

人工智能与大数据在传染病防控中的理论研究*

安明扬

赵增仁

谭友文

(北京大学公共卫生学院社会医学
与健康教育系 北京 100191)

(河北医科大学第一医院
石家庄 050030)

(江西省人民医院 南昌 330006)

潘习龙 王培玉

安 鹏

(北京大学公共卫生学院社会医学与健康教育系
北京 100191)

(北京京东叁佰陆拾度电子商务有限公司
北京 100176)

[摘要] 介绍人工智能与大数据在国内外的研究现状, 阐述人工智能与大数据在呼吸道传染病方面的理论研究情况, 提出构建传染性疾病预防平台, 对关键体征、关键症状、关键事件进行监听与预警, 取得一定成效, 未来将对该平台进行进一步研究与完善。

[关键词] 人工智能; 信息系统; 公共卫生; 传染性疾病预防

[中图分类号] R-058 **[文献标识码]** A **[DOI]** 10.3969/j.issn.1673-6036.2022.10.008

Theoretical Research on Infectious Disease Prevention and Control Based on Artificial Intelligence and Big Data AN Mingyang, Department of Social Medicine and Health Education, School of Public Health, Peking University, Beijing 100191, China; ZHAO Zengren, The First Hospital of Hebei Medical University, Shijiazhuang 050030, China; TAN Youwen, Jiangxi Provincial People's Hospital, Nanchang 330006, China; PAN Xilong, WANG Peiyu, Department of Social Medicine and Health Education, School of Public Health, Peking University, Beijing 100191, China; AN Peng, Beijing Jingdong San Bai Lu Shi E-commerce Co. Ltd., Beijing 100176, China

[Abstract] The paper introduces the research status of Artificial Intelligence (AI) and big data at home and abroad, expounds the theoretical research status of AI and big data in respiratory infectious diseases, and proposes that the established quality monitoring platform for infectious diseases can monitor and warn key signs, symptoms and events, and has achieved certain results. Further research and improvement of this platform will be carried out in the future.

[Keywords] Artificial Intelligence (AI); information system; public health; infectious diseases

[修回日期] 2022-01-27

[作者简介] 安明扬, 硕士研究生, 发表论文 1 篇; 通信作者: 潘习龙, 教授。

[基金项目] 陕西省 2021 年自然科学基金基础研究计划项目“基于 5G 与 AI 的医疗质量监控平台研究”(项目编号: 2022JM-592)。

1 引言

呼吸道传染性疾病的产生、传播、影响与防范等环节都具有很强的社会学特性。在传染性疾病管理过程中,常常在病例出现之后相关部门才通过传染病申报系统逐级上报,很难做到防患于未然。在人工智能^[1-2]、大数据^[3]和云计算^[4]等互联网技术飞速发展的背景下,管理部门可以改变高投入、低效率、迟响应的传染性疾病管理现状。为了消除或缓解呼吸道传染性疾 病给人民健康、生命财产、国家经济带来的巨大损失,可以运用大数据与人工智能技术,根据气候、环境、病原学、人群等要素提前做出预警;当病例出现后,根据患者体征、传播速度、传播途径、影响范围等要素,即刻启动相应防控级别,将损失控制在最小范围。因此本研究具有重要现实意义。

2 国内外人工智能与大数据研究现状

人工智能(Artificial Intelligence, AI)技术是指深入研究、发展并用于模拟、扩充和拓展人的智慧的理论、方式、技能和运用体系的一种新兴技术。随着“人工智能+X”上升为国家战略,我国开始全力推进人工智能在各个领域的创新融合。人工智能技术在健康保健、精准医疗、患者治疗、新药开发等领域发挥着显著作用。1999 年 1 月研制的“达芬奇”智能机器人手术系统,是世界上第 1 套真正在手术室中应用的人工智能机器人手术控制系统。在微创、无创手术治疗领域获得巨大成功,在某种程度上代表着世界范围内智能医疗手术机器人的最高水平。国内研究者将人工智能技术运用到传染性疾 病监控领域,取得了一些理论和实践方面的成果^[5-6]。

