

# ·医学信息· 基于深度学习技术中医优势病种证候到治法方剂算法模型的构想

曹继忠, 赵亮, 温川飙\*

(成都中医药大学, 四川 成都 610075)

**摘要:** 目前开展的中医辩证论治信息化研究都只针对“辨证”过程进行探索, 而同样重要“论治”过程(即选法定方)的信息化却被相关研究人员所忽略。目前, “论治”过程的研究只是运用浅层算法进行简单关联, 推导出结果的准确性难以得到验证。本项目以神经网络作为改进算法的基本架构, 采用一种学习深层非线性网络结构, 实现复杂函数逼近的深度学习技术与中医的“因位性势”相结合, 利用挖掘出的信息建立的数据仓库, 训练模型, 确定新型算法中更为高效的网络权重。在训练模型的过程中采用具有较强的非线性映射能力、高度自学习和自适应的能力的共轭梯度 BP 神经网络以反向传播学习算法的方式, 从输出层经各中间层逐层修正各连接权值, 得出最为优化的网络权重, 推荐出准确、合理的治法与方剂。本算法模型探索了中医名老专家选法定方的思维模式, 并核查中医治疗过程是否规范合理, 检出误治、失治病例, 从而提高临床疗效, 使中医治疗过程实现规范化与标准化。

**关键词:** 因位性势; 深度学习; 算法模型; 选法定方

**中图分类号:** R857.11

**文献标识码:** B

**DOI:** 10.19613/j.cnki.1671-3141.2018.28.133

**本文引用格式:** 曹继忠, 赵亮, 温川飙. 基于深度学习技术中医优势病种证候到治法方剂算法模型的构想 [J]. 世界最新医学信息文摘, 2018, 18(28): 252-253, 258.

## The Conception of the Model of the Algorithm of TCM Prescription Based on Deep Learning Technology

CAO Ji-zhong, ZHAO Liang, WEN Chuan-biao\*

(Chengdu University of Traditional Chinese Medicine, Chengdu Sichuan 610075)

**ABSTRACT:** At present, the research on TCM dialectical and treatment informationization is only aimed at “dialectical” process to explore, and the same important “treatment” process (that is, the election of the therapies and prescription) information has been ignored by the relevant researchers. At present, the study of the process of “treatment” is only a simple association using shallow algorithm, and it is difficult to deduce the accuracy of the results. In this paper, the neural network is used as the basic architecture of the new algorithm. A deep learning network is used to realize the deep learning technique of complex function approximation and the “Etiology, Lesion site, Disease nature and Disease trends” of traditional Chinese medicine, and the data created by excavating the information Warehouse, training model, to determine the new algorithm in a more efficient network weight. In the process of training the model, a conjugate gradient BP neural network with strong non-linear mapping ability, high self-learning and adaptive ability is used to reverse-propagate the learning algorithm, and The weight of each connection is corrected from the output layer to the middle layer. The most optimal network weight recommend accurate and reasonable treatment and prescription. The algorithm model explores the thinking mode of the prescriptions of traditional Chinese medicine prescriptions, and verifies whether the treatment process of Chinese medicine practitioners is reasonable and reasonable, and detects the misdiagnosis and treatment cases, so as to improve the clinical curative effect and standardize the traditional Chinese medicine treatment process.

**KEY WORDS:** Etiology, Lesion site; Disease nature and Disease trends; Deep learning, algorithm model; Choose the therapy and determine the prescription

## 0 引言

近年随着国家推进中医信息化的出台, 中医工作人员与信息化工作者对此进行了大量的研究与探索。“整体观”与“辩证论治”是中医的两大特点, 其中“整体观”是中医诊病的一种基本思想; “辩证论治”则是目前中医界普遍认可的一种“规范化流程”, 是中医诊病的关键。“辩证论治”, 根据其思维理论模式分为“辨证”与“论治”两个主要阶段<sup>[1]</sup>。而目前的大多数研究都只针对“辨证”过程进行信息化探索, 并取得一定成果<sup>[2-3]</sup>。但关于“论治”部分的信息化研究尚较少开展。“论治”即是根据“证型”以确定“治法”、“方剂”的过程。我国中药资源种类有 12807 种<sup>[4]</sup>, 再将它们以不定的数目相互配伍组成方剂, 所以“证型”与“方剂”的对应是非常复杂的, 其整理数据量较庞大。因此, 我们要设计出适合处理“证”到“方”具有复杂线性关系与数据量大等特点的新型算法, 来实现“证”到“方”的准确对应。以探索中医名老专家选法定方的思维模式, 并检验在职中医生是否按中医“论治”的规范流程诊病。核查中医生治疗过程是否规范合理, 检出误治、

