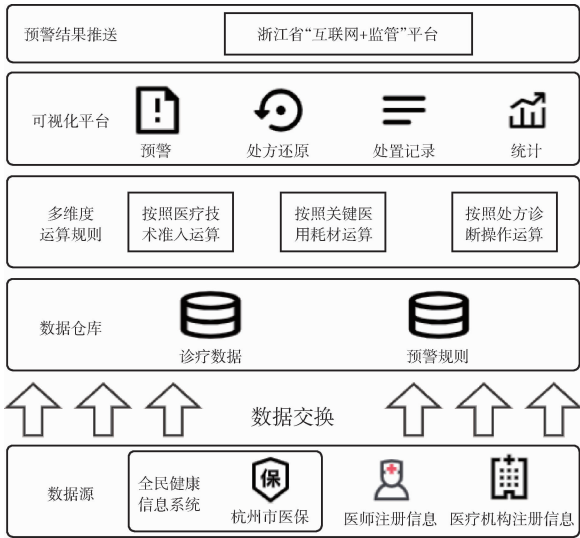


图 1 系统业务架构

2.1.2 开发环境 系统采用 Window + Eclipse + Vue + Java + Web + Android + Mysql + Hadoop + Redis 开发环境。Vue 是具有轻量级、双向数据绑定、指令、组件化等特性的渐进式框架^[5]，通过简单化应用程序接口（Application Programming Interface, API）实现响应的数据绑定和组合的视图组件。安卓（Android）是一种基于 Linux 平台的开源操作系统，应用于移动设备，具备较为优秀的交互界面以及交互逻辑，同时其较为开放的特性给予开发者更多资源^[6]。结合开发工具 Eclipse 在 Window 系统中集成。Eclipse 采用“平台+插件”的体系结构，平台仅作为容器，所有业务功能可封装在插件中应用。通过不同插件扩展实现所需功能，满足开发者需求。采用 Mysql 创建结构式数据存储，通过 Hadoop 对海量大文件和复杂、庞大的数据集进行存储处理；Redis 为临时数据提供缓存，可快速查询、处理数据^[7]。

2.2 系统功能模块

2.2.1 概述 系统主要有 5 项功能，包括预警展示、全景地图、数据仓库、风险管理和预警调度。各模块之间关联互通，贯穿卫生健康综合执法业务，实现卫生监督数据在线云计算、智能预警和分级管理等。

2.2.2 预警展示 建立全市 6 000 多家医疗机构诊疗行为数据风险预警展示页面,采用大屏进行数据可视化监控,监控内容包括全区各监督对象数据总览、执法记录数据总览、在线监测数据图表等,并实时更新动态。

2.2.3 全景地图 建立完善的监管对象全景档案。监管对象是以管理相对人一户一档为中心,将各县市区管理对象集中整合进一户一档。最终形成以管理对象为中心,以地理信息系统 (Geographic Information System, GIS) 定位信息为节点的全景档案,实现管理对象定位显示及筛选功能。

2.2.4 数据仓库 集中展示医疗机构诊疗行为数据,包括医疗机构诊疗信息、人员表、处方分类表、接诊类型表、机构输液许可、限制级药品、医护注册系统数据等,便于对监管对象进行预警行为分析。

2.2.5 风险管理 对辖区内 5 000 家医疗机构的诊疗超范围、执业异常、处方异常、抗菌药品警示、输液非许可、基因检测资质异常、限制级临床技术资质异常和麻醉行为资质异常等进行风险管理,具体内容如下:诊疗超范围预警是通过自研诊疗范围模型和实际诊疗行为数据进行比对,对辖区内医疗机构以及人员的超范围诊疗行为进行预警;医师执业预警是通过算法对医疗机构医护基本信息、医师多点执业信息和处方信息建立医师注册异常、诊疗行为异常、医师空挂等预警信息;处方异常预警是通过逻辑推理智能算法对注册医师的处方信息和处方内药品数据进行判断,获得处方异常记录,实时提示机构或人员违规行为;抗菌药超常使用预警,通过机器学习、卷积神经网络算法对医疗机构的诊疗数据和抗菌药物库进行判断,标记抗菌药物异常使用的处方记录并产生机构预警提示信息;输液许可预警,建立医疗机构输液许可数据、匹配规则及预警提示;母婴保健预警,分析机构资质,当不具有母婴保健服务许可以及计划生育专业的机构开展人流诊疗行为时,对该机构预警;基因检测预警,当不具有临床细胞分殖遗传学专业的机构进行基因检测行为时,对该机构预警;限制级临床技术预警,分析机构资质,当不具备限制级临床技术备案的机构开展限制级临床技术时,对该机构预警;麻醉行为预警,分析机构资质,当不具备麻醉专业资质的机构开展麻醉行为