云计算作为计算机领域的一种新兴科技,受到商业界和学术界的关注。2008 年谷歌集团通过云计算平台发布谷歌健康(Google Health)服务,其目的是将海量医疗数据存储于云端,通过云服务供患者查看。微软公司也研发过类似的健康管理服务

(HealthVault)。在医疗大数据应用领域,许多研发机构和人员已经开始利用 Hadoop 大数据框架开展医疗服务与医疗临床项目管理研发工作^[7-8]。Taylor R C^[9]详尽阐述 Hadoop 在生物信息学中的广泛应用,Schatz M C^[10]研发一款名为 CloudBurst 的开源软件包,为生物医学信息分析提出一种采用 Hadoop MapReduce 技术的并行算法。

国内学者对医疗大数据的学术研究与应用探讨也非常广泛。中国工程院院士邬贺铨^[11]在 2012 年中国科技论坛上提出,智慧医疗应该是大数据应用研究的前沿领域,重点应用领域为临床决策支持、药品分析、误诊分析等多个方面。姚志洪^[12]提出医疗信息化发展的 10 大重点,包括电子健康档案标准化、大数据与 NoSQL、健康物联网、健康云和个性化医疗等方面。周光华、辛英和张雅洁等^[13]在综合论述我国医疗数据资源现状的基础上,详细分析大数据在医药研发、公共卫生、疾病管理和健康管理 4 个方面的应用。陈鹤群^[14]提出医疗大数据面临的两大挑战:一方面患者隐私数据保护应细化到标识隐私匿名保护、位置隐私保护、关系匿名保护等方面;另一方面应促进医疗数据标准化,做好医疗大数据共享和开放。张振、周毅和杜守洪等^[15]论述医疗大数据带给医药研究、智慧医疗和商业智能的发展机遇,提出医疗大数据实施过程中面临信息孤岛、医疗数据标准化和医务人员信息意识缺乏等难题,同时在技术应用上存在数据集成、数据存储、数据检索和数据分析利用等方面的挑战。

3 人工智能与大数据在呼吸道传染病方面的应用研究

3.1 概述

基于大数据和人工智能算法,预计部署 22 个模块及 3 万个感染类疾病监控预警节点,实现全过程实时闭环监控,结合大数据与人工智能技术对传染性疾病的计算,构建事前异常状态预警体系。融合信息技术、临床医学与重大公共卫生事件相关知识,实现云部署、跨平台、多医院、多区域的数据同步,达到系统云端一体化、线上线下一体化,建

立标准化、智能化、云服务化的呼吸道传染性疾病防范与管理 体系。

3.2 传染性疾病监控点研究方法

3.2.1 建立质控云平台 根据呼吸道传染病 SAF-ECARE 体系设置, 结医院个性化管理需求, 以敏感事件、敏感症状及敏感体征为切入点定义相关数据源, 触发数据采集监听模块, 形成符合传染性疾病特点的质控云平台。实时监测所有影响传染性疾病事件的重要节点与重要过程, 记录所有事件从启动到终止的全过程, 保存可跟踪的历史数据, 做到有据可查。AI 在控制各个环节时并非按照一个质量控制点进行管控而是按照功能模块, 各个模块均可对应多个质量监控点。

3.2.2 建立事前提醒和预警、事中监护、事后评价管理体系 除已实现的质控预警平台以及部分监控点之外, 还计划开发危急值立体闭环监控、发热门诊管控、传染病管理、重大公共卫生事件地区化

管控等模型。这套体系将有助于地区疾病预防控制中心、综合性医疗机构、门诊急诊医生和临床医务人员全面掌握对呼吸道等感染性疾病的筛查, 在管理工作中把握主动权, 将传统事后申报转为建立事前提醒和预警、事中监护、事后评价的管理体系。