失治病例, 从而提高临床疗效, 使中医治疗过程实现规范化与标准化。

## 1 目前对“论治”部分的信息化研究中还存在以下问题

(1) 结果的准确性。要么是基于中医专家知识与经验的中医专家系统<sup>[5-6]</sup>, 这种“论治”过程过于局限某个中医专家的临床经验, 中医知识广度不够; 要么是某些用智能算法模拟中医临床经验进行“论治”过程, 这种“论治”过程只是运用浅层算法进行简单关联, 推导出结果的准确性难以得到验证。

(2) 样本采集的质量和容量。建立系统的稳定性与样本质量好坏和容量大小关系密切。对于小样本数据, 模型不能得到充分学习, 会出现过拟合的情况, 因此而产生不稳定的系统。当前, 中医数据集由于受到采集成本等现实条件的制约, 相对于数据维数而言, 是一个小样本数据集。

本课题拟以四川省中医药管理局的“中医数字化诊疗平台”门诊电子病历的数据为基础, 目前该平台的电子病历数据来源于德阳市、绵阳市、成都市等 70 个市以及其下辖县的中医医院, 上传数据量已有 100 多万条。然后采用大数据的海量数据挖掘技术, 选择合适的数据挖掘方法(如关联规则、分类和聚类等), 针对中医的优势病种如: 胃炎, 以电子病历的海量真实数据进行数据挖掘, 从而挖掘出中医证候、治法

**作者简介:** 第一作者: 曹继忠 (1991-), 男, 硕士, 研究方向: 中医药数字化研究。

**通讯作者:** 温川飙 (1970-), 男, 研究员; 成都中医药大学数字医药研究所所长, 研究方向: 中医药数字化研究。

以及处方等信息,建立数据仓库。以上解决了样本采集的质量和容量的问题。

再以神经网络作为新算法的基本架构,采用一种学习深层(深层的好处是可以较少的参数表示复杂的函数)非线性网络结构,实现复杂函数逼近的深度学习技术与中医的“因位性势”相结合,利用挖掘出的信息建立的数据仓库,训练模型,确定新型算法中更为高效的网络权重,提高“证”到“方”的匹配准确度。

## 2 研究现状及发展动态分析

“辨证论治”,根据中医思维理论模式分为“辨证”与“论治”两个主要阶段。目前,BP 神经网络在中医证候诊断的应用中是最为广泛的一类<sup>[7-8]</sup>。BP 网络模型虽也被称作多层感知机,但实际是只含有一层隐层节点的浅层模型<sup>[9]</sup>。且 BP 网络只能有标签数据来训练,但大多数数据是没有标签的。浅层结构算法,其局限性在于有限样本和计算单元情况下对复杂函数的表示能力有限,针对复杂分类问题其泛化能力受到一定制约<sup>[10]</sup>。而当前的大多数研究都只针对“辨证”过程进行信息化探索,并取得一定成果<sup>[11-13]</sup>。但关于“论治”部分的信息化研究尚较少开展,而选法定方是中医拥有良好疗效的关键。

中医是一个复杂系统,证候与证候间,证候与治法间,证候与疾病间往往具有模糊性、非线性、复杂性的特点。深度学习是机器学习的一个新领域,是关于自动学习要建模的数据的潜在(隐含)分布的多层(复杂)表达的算法。深度学习通过学习一种深层(深层优势为可以用较少的参数表示复杂的函数)非线性网络结构,以实现复杂函数逼近,从少数样本集中学习数据集的本质特征。与浅层学习相比,深度学习强调了模型结构的深度,明确突出了特征学习的重要性,通过逐层特征变换,逐层学习并把学习到的知识传递给下一层,使下一层能够得到更高级别的特征,将样本在原空间的特征表示变换到一个新特征空间,从而使分类或预测更加快速准确<sup>[14]</sup>。深度学习与中医辨证论治相互融合是当前实现中医辨证论治信息化的研究走向与热点。

## 3 新型算法模型的构建

### 3.1 基于大数据技术建立数据仓库,实现对中医“论治”客观数据的支持向量特征提取

以四川省中医药管理局的“中医数字化诊疗平台”的门诊电子病历数据为依据,进行整理归纳。将整理好的数据进行预处理,再对数据挖掘方法的分析比较,并选择合适的数据挖掘方法进行数据挖掘。根据中医辨证论治的特点,拟用关联规则和分类的数据挖掘方法。