时,对该机构预警。

2.2.6 预警调度 系统根据所产生预警信息分配至相应人员,对预警信息进行人工校验。相关人员核验信息属实后则组织相关执法人员进行现场取证并做相关处罚。行政处罚结束后可以在系统中录入处置结果。

2.3 数据层架构

2.3.1 数据采集 医疗大数据采集和分析是系统建设的基础。数据源头较多,数据采集任务较重,根据不同类型数据采集需要,主要采用数据库数据交换、录入、导入、接口等采集方式。第 1 种方式为数据库交换技术,主要采集来自全民健康信息平台的数据,如医院科室分类、个人电子病历信息等。第 2 种方式是搜索引擎技术。互联网是开放、巨大的资源库,互联网资源具有信息时效性和内容广泛性等特点,可通过搜索技术获得大量有效信息。根据医疗监管范围制定检索边界和规则,筛选出可用信息作为数据储备。第 3 种方式为批量导入,主要采集数据量较大的结构化数据。第 4 种方式为应用程序接口,将其他系统数据采集到本系统。使用数据仓库对数据进行清洗、导入、检索、结构呈现等,并快速筛选业务逻辑中所需数据,实现产生精准数据的目的。

2.3.2 数据存储 采集的数据根据结构化程度分为非结构化、半结构化和结构化数据。对于结构化数据,结合病历、科室、病种、药品名称等出现频率较高的高效值数据,优先处理并采用 Mysql 等传统存储技术进行存储;对于非结构化、半结构化和部分结构化的数据则采用 Hadoop 等大数据处理技术进行分布式、云存储处理并进行深入分析、挖掘、应用,为监管部门提供预警信息。

2.3.3 数据算法 医疗大数据具有海量性、复杂性、异构性、精确性等特点,传统数据处理技术无法满足其应用需求,需借助数据算法进行处理才可实现数据驱动,最大化发挥数据价值。本研究采用 Apache Mahout 技术,其主要提供推荐引擎、聚类和分类等可扩展机器学习领域的经典算法,可以实现文档聚类分类、推荐过滤、频繁子项挖掘等功能。Mahout 遵从 Apache 开源协议,充分利用 MapReduce 和分布式文件系统 (Hadoop Distributed File

System, HDFS) 的扩展性和容错性。根据医疗大数据构建分类模型, 对分类模型进行管理, 较为常用的 Mahout 贝叶斯分类算法用于对数据的预处理。其中聚类算法 K - means 用于提取特征并分词, 对词编码, 并将特征向量化, 向量化表现形式采用更省空间的 hashmap 保存, 将数据保存为 SequenceFile 格式。

2.3.4 硬件架构 服务器采用多节点分布式系统架构, 在业务支撑能力上, 有着良好的自主性和灵活性。对于后续的运维和故障恢复较为便捷。为应对更大的数据量配备至少 4 核心且时钟频率在 2.5GHZ 及以上的中央处理器, 8GB 内存。在系统软件上需支持 Hadoop 框架, 其核心功能是面向大数据处理的并行计算框架^[7], 核心组件有 HDFS、Hadoop MapReduce、YARN, 在 MapReduce 的思想下, Handoop 具备高可靠性、高扩展性、高效性、高容错性, 以应对海量数据存储和分析计算需求。