3.3 AI 质控云平台系统架构优化升级方法

3.3.1 建立标准化、智能化、云服务的传染性疾病预警、防控与管理 系统 从业务层面梳理, 按行为活动划分, 融合了丰富完备的临床知识库、质控系统、管理监控、辅助策略等。医院管理系统将整体拆分成对接层、大数据管理层、决策层、多功能中台业务层、应用层 5 个层面, 实现医疗系统平台即服务 (Platform as a Service, PaaS), 完成服务平台化设计, 支持业务云端部署, 支持多院区、多平台、多机构大数据同步、即时管控。实现系统云端一体化、线上线下一体化、临床管理一体化, 见图 1。

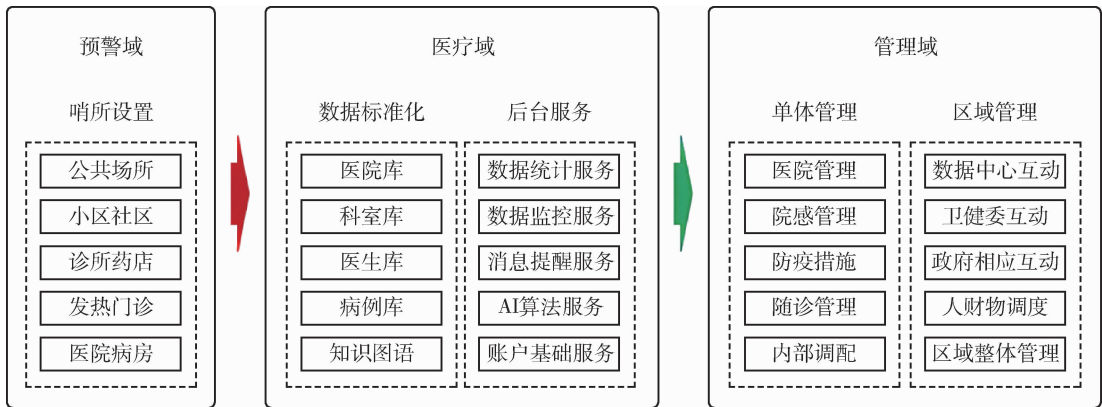


图 1 预警、医疗、管理运作框架

3.3.2 数据对接及数据标准化 为解决多源异构造成的诊疗数据分散、难以共享等问题, 采用创建主数据管理 (Master Data Management, MDM) 中心、建立临床数据中心 (Clinical Data Repository, CDR) 的方法, 进行数据治理、清洗和集成。采用大数据分析技术和人工辅助的形式, 形成标准化、结构化的医院系统底层数据结构, 通过 CDR、运营数据中心 (Operational Data Repository, ODR), 支持中台服务层的深度学习策略模型调用。中台服务通过先拆分、抽离等方式拆分为单独的支持点, 然

后对不同业务场景梳理运算、聚合成为中台服务点, 以支持 WAP、H5、APP 等不同应用层展示, 支持不同场景独立运行。

3.3.3 遵循系统独立性、数据低冗余和资源剥离的原则建立医院监控系统 运用 Hadoop 体系架构进行分布式数据存储、运算, 底层通过 Hdfs 搭建分布式存储; 运用 MapReduce 映射和化简编程模型提高数据处理效率, 解决单机性能瓶颈限制。通过 Hive 建立数据仓库。策略层引入深度学习中 TensorFlow 框架自然语言处理算法, 以及部分受限波尔兹

曼机 (Restricted Boltzmann Machine, RBN) 算法, 提供策略模型算法支持, 通过 Hue 算法工具支持数据可视化展示。使用“大数据 + AI + 知识图谱 + 语言识别”等技术手段, 使机器可以读懂大量病例从