从而挖掘出“证型”、“治法”以及“方剂”等数据信息,然后建立用于深度学习数据仓库。

**3.2 模拟中医“论治”特点,将神经网络与多维空间数学辨证论治模型中“证”的因、位、性、势量化码结合,产生一种新的基于深度学习的中医智能辨证论治算法,并构建算法模型。**

以“证型”作为输入层,“因位性势”、“治法”依次作为隐含层,最后的得出的“方剂”作为输出层,构建神经网络模型。根据中医“因、位、性、势”理论,把证型分解为“因、位、性、势”,通过编码量化,实现“病位归经”。再使用建立好的深度学习数据仓库中的数据,对神经网络模型进行训练,不断调整编码值,得出最佳编码值。

已有中医辨证论治平台的量化码在准确性上还需提升优化,同时,其算法为浅层结构算法,浅层结构算法的局限性在于有限样本和计算单元情况下对复杂函数的表示能力有限,针对复杂分类问题其泛化能力受到一定制约。与浅层

学习相比,深度学习强调了模型结构的深度,明确突出了特征学习的重要性,通过逐层特征变换,逐层学习并把学习到的知识传递给下一层,使下一层能够得到更高级别的特征,将样本在原空间的特征表示变换到一个新特征空间,从而使分类或预测更加快速准确。

### 3.3 开展深度学习的算法训练,优化权重估值网络

将设计好的神经网络深度学习模型放入中医优势病“证”、“法”、“方”的大量临床真实数据中进行训练。应用非监督逐层训练算法,解决深层结构的优化难题。每次无监督学习只训练一层,再将此层的输出作为下一层的输入。对中医常见病数据库多个层次变量之间的内在关系,进行机器学习,得出神经网络的权重估值。初步深度学习和训练算法如下:首先给神经网络提供一定量数据的训练样本组,每组样本由输入样本数据与期望输出数据组成。通过修改网络权值使网络的期望输出和实际输出趋向一致,当网络最终的实际输出与理想输出结果基本一致时,训练结束。

选用 Sigmoid 函数,作为激活函数,本函数连续可微分并且更接近生物神经元的信号输出形式,所以本研究中激活函数亦选用此函数。

## 4 算法模型构建的可行性

### 4.1 具有涵盖中医重点学科、中医药信息学、大数据研究、软件工程等多学科交叉的优秀专业团队

课题研究团队由 10 人组成,其中高级职称 2 人,课题组成员全部为硕士或博士学历,课题组成员主持或承担过国家自然科学基金、科技部 973 计划项目、科技支撑计划及多个省级科技项目,具有较强的科研能力和较高的研究水平。项目的设计、分析、执行、测试为多学科的相互配合,这种优势互补,联合攻关的形式将有助于项目的研究和实施。

### 4.2 项目已具备成功实施深度学习算法的三个关键条件

足够多的数据、足够大的计算能力和专业化的初始权重。本课题有前期项目的支持,有 100 万条真实中医电子病历数据对算法模型进行训练,不断调整“因位性势”编码值。更有专业的中医团队在分层特征提取初始化阶段提供专业化的初始权重处理能力;课题负责人所在的数据中心云计算平台是目前中医药大数据在西南片区最完备的大数据研究和应用机构,由于硬件的高速发展,计算机已经能提供近乎无限的计算能力,对权重计算和深度学习的特征提取提供可靠的保障。

## 5 结语

本文提出采用深度学习技术与中医“因位性势”的方法,逐层学习主要特征,再以“因位性势”理论作为隐含层在“证型”、“治法”之间作紧密关联,贴切的模拟中医“论治”过程,探索中医“选法定方”的理论思维,规范中医“论治”过程。并为辨证论治的信息化研究提供新的研究方式。为中医辨证论治体系与神经网络深度学习模型的结合运用研究新思路与新方法。

## 参考文献

- [1] 任廷革,李宇航,陈永义,等.中医辨证论治信息研究评述[J].中国中医药信息杂志,2006,(03):94-96.
- [2] 倪荣,彭明德,沈玉强,等.模式识别与中医智能化辨证[J].中华中医药学刊,2015,(05):1061-1064.
- [3] 吴芸.中医辨证融合型软计算方法研究[D].导师:周昌乐,厦门大学,2007.
- [4] 张惠源,赵润怀,袁昌齐,等.我国的中药资源种类[J].中国中药杂志,1995,(07):387-390.
- [5] 王瑞祥,刘晓玉.一种基于粗集简约的中医专家系统设计和实现[J].时珍国医国药,2013,(10):2476-2477.