2.4 服务层架构

2.4.1 面向用户 一是根据关键词匹配医疗范围。在医疗范围中, 结合实际选择特有医疗术语输入智能监管预警系统中, 将用户超诊疗范围预警机构以低、中、高 3 个等级类型显示。二是可设置系统白名单。线下核实后如发现预警中信息不符合许可一致白名单, 在有效时间内不再匹配与预警该机构。系统提供统计、预警、处方还原、处置记录的永久查看功能, 并可跨平台、多用户、实时显示。

2.4.2 面向大数据平台 因医疗数据敏感性较高, 平台进行严格的权限管理, 除指定场所外不提供对接相关接口权限, 为第 3 方提供接口服务。服务内容仅包括相关统计数据、预警数据、加密后的处方还原等。终端设备通过市医保等系统的安全网络访问人工智能核心服务器获取对应数据, 经系统后台多重计算并采用 Vue + HTML5 技术将数据计算结果转换为各种统计图表直观展示。客户端只需配备通用浏览器软件, 基于 Web 服务的图形 CAD 网络化协同设计框架, 用户使用任意设备和操作系统都能获得良好体验。

2.4.3 面向数据智能处理 在实现信息系统数据收集的前提下, 针对大数据仓库中的数据进行多维度运算分析。主要包括按照医疗技术准入运算、按

照处方诊断操作运算、按照关键医用耗材运算、快速精准定位预警机构信息并提供可追溯查看的依据记录。配备集群计算服务器, 提高运算效率和处理速度。配备支持统一计算设备架构 (Compute Unified Device Architecture, CUDA), 具有图形处理单元的图形处理器 (Graphic Process Unit, GPU), 从而将特定计算分解为独立、可同时执行的较小计算, 并将计算重新组合或同步以形成计算结果。

3 实践结果

3.1 充分利用医疗大数据, 提高数据效能

系统基于杭州市全民健康信息平台上归集的全市医疗机构医保诊疗行为数据, 接入情况如下: 总医疗机构数 7 115 家, 累计 333 255 052 条, 药品信息 208 887 条。截至 2021 年 9 月查处包括限制临床应用医疗技术、违法开展放取环技术、门诊违规使用特殊级抗菌药物、违反基层医疗机构停止接诊规定、违规接诊发热等问题事件共计 123 起, 监管数据包括定点医疗服务机构、人员就诊登记、项目明细、手术、住院医嘱、病案主页、检查单和检验单等信息, 建立面向全市医疗机构电子病历数据的信息共享交换平台, 运用大数据智能分析技术, 实现大数据高效挖掘、分析利用, 为精准执法提供确凿证据。

3.2 有效利用智能分析技术, 提升监管效率

系统覆盖全市 90% 以上的医疗机构, 通过基础支持库与多维度诊疗行为监管算法, 智能筛选出诊疗行为违法预警信息。以使用抗菌药物开展静脉输液未经核准预警为例, 目前杭州市医疗机构诊疗行为数据有 300 032 423 423 条, 发现门诊使用特殊级抗菌药物异常报警 2 286 条, 涉及医疗机构 49 家。按照相关条例和文件要求, 从抗菌药物、医疗机构类别、静脉输液、县级卫生行政部门核准 4 个维度进行分析; 将提取的各维度信息进行完善, 形成相对应的规则, 同时将规则与预警进行关联; 系统按照规则对医保数据分析计算得出预警; 监督员依照系统提供的预警线索前往现场核查。通过上述流程数分钟即可实现对上百万条电子病历和处方进行监管、快速非现场执法, 使卫生执法人员有更多时间投入到监督执法活动中。

3.3 实现实时监控数据, 提高执法精准性

应用该系统既可实现卫生监督执法工作联动, 又可及时归集数据, 做到实时掌握风险、发现问题, 使卫生监督工作真正摆脱以往信息不畅、无法从海量诊疗数据中查找潜在风险信息和违法事实信息的局面, 实现由传统的人工查找工作模式向线上执法、“互联网+监管”的模式转变。