而实现信息标准化, 同时通过数据逻辑判断, 将更多信息自动转换为可分析、可计算、可利用的数据, 见图 2。

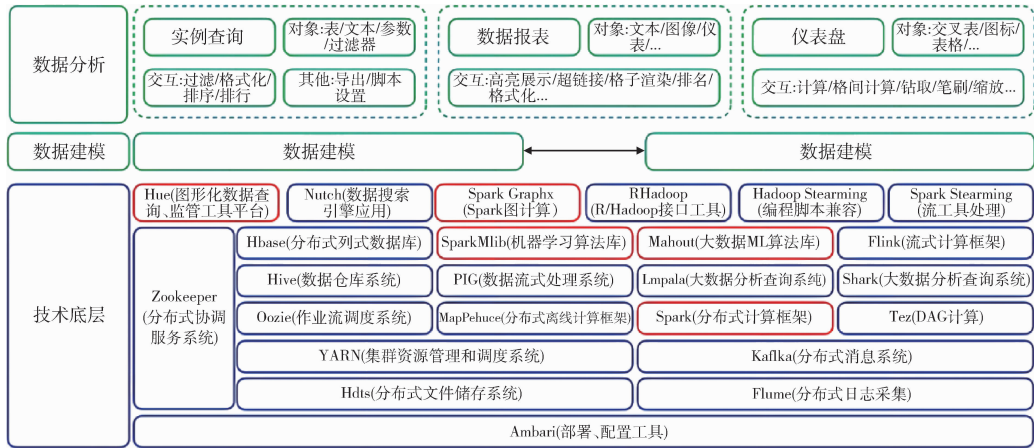


图 2 医院监控微服务架构

3.3.4 实现消息系统与策略系统无缝对接以及管理结果快速提醒 服务于患者, 构建简洁、易用、方便的医疗监测体系; 服务于医务人员, 设置便于应用、高效管理、少差错的信息平台功能; 服务于管理人员, 提供信息监控、工作流程管控、决策过

程综合分析的管理工具。从部分应用到全部应用, 从信息分散到大数据融合, 从事务处理到智能控制, 提供完整、安全、可靠、实时的患者信息访问和档案管理系统, 进一步优化内部工作流程, 有效减少医疗差错, 见图 3。

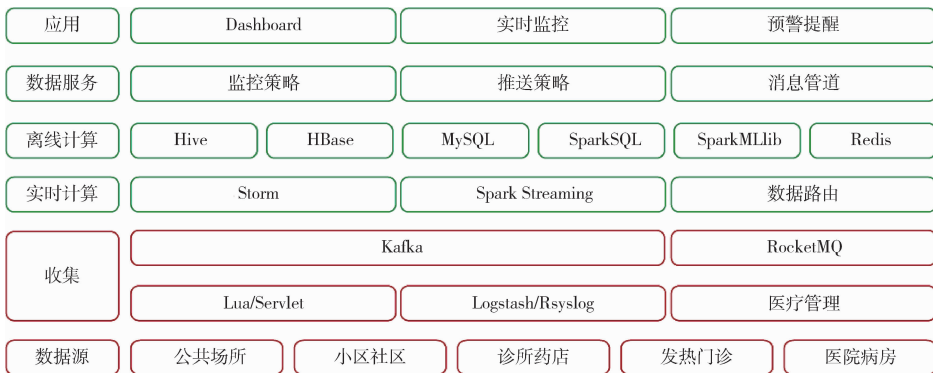


图 3 实时消息系统架构

4 结语

本研究基于大数据和 AI 算法, 预计部署 22 个模块及 3 万个感染类疾病监控预警节点, 实现全过程实时闭环监控, 结合大数据与人工智能技术对传染病进行计算, 构建事前异常状态预警体系。自 2015

年起课题组联合国内多所大中型公立医院的临床专家, 按照我国医院质量监督管理有关文件精神初步建立了医疗质量监控平台。该平台融合信息技术、临床医学与重大公共卫生事件相关知识, 实现云部署、跨平台、多医院、多区域数据同步, 达到系统云端一体化、线上线下一体化, 实现标准化、智能化、云服务的呼吸道传染性疾病预防与管理。该平台有

效打通医院各信息系统, 实现各功能系统的数据整合与清洗并对收集到的数据进行监听和预警。目前已在中南大学湘雅二院、广东省第二人民医院、山东枣庄市立医院、山西省长治人民医院、郑州市第六人民医院 5 家医院建立疫情应急医院合作研究基地。该系统监控点更多侧重于医疗质量, 在呼吸性传染病的数据监控精准性、信息共享延时性等方面有待完善。为此课题组将联合传染病学专家与人工智能专家, 从理论与实践方面进一步完善, 提出传染性疾病监控点标准体系, 部署传染性疾病监控关键节点, 优化架构, 研究人工智能深度学习功能, 搭建更广泛、智能化的传染性疾病监控系统。

参考文献

- 1 唐珂. 人工智能未来发展前景 [J]. 人民论坛, 2018 (2): 24-25.
- 2 汤建. 基于机器学习的临床决策支持系统的设计 [J]. 计算机光盘软件与应用, 2014 (14): 36-45.
- 3 Lynch C. Big Data: How Do Your Data Grow? [J]. Nature, 2008, 455 (7209): 28-29.
- 4 Dudin E B, Smetanin Y G. A Review of Cloud Computing [J]. Scientific and Technical Information Processing, 2012, 8 (4): 280-284.
- 5 王晨阳, 潘习龙, 吴曼琪, 等. 人工智能在医学领域应

- 用浅析 [J]. 中华医院管理杂志, 2020 (1): 50-52.
- 6 吴曼琪, 潘习龙, 程迪, 等. 基于临床大数据的传染性
疾病监控辅助系统探讨 [J]. 中华医院管理杂志, 2020
(1): 53-55.
 - 7 Liu B, Madduri R K, Sotomayor B, et al. Cloud-based
Bioinformatics Workflow Platform for Large-scale Next-
generation Sequencing Analyses [J]. Journal of Biomedical
Informatics, 2014, 49 (6): 119-133.
 - 8 Santana Q L. HTVE-Hexagon: High-performance, Paral-
lelized Sequence Alignment for Next-generation Sequencing
Data Analysis [J]. Plos One, 2014, 9 (6): e99033.
 - 9 Taylor R C. An Overview of the Hadoop/MapReduce/HBase
Framework and Its Current Applications in Bioinformatics
[J]. BMC Bioinformatics, 2010, 12 (6): 3395-3407.
 - 10 Schatz M C. CloudBurst: Highly Sensitive Read Mapping with
MapReduce [J]. Bioinformatics, 2009, 25 (11): 1363-1369.
 - 11 郭贺铨. 大数据时代的机遇与挑战 [J]. 中国储运,
2013 (3): 53-54.
 - 12 姚志洪. 医疗卫生信息化 10 大视点 [J]. 医学信息学
杂志, 2013, 34 (1): 2-9.
 - 13 周光华, 辛英, 张雅洁, 等. 医疗卫生领域大数据应用探
讨 [J]. 中国卫生信息管理杂志, 2013 (4): 296-300.
 - 14 陈鹤群. 大数据环境下医疗数据隐私保护面临的挑战及相
关技术梳理 [J]. 电子技术与软件工程, 2014 (16): 51-53.
 - 15 张振, 周毅, 杜守洪, 等. 医疗大数据及其面临的机遇
与挑战 [J]. 医学信息学杂志, 2014, 35 (6): 2-8.

2023 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内创刊最早的医学信息学方面的专业指导性刊物。主管：国家卫生健康委员会；主办：中国医学科学院。中国科技核心期刊（中国科技论文统计源期刊），RCCSE 中国核心学术期刊，美国《化学文摘》《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息研究，医学信息技术，医学信息资源管理与利用，医学信息教育等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业师生及广大医教研人员。

2023 年《医学信息学杂志》国内外公开发行，每册定价：15 元（月刊），全年 180 元。邮发代号：2-664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号（100020）医科院信息所《医学信息学杂志》编辑部；电话：010-52328672，52328686，52328687。

《医学信息学杂志》编辑部