(下转第 258 页)

合理作息、科学运动、保持健康乐观心理状态、戒烟、控制体重、控制血压、血糖、血脂、采用合理规范的药物治疗,包括:阿司匹林、氯吡格雷、他汀类药物、 $\beta$ 受体阻滞剂、ARB/ACEIs 治疗等。

本调查研究发现,奉贤地区冠心病 PCI 术后患者以 60-75 岁人群居多,且男性多于女性。因心肌梗死行冠脉 PCI 术患者占 48.24%,其中急性心肌梗死患者占 28.17%,因此,发生心肌梗死患者占冠心病 PCI 患者的比例极高,其中急性心梗患者比例很高,危害大,提高冠心病心肌梗死的早期诊断治疗及预防极其重要。大约有 90% 左右的患者合并一种或多种慢性病,如高血压、糖尿病、高脂血症等其他疾病,其中高血压的患病率最高,故将高血压病、糖尿病等慢性疾病的进行有效的预防控制极其重要。本研究中,服用阿司匹林的患者占比为 98.33%,服用氯比格雷的患者占比为 96%,服用他汀类药物的患者占比为 96.34%,这几类药物使用的依从性较高,但服用  $\beta$ 受体阻滞剂类的比例则仅为 48.67%。因此,大力宣传普及冠心病二级预防用药如: $\beta$ 受体阻滞剂的合理使用也是非常必要的。此外,吸烟严重影响

冠心病 PCI 患者的康复<sup>[3]</sup>,本研究的数据显示,在强化健康教育等有效方式的劝导下,患者戒烟情况比较好。调查显示,在冠心病 PCI 患者需要调整用药的占比为 18.5%。

综上所述,冠心病 PCI 患者现况并不理想,加强冠心病二级预防规范用药和建立家庭医生服务模式下的规范化管理非常重要。除年龄、性别和家族史外,冠心病的主要危险因素为高血压病、血脂异常、糖尿病、过量饮酒、吸烟、肥胖以及社会环境等,这些因素增加了冠心病 PCI 患者致残致死的风险。通过社区家庭医生进行积极、专业的健康干预和指导,使得患者提高用药依从性和自我管理能力,降低再次心肌梗死的发生率和死亡率,改善患者的生活质量。

### 参考文献

- [1] 陈伟伟,高润霖,刘力生,等.中国心血管病报告 2016[J].中国循环杂志,2017,32(6):521-530.
- [2] 虞琪,王长谦.冠心病二级预防的规范化药物治疗[J].中华全科医师杂志,2015,14(11):820-822.
- [3] 陆丕能,孙宁玲,陆肇,等.吸烟量与冠心病关系的病例对照研究[J].中华流行病学杂志,2002,23(4):297-300.

(上接第 253 页)

- [6] 曾照芳,安琳.新型中医专家系统的构建及应用研究[J].辽宁中医药大学学报,2009,(10):9-10.
- [7] 洪芳,何建成,曹雪滨.人工神经网络在中医证候研究中的应用现状与趋势[J].辽宁中医杂志,2013,(01):13-15.
- [8] 孙贵香,袁肇凯.人工神经网络在中医证候研究中的应用[J].中华中医药学刊,2007,(07):1450-1452.
- [9] 陈曦.深度学习(DL)算法在中医药领域的应用探索[A].中国中医药信息研究会.第一届中国中医药信息大会论文集[C].中国中医药信息研究会,2014:4.

- [10] 郭丽丽,丁世飞.深度学习研究进展[J].计算机科学,2015,(05):28-33.
- [11] 陈菊,严小英,裴敬,等.中医临床辨证论治辅助诊疗系统[J].世界科学技术-中医药现代化,2015,(12):2436-2442.
- [12] 魏鲁霞,彭翔.中医辨证施治信息化的方法探讨[J].辽宁中医药大学学报,2008,(01):151-153.
- [13] 杨殿兴,彭明德,林红,等.辨证论治数字化研究思路[J].中国中医药现代远程教育,2003,(06):32-34.
- [14] 张建明,詹智财,成科扬,等.深度学习的研究与发展[J].江苏大学学报(自然科学版),2015,(02):191-200.