4 结语

本文提出基于全市诊疗数据、通过多维度运算规则得到精准预警数据的处理方式, 详细介绍精准预警监管系统整体架构和关键技术。制定一整套运算规则, 进一步加强对医疗技术临床应用的监管。通过“互联网+监管”模式, 将原本分散的监管方式向统一化监管方向转变, 监管范围提升到全市 95% 以上的医疗机构, 智能化的精准预警提升了监管效率。本系统尚存在一定缺陷, 如设计架构适合就诊记录、检验、处方等具有固定格式的结构化数据, 对于遗嘱等半结构化或非结构化数据无法对其内容进行精准判断, 存在监管漏洞。下一步将针对

此类问题采用自然语言处理技术进行预分析, 提取内容标签并进行对比分析, 提高系统监管准确性和全面性。

参考文献

- 1 国务院. 关于加强和规范事中事后监管的指导意见 (国发〔2019〕18 号) [EB/OL]. [2021-07-20]. http://www.gov.cn/zhengce/content/2019-09/12/content_5429462.htm.
- 2 尹佳, 张筱烽, 鲜明. 基于史密斯模型的四川省“医疗三监管”政策执行情况研究 [J]. 实用医院临床杂志, 2019, 16 (2): 241-243.
- 3 宋世贵. 四川: “三监管”瞄准全程动态精准 [J]. 中国卫生, 2019 (6): 45-46.
- 4 朱红. 展现新时代卫生计生行政执法新作为 [J]. 中国卫生监督杂志, 2018, 25 (5): 403-407.
- 5 胡川, 罗浩, 汪鹏. 医院智能血糖监测管理系统设计与应用 [J]. 医学信息学杂志, 2020, 41 (7): 68-71.
- 6 刘洋, 史森中. 基于智能手机操作平台的移动医学影像系统软件设计 [J]. 医学信息学杂志, 2019, 40 (2): 25-28.
- 7 孙荣明. Hadoop 框架核心技术在高校大数据教学系统中的应用分析 [J]. 电脑知识与技术 2018, 25 (14): 129-130.

(上接第 41 页)

病心绞痛的病机。现代医家在传承古人治疗冠心病用药特点及规律的同时, 在辨证和治法方面发生较大改变。现代医家在冠心病治疗方面更侧重于“清”法运用^[10], 治疗冠心病使用较多的是活血化瘀、清热解毒、行气活血等方法。现代医家发现, 随着生活环境的变化现代人的体质较古人改善较多, “热”“毒”已成为导致冠心病的更大诱因。疾病的发生、发展与转归受多方面因素影响, 这些辨证及治法的改变正是体现了中医因时、因地、因人制宜的辨证精髓。

参考文献

- 1 中国中医科学院中医药信息研究所科学数据研究室. 医案大数据分析系统 [EB/OL]. [2021-07-05]. <http://academy.cintcm.com/channel-about.jsp?channelId=10771>.
- 2 中国中医科学院中医药信息研究所科学数据研究室. 中

医古方数据库 [EB/OL]. [2021-07-05]. <http://zyyqx.cintcm.ac.cn:9698/zyyqx/manager/timeOut.html>.

- 3 朱文峰. 中医诊断学 [M]. 北京: 中国中医药出版社, 2017.
- 4 姚乃礼. 中医症状鉴别诊断学 (第 2 版) [M]. 北京: 人民卫生出版社, 2000.
- 5 高学敏. 中药学 [M]. 北京: 中国中医药出版社, 2007.
- 6 邓中甲. 方剂学 [M]. 北京: 中国中医药出版社, 2017.
- 7 李园白, 杨阳. 中医药科学数据 [M]. 北京: 科学出版社, 2016.
- 8 中国中医科学院中医药信息研究所. 中国中药药对数据库 [EB/OL]. [2021-07-05]. <http://cintmed.cintcm.com/cintmed/searchIndex/basic.html?dbtype=zyyd&pdl=6>.
- 9 李柳骥. 冠心病心绞痛古今中医文献整理与研究 [D]. 北京: 北京中医药大学, 2007.
- 10 屈茜茜. 清热解毒法辨治冠心病的中医文献研究 [D]. 济南: 山东中医药大学, 2